HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 44 Issue 3
2015

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 44 Issue 3
June 2015

A Peer Reviewed Journal
Published Bimonthly by the
Faculty of Science of Hacettepe University

Abstracted/Indexed in

SCI-EXP, Journal Citation Reports, Mathematical Reviews,
Zentralblatt MATH, Current Index to Statistics,
Statistical Theory & Method Abstracts,
SCOPUS, Tübitak-Ulakbim.

ISSN 1303 5010

This Journal is typeset using LaTeX.

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Cilt 44 Sayı 3 (2015)

ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2015 SAYI : 44 - 3 AY : Haziran
YAYIN SAHİBİNİN ADI : H. Ü. Fen Fakültesi Dekanlığı adına Prof. Dr. Bekir Salih
SORUMLU YAZI İŞL. MD. ADI : Prof. Dr. Yücel Tıraş
YAYIN İDARE MERKEZİ ADRESİ : H. Ü. Fen Fakültesi Dekanlığı
YAYIN İDARE MERKEZİ TEL. : 0 312 297 68 50
YAYININ TÜRÜ : Yaygın
BASIMCININ ADI : Hacettepe Üniversitesi Hastaneleri Basımevi.
BASIMCININ ADRESİ : 06100 Sıhhıye, ANKARA.
BASIMCININ TEL. : 0 312 305 1020
BASIM TARİHİ - YERİ : - ANKARA

Hacettepe Journal of Mathematics and Statistics

A Bimonthly Publication – Volume 44 Issue 3 (2015)

ISSN 1303 – 5010

Published by Hacettepe University
Faculty of Science

CONTENTS

Mathematics

Statistics

MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 513 – 519

Generalized statistical convergence and some sequence spaces in 2-normed spaces

Cemal Belen * and Mustafa Yildirim †

Abstract

In this work, we first define the concepts of A-statistical convergence and $A^{\mathcal{I}}$-statistical convergence in a 2-normed space and present an example to show the importance of generalized form of convergence through an ideal. We then introduce some new sequence spaces in a 2-Banach space and examine some inclusion relations between these spaces.

2000 AMS Classification: 40A35, 40C05, 40G15.

Keywords: Statistical convergence, Ideal convergence, 2-normed space, Modulus function, Sequence space.

Received 03/02/2012 : Accepted 24/02/2014 Doi : 10.15672/HJMS.2015449421

1. Introduction

The idea of statistical convergence was first introduced by Fast [6] and also independently by Buck [2] and Schoenberg [22] for real and complex sequences, but the rapid developments started after the papers of Šalát [18], Fridy [8] and Connor [3].

Let $K \subseteq \mathbb{N}$ and $K_n = \{k \leq n : k \in K\}$. Then the natural density of K is defined by $\delta(K) = \lim_n n^{-1} |K_n|$ if the limit exists, where $|K_n|$ denotes the cardinality of K_n.

The number sequence $x = (x_k)$ is said to be statistically convergent to the number L provided that for every $\varepsilon > 0$ the set $K(\varepsilon) := \{k \in \mathbb{N} : |x_k - L| \geq \varepsilon\}$ has natural density zero. In this case we write $st - \lim x = L$.

Let X, Y be two sequence spaces and $A = (a_{nk})$ be an infinite matrix. If for each $x \in X$ the series $A_n(x) = \sum_{k=1}^{\infty} a_{nk} x_k$ converges for all n and the sequence $Ax = (A_n(x)) \in Y$, then we say that A maps X into Y. By (X, Y) we denote the set of all matrices which maps X into Y, and in addition if the limit is preserved then we denote the class of such matrices by $(X, Y)_{reg}$. A matrix A is called regular if $A \in (c, c)_{reg}$, where c denotes the space of all convergent sequences.

*Ordu University, Faculty of Education, 52200 Ordu, Turkey
Email: cbelen52@gmail.com

†Department of Mathematics, Faculty of Science, Cumhuriyet University, 58140 Sivas, Turkey
Email:yildirim@cumhuriyet.edu.tr

The well-known Silverman-Toeplitz theorem asserts that A is regular if and only if
(R_1) $\|A\| = \sup_n \sum_k |a_{nk}| < \infty$;
(R_2) $\lim_n a_{nk} = 0$, for each k;
(R_3) $\lim_n \sum_k |a_{nk}| = 1$.

Following Freedman and Sember [7], we say that a set $K \subset \mathbb{N}$ has A-density if

$$\delta_A(K) = \lim_n \sum_{k \in K} a_{nk}$$

exists, where $A = (a_{nk})$ is nonnegative regular matrix.

The idea of statistical convergence was extended to A-statistical convergence by Connor [3] and also independently by Kolk [12]. A sequence x is said to be A-statistically convergent to L if $\delta_A(K(\varepsilon)) = 0$ for every $\varepsilon > 0$. In this case we write $st_A - \lim x = L$.

Let $X \neq \emptyset$. A class $\mathcal{I} \subset 2^X$ of subsets of X is said to be an ideal in X provided; **(i)** $\emptyset \in \mathcal{I}$; **(ii)** $A, B \in \mathcal{I}$ implies $A \cup B \in \mathcal{I}$; **(iii)** $A \in \mathcal{I}$, $B \subset A$ implies $B \in \mathcal{I}$. $\mathcal{I}$ is called a nontrivial ideal if $X \notin \mathcal{I}$, and a nontrivial ideal $\mathcal{I}$ in X is called admissible if $\{x\} \in \mathcal{I}$ for each $x \in X$.

Let $\mathcal{I} \subset 2^{\mathbb{N}}$ be a nontrivial ideal. Then the sequence $x = (x_k)$ of real numbers is said to be ideal convergent or $\mathcal{I}$-convergent to a number L if for each $\varepsilon > 0$ the set $\{k \in \mathbb{N} : |x_k - L| \geq \varepsilon\} \in \mathcal{I}$ (see [15]).

Note that if $\mathcal{I}$ is an admissible ideal in $\mathbb{N}$, then usual converges implies $\mathcal{I}$-convergence.

If we take $\mathcal{I}$=$\mathcal{I}_f$, the ideal of all finite subsets of $\mathbb{N}$, then $\mathcal{I}_f-$convergence coincides with usual convergence. We also note that the ideals $\mathcal{I}_\delta = \{B \subset \mathbb{N} : \delta(E) = 0\}$ and $\mathcal{I}_{\delta_A} = \{B \subset \mathbb{N} : \delta_A(B) = 0\}$ are admissible ideals in $\mathbb{N}$, also $\mathcal{I}_\delta$-convergence and $\mathcal{I}_{\delta_A}$-convergence coincide with statistical convergence and A-statistical convergence respectively.

Savaş et al. (see [21]) have generalized A -statistical convergence by using ideals. Let $A = (a_{nk})$ be a nonnegative regular matrix. A sequence $x = (x_k)$ is said to be $A^{\mathcal{I}}$-statistically convergent (or $S_A(\mathcal{I})$ -convergent) to L if for any $\varepsilon > 0$ and $\delta > 0$,

$$\left\{ n \in \mathbb{N} : \sum_{k \in K(\varepsilon)} a_{nk} \geq \delta \right\} \in \mathcal{I}.$$

In this case we shall write $S_A(\mathcal{I}) - \lim x = L$.

Note that if we take $\mathcal{I}$=$\mathcal{I}_f$, then $A^{\mathcal{I}}$-statistical convergence coincides with A-statistical convergence. Furthermore, the choice of $\mathcal{I}$=$\mathcal{I}_f$ and $A = C_1$, the Cesàro matrix of order one, give us $\mathcal{I}$-statistical convergence introduced in [5] and [20].

Let X be a real vector space of dimension d, where $2 \leq d < \infty$. A 2-norm on X is a function $\|.,.\| : X \times X \to \mathbb{R}$ which satisfies **(i)** $\|x, y\| = 0$ if and only if x and y are linearly dependent; **(ii)** $\|x, y\| = \|y, x\|$; **(iii)** $\|\alpha x, y\| = |\alpha| \|x, y\|$, $\alpha \in \mathbb{R}$; **(iv)** $\|x, y + z\| \leq \|x, y\| + \|x, z\|$. The pair $(X, \|.,.\|)$ is then called a 2-normed space [9]. As an example of a 2-normed space we may take $X = \mathbb{R}^2$ being equipped with the 2-norm $\|x, y\| :=$ the area of parellelogram spanned by the vectors x and y, which may be given explicitly by the formula

$$\|x, y\| = |x_1 y_2 - x_2 y_1|, \quad x = (x_1, x_2),\ y = (y_1, y_2). \tag{1.1}$$

Recall that $(X, \|.,.\|)$ is a 2-Banach space if every Cauchy sequence in X is convergent to some x in X.

The concept of statistical convergence in 2-normed spaces has been introduced and examined by Gürdal and Pehlivan [10]. Let (x_n) be a sequence in 2-normed space $(X, \|.,.\|)$. The sequence (x_n) is said to be statistically convergent to L if for every $\varepsilon > 0$

$$\lim_n \frac{1}{n} |\{n : \|x_n - L, z\| \geq \varepsilon\}| = 0$$

for each nonzero z in X. In this case we write $st-\lim_n \|x_n, z\| = \|L, z\|$.

Finally, we recall that a modulus f is a function from $[0,\infty)$ to $[0,\infty)$ such that **(i)** $f(x)=0$ if and only if $x=0$; **(ii)** $f(x+y)\le f(x)+f(y)$ for all $x\ge 0$ and $y\ge 0$; **(iv)** f is increasing and (iv) f is continuous from the right at 0.

2. $A^{\mathcal{I}}$-statistical convergence in 2-normed spaces

In this section we introduce the concepts of A-statistical convergence and $A^{\mathcal{I}}$-statistical convergence in a 2-normed space when $A=(a_{nk})$ is a nonnegative regular matrix and $\mathcal{I}$ is an admissible ideal of $\mathbb{N}$.

2.1. Definition. Let (x_k) be a sequence in 2-normed space $(X, \|.,.\|)$. Then (x_k) is said to be A-statistically convergent to L if for every $\varepsilon>0$

$$\lim_n \sum_{k:\|x_k-L,z\|\ge\varepsilon} a_{nk}=0$$

for each nonzero z in X, in other words, (x_k) is said to be A-statistically convergent to L provided that $\delta_A(\{k\in\mathbb{N}:\|x_k-L,z\|\ge\varepsilon\})=0$ for every $\varepsilon>0$ and each nonzero z in X. In this case we write $st_A-\lim_k \|x_k,z\|=\|L,z\|$.

We remark that if we take $A=C_1$ in Definition 2.1, then A-statistical convergence coincides with the concept of statistical convergence introduced in [10].

Now we introduce the concept of $A^{\mathcal{I}}$-statistical convergence in a 2-normed space.

2.2. Definition. A sequence (x_k) in 2-normed space $(X, \|.,.\|)$ is said to be $A^{\mathcal{I}}$-statistically convergent to L provided that for every $\varepsilon>0$ and $\delta>0$

$$\left\{n\in\mathbb{N}: \sum_{k:\|x_k-L,z\|\ge\varepsilon} a_{nk}\ge\delta\right\}\in\mathcal{I}$$

for each nonzero z in X. In this case we write $S_A(\mathcal{I})-\lim_k \|x_k,z\|=\|L,z\|$.

We shall denote the space of all A-statistically convergent and $A^{\mathcal{I}}$-statistically convergent sequences in a 2-normed space $(X,\|.,.\|)$ by $S_A(\|.,.\|)$ and $S_A(\mathcal{I},\|.,.\|)$, respectively. It is clear that if $\mathcal{I}=\mathcal{I}_f$, then the space $S_A(\mathcal{I},\|.,.\|)$ is reduced to $S_A(\|.,.\|)$.

Example. Let $X=\mathbb{R}^2$ be equipped with the 2-norm by the formula (1.1). Let $\mathcal{I}\subset 2^{\mathbb{N}}$ be an admissible ideal, $C=\{p_1<p_2<\ldots\}\in\mathcal{I}$ be an infinite set and define the matrix $A=(a_{nk})$ and the sequence (x_k) by

$$a_{nk}=\begin{cases}1 & ;\text{if } n=p_i,\ (i\in\mathbb{N}),\ k=2p_i\\ 1 & ;\text{if } n\ne p_i,\ k=2n+1\\ 0 & ;\text{otherwise.}\end{cases}$$

and

$$x_k=\begin{cases}(0,k) & ;\text{ if } k \text{ is even}\\ (0,0) & ;\text{ otherwise}\end{cases}$$

respectively. Also let $L=(0,0)$ and $z=(z_1,z_2)$. If $z_1=0$ then

$$\{k:\|x_k-L,z\|\ge\varepsilon\}=\emptyset$$

for each z in X. Then $\delta_A(\{k\in\mathbb{N}:\|x_k-L,z\|\ge\varepsilon\})=0$. Hence we have $z_1\ne 0$. For each $\varepsilon>0$

$$\{k:\|x_k-L,z\|\ge\varepsilon\}\overset{\text{if } k \text{ is even}}{=}\left\{k:k\ge\frac{\varepsilon}{|z_1|}\right\},$$

hence for each $\delta > 0$ we obtain

$$\left\{ n \in \mathbb{N} : \sum_{k:\|x_k - L, z\| \geq \varepsilon} a_{nk} \geq \delta \right\} = \{n \in \mathbb{N} : n = p_i\} = C \in \mathcal{I}.$$

This means that $S_A(\mathcal{I}) - \lim_k \|x_k, z\| = \|(0,0), z\|$, but $st_A - \lim_k \|x_k, z\| \neq \|(0,0), z\|$ since

$$\lim_n \sum_{k:\|x_k - L, z\| \geq \varepsilon} a_{nk} = 1 \neq 0.$$

This example also shows that $A^{\mathcal{I}}$-statistical convergence is more general than A-statistical convergence in a 2-normed space.

3. Some New Sequence Spaces

Following the study of Maddox [16], who introduced the notion of strongly Cesàro summability with respect to a modulus, several authors used modulus function to construct some new sequence spaces by using different methods of summability. For instance, see [4], [19] and [1]. Also in [11, 13, 14, 17] some new sequence spaces are defined in a Banach space by means of sequence of modulus functions $\mathcal{F} = (f_k)$.

In this section, we introduce some new sequence spaces in a 2-Banach space by using sequence of modulus functions and ideals. We further examine the inclusion relations between these sequence spaces.

Let $A = (a_{nk})$ be a nonnegative regular matrix, $\mathcal{I}$ be an admissible ideal of $\mathbb{N}$ and let $p = (p_k)$ be a bounded sequence of positive real numbers. By $s(2-X)$ we denote the space of all sequences defined over $(X, \|.,.\|)$. Throughout the paper $\mathcal{F} = (f_k)$ is assumed to be a sequence of modulus functions such that $\lim_{t\to 0^+} \sup_k f_k(t) = 0$ and further let $(X, \|.,.\|)$ be a 2-Banach space. Now we define the following sequence space:

$$w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|) = \Big\{ x \in s(2-X) : \big\{ n \in \mathbb{N} : \sum_k a_{nk} \left[f_k\left(\|x_k - L, z\| \right) \right]^{p_k} \geq \delta \big\} \in \mathcal{I} \text{ for each } \delta > 0 \text{ and } z \in X, \text{ for some } L \in X \Big\}.$$

If $x \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$ then x is said to be strongly $(A, \mathcal{F}, \|.,.\|)$-summable to $L \in X$.

Note that if $0 < p_k \leq \sup_k p_k =: H$, $D := \max(1, 2^{H-1})$, then

$$|a_k + b_k|^{p_k} \leq D\left\{ |a_k|^{p_k} + |b_k|^{p_k} \right\} \tag{3.1}$$

for all k and $a_k, b_k \in \mathbb{C}$.

3.1. Theorem. *$w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$ is a linear space.*

Proof. *Assume that the sequences x and y are strongly $(A, \mathcal{F}, \|.,.\|)$-summable to L and L', respectively and let $\alpha, \beta \in \mathbb{C}$. By using the definitions of modulus function and 2-norm and also from (3.1), we have*

$$\sum_{k=1}^{\infty} a_{nk} \left[f_k\left(\|(\alpha x_k + \beta y_k) - (\alpha L + \beta L'), z\| \right) \right]^{p_k} \leq DM_\alpha^H \sum_{k=1}^{\infty} a_{nk} \left[f_k\left(\|x_k - L, z\| \right) \right]^{p_k} + DM_\beta^H \sum_{k=1}^{\infty} a_{nk} \left[f_k\left(\|y_k - L, z\| \right) \right]^{p_k}$$

where M_α and M_β are positive numbers such that $|\alpha| \leq M_\alpha$ and $|\beta| \leq M_\beta$. From the last inequality, we conclude that $\alpha x + \beta y \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$.

If we take $f_k(t) = t$ for all k and t, then the space $w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$ is reduced to

$$w^{\mathcal{I}}(A, p, \|.,.\|) = \Big\{x \in s(2-X) : \big\{n \in \mathbb{N} : \sum_k a_{nk} (\|x_k - L, z\|)^{p_k} \geq \delta\big\} \in \mathcal{I} \text{ for each } \delta > 0 \text{ and } z \in X, \text{ for some } L \in X\Big\}.$$

If $x \in w^{\mathcal{I}}(A, p, \|.,.\|)$ then we say that x is strongly $(A, \|.,.\|)$-summable to $L \in X$.

3.2. Lemma. *Let f be any modulus function and $0 < \delta < 1$. Then for each $t \geq \delta$ we have $f(t) \leq 2f(1)\delta^{-1}t$ [16].*

3.3. Theorem. *If x is strongly $(A, \|.,.\|)$-summable to L then x is strongly $(A, \mathcal{F}, \|.,.\|)$-summable to L, i.e. the inclusion*

$$w^{\mathcal{I}}(A, p, \|.,.\|) \subset w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$$

holds.

Proof. *Let $x = (x_k) \in w^{\mathcal{I}}(A, p, \|.,.\|)$. Since a modulus function is continuous at $t = 0$ from the right and $\lim_{t\to 0^+} \sup_k f_k(t) = 0$, then for any $\varepsilon > 0$ we can choose $0 < \delta < 1$ such that for every t with $0 \leq t \leq \delta$, we have $f_k(t) < \varepsilon$ $(k \in \mathbb{N})$. Then, from Lemma 3.2, we have*

$$\begin{aligned}\sum_{k=1}^{\infty} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} &= \sum_{k:\|x_k - L, z\| \leq \delta} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \\ &+ \sum_{k:\|x_k - L, z\| > \delta} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \\ &\leq \max\left(\varepsilon^{\inf p_k}, \varepsilon^{\sup p_k}\right) \sum_{k=1}^{\infty} a_{nk} \\ &+ \max(M_1, M_2) \sum_{k=1}^{\infty} a_{nk} (\|x_k - L, z\|)^{p_k}\end{aligned}$$

where $M_1 = \left(2\sup f_k(1)\delta^{-1}\right)^{\inf p_k}$ and $M_2 = \left(2\sup f_k(1)\delta^{-1}\right)^{\sup p_k}$. Let $M := \max(M_1, M_2)$ and $N := \max\left(\varepsilon^{\inf p_k}, \varepsilon^{\sup p_k}\right)$. Now by considering the inequality $\sum_k a_{nk} \leq \|A\|$ for each $n \in \mathbb{N}$, choose a $\sigma > 0$ such that $\sigma - N\|A\| > 0$. Then we obtain

$$\begin{aligned}&\left\{n \in \mathbb{N} : \sum_k a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \geq \sigma\right\} \\ &\quad \subset \left\{n \in \mathbb{N} : \sum_k a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \geq \frac{\sigma - N\|A\|}{M}\right\}\end{aligned}$$

From the assumption we conclude that $x \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$.

3.4. Theorem. *Let $\mathcal{F} = (f_k)$ be the sequence of modulus functions such that $\lim_{t\to\infty} \inf_k \frac{f_k(t)}{t} > 0$. Then $w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|) \subset w^{\mathcal{I}}(A, p, \|.,.\|)$.*

Proof. *Let $x \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$. If $\lim_{t\to\infty} \inf_k \frac{f_k(t)}{t} > 0$ then there exists a $c > 0$ such that $f_k(t) > ct$ for every $t > 0$ and for all $k \in \mathbb{N}$. Thus, for each $\delta > 0$ we have*

$$\begin{aligned}&\left\{n \in \mathbb{N} : \sum_{k=1}^{\infty} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \geq \delta\right\} \\ &\quad \supset \left\{n \in \mathbb{N} : \min\left(c^{\inf p_k}, c^{\sup p_k}\right) \sum_{k=1}^{\infty} a_{nk} (\|x_k - L, z\|)^{p_k} \geq \delta\right\}.\end{aligned}$$

Hence $x \in w^{\mathcal{I}}(A, p, \|.,.\|)$ *and this compltes the proof of theorem.*

Finally, we establish the relations betwen the spaces $S_A(\mathcal{I}, \|.,.\|)$ and $w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$.

3.5. Theorem. *Let* $\mathcal{F} = (f_k)$ *be a sequence of modulus functions such that* $\inf_k f_k(t) > 0$. *Then* $w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|) \subset S_A(\mathcal{I}, \|.,.\|)$.

Proof. *Let* $x \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$ *and* $\varepsilon > 0$. *If* $\inf_k f_k(t) > 0$ *then there exists* $c > 0$ *such that* $f_k(\varepsilon) > c$ *for all* k. *If we write* $K(\varepsilon) = \{k : \|x_k - L, z\| \geq \varepsilon\}$, *then*

$$\sum_{k=1}^{\infty} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \geq \min\left(c^{\inf p_k}, c^{\sup p_k}\right) \sum_{k \in K(\varepsilon)} a_{nk}.$$

Let $C := \min\left(c^{\inf p_k}, c^{\sup p_k}\right)$. *Thus we have*

$$\left\{n \in \mathbb{N} : \sum_{k \in K(\varepsilon)} a_{nk} \geq \delta\right\} \subset \left\{n \in \mathbb{N} : \sum_{k=1}^{\infty} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \geq \frac{\delta}{C}\right\}$$

for all $\delta > 0$. *Since the set on the right-hand of the above inclusion belongs to* $\mathcal{I}$, *we conclude that* $x \in S_A(\mathcal{I}, \|.,.\|)$. *This completes the proof.*

3.6. Theorem. *Let* $\mathcal{F} = (f_k)$ *be a sequence of modulus functions such that* $\sup_t \sup_k f_k(t) > 0$. *Then* $S_A(\mathcal{I}, \|.,.\|) \subset w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$.

Proof. *Let* $x \in S_A(\mathcal{I}, \|.,.\|)$ *and* $h(t) := \sup_k f_k(t)$, $M := \sup_t h(t)$. *Then for every* $\varepsilon > 0$, *we have*

$$\begin{aligned}
\sum_{k=1}^{\infty} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} &= \sum_{k:\|x_k - L, z\| \geq \varepsilon} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \\
&+ \sum_{k:\|x_k - L, z\| < \varepsilon} a_{nk} [f_k(\|x_k - L, z\|)]^{p_k} \\
&\leq \max\left(M^{\inf p_k}, M^{\sup p_k}\right) \sum_{k:\|x_k - L, z\| \geq \varepsilon} a_{nk} \\
&\quad + h(\varepsilon) \sum_{k:\|x_k - L, z\| < \varepsilon} a_{nk} \\
&\leq M_0 \sum_{k:\|x_k - L, z\| \geq \varepsilon} a_{nk} + \varepsilon_1 \|A\|,
\end{aligned}$$

where $M_0 = \max\left(M^{\inf p_k}, M^{\sup p_k}\right)$ *and* ε_1 *is a positive number such that* $h(\varepsilon) < \varepsilon_1$, *which can be obtained from the condition* $\lim_{t \to 0^+} h(t) = 0$. *Hence, from the last inequality we obtain that* $x \in w^{\mathcal{I}}(A, \mathcal{F}, p, \|.,.\|)$.

Acknowledgment. The authors were supported by the Scientific Research Project Fund of Cumhuriyet University under the project number F334.

References

[1] T. Bilgin, Lacunary strong A-convergence with respect to a sequence of modulus functions, Appl. Math. Comput. 151 (2004) 595-600.

[2] R.C. Buck, Generalized asymptotic density, Amer. J. Math. 75 (1953) 335–346.

[3] J. Connor, The statistical and strong p-Cesaro convergence of sequences, Analysis 8 (1988) 47-63.

[4] J. Connor, On strong matrix summability with respect to a modulus and statistical convergence, Canad. Math. Bull. 32 (1989) 194-198.
[5] P. Das, E. Savaş, S. Kr. Ghosal, On generalizations of certain summability methods using ideals, Appl. Math Lett. 24(9) (2011) 1509-1514.
[6] H. Fast, Sur la convergence statistique, Colloq. Math. 2 (1951) 241-244.
[7] A.R. Freedman, J.J. Sember, Densities and summability, Pacific J. Math. 95 (1981) 293-305.
[8] J.A. Fridy, On statistical convergence, Analysis 5 (1985) 301-313.
[9] S. Gähler, 2-Metrische Räume und ihre Topologische Struktur, Math.Nachr., 26 (1963), 115-148.
[10] M. Gürdal, S. Pehlivan, Statistical convergence in 2 -normed spaces, Southeast Asian Bull. Math. 33(2) (2009) 257-264.
[11] E. Kolk, The statistical convergence in Banach spaces, Tartu Ül. Toimetised 928 (1991) 41-52.
[12] E. Kolk, Matrix summability of statistically convergent sequences, Analysis 13 (1993), 77-83.
[13] E. Kolk, Inclusion theorems for some sequence spaces defined by a sequence of moduli, Tartu Ül. Toimetised 970 (1994) 65-72.
[14] E. Kolk, Inclusion relations between the statistical convergence and strong summability, Acta Comment. Univ. Tartu. Math. 2 (1998) 39–54.
[15] P. Kostyrko, T. Šalát, W. Wilczynski, $\mathcal{I}-$ convergence, Real Anal. Exchange 26 (2000/2001) 669-686.
[16] I.J. Maddox, Some sequence spaces defined by a modulus , Math. Proc. Cambridge Philos Soc. 100(1) (1986) 161-166.
[17] S. Pehlivan, B. Fisher, Lacunary strong convergence with respect to a sequence of modulus functions, Comment Math. Univ. Carolinae 36 (1995) 69-76.
[18] T. Šalát, On statistically convergent sequences of real numbers, Math Slovaca 30 (1980) 139-150.
[19] E. Savaş, Some sequence spaces and statistical convergence, Int. J. Math. and Math. Sci. 29(5) (2002), 303-306.
[20] E. Savaş, P. Das, A generalized statistical convergence via ideals, Appl. Math Lett. 24 (2011), 826-830.
[21] E. Savaş, P. Das, S. Dutta, A note on strong matrix summability via ideals, Appl. Math Lett. 25(4) (2012) 733-738.
[22] I.J. Schoenberg, The integrability of certain functions and related summability methods, Amer. Math. Mon. 66 (1959) 361–375.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 521 – 533

n-coherent rings in terms of complexes

Lu Bo * and Liu Zhongkui †

Abstract

The aim of this paper is to investigate n-coherent rings using complexes. To this end, the concepts of n-injective complexes and n-flat complexes are introduced and studied.

2000 AMS Classification: 16E05, 18G35, 18G15.

Keywords: n-Presented complex, n-Injective complex, n-Flat complex, n-Coherent ring.

Received 20/06/2012 : Accepted 26/02/2014 Doi : 10.15672/HJMS.2015449422

1. Introduction. The notion of coherent rings was first appear in Chase's paper [[3]] without being mentioned by name. The term coherent was first used by Bourbaki in [[2]]. Since then, coherent rings have became a vigorously active area of research, see [[13]].

Coherent rings have been characterized in various ways using modules by many authors such as Chase, Cheatham, Ding, Stone, Stenström and Vasconcelos (see [[3, 4, 8, 14, 16]]). For example, a ring R is left coherent if and only if the direct product of any flat right R-modules is flat if and only if the direct limit of FP-injective left R-modules is FP-injective [[3, 14]]. In [[4], Theorem 1], Cheatham and Stone characterized coherent rings using the notion of character module as follows:
The following statements are equivalent:
(1) R is a left coherent ring;
(2) A left R-module M is injective if and only if M^+ is flat;
(2) A left R-module M is injective if and only if M^{++} is injective;
(2) A right R-module M is flat if and only if M^{++} is flat.

The homological theory of complexes of modules has been studied by many authors such as Christensen, Enochs, Foxby, Garcßa Rozas, Holm, Liu and Wang. Several characterizations of coherent rings also have been done in various ways using complexes. For instance, a ring R is right coherent if and only if any complex of left R-modules has a flat

---

*College of Mathematics and Computer Science, Northwest University for Nationalities Lanzhou 730124, Gansu, China.
Email: lubo55@126.com Corresponding Author.
†Department of Mathematics, Northwest Normal University Lanzhou 730070, Gansu, China.
This research was supported by National Natural Science Foundation of China (No.11201376, 11201377, 11261050).

preenvelope [[12], Theorem 5.2.2]; a ring R is right coherent if and only if the direct limit of FP-injective complexes of left R-modules is FP-injective [[17], Proposition 2.30].

The concept of n-coherent rings was introduced by Costa in [[7]]. In [[1]], Bennis introduced the notion of n-$\mathscr{X}$-coherent rings and gave some characterizations of it using n-$\mathscr{X}$-injective and n-$\mathscr{X}$-flat modules for a class of R-modules $\mathscr{X}$.

Motivated by the above work, the object of this paper is to characterize left n-coherent rings using complexes. To this end, we firstly introduce and study n-injective and n-flat complexes for a fixed positive integer n. We show the following results as our main results in this note (cf. Theorem 4.11).

1.1. Theorem. *Let R be a ring and n a fixed positive integer. Then the following are equivalent*:
(1) *R is left n-coherent*;
(2) *Every direct product of n-flat complexes of right R-modules is n-flat*;
(3) *Every direct limit of n-injective complexes of left R-modules is n-injective*;
(4) $\mathrm{Ext}^n(A, \varinjlim C^i) \cong \varinjlim \mathrm{Ext}^n(A, C^i)$ *for every n-presented complex A of left R-modules and direct system $\{C^i\}_{i\in I}$ of complexes of left R-modules*;
(5) $\overline{\mathrm{Tor}}_n(\prod_{\alpha\in I} D^\alpha, A) \cong \prod_{\alpha\in I} \overline{\mathrm{Tor}}_n(D^\alpha, A)$ *for any family $\{D^\alpha\}_{\alpha\in\Lambda}$ of complexes and any n-presented complex A of left R-modules*;
(6) *A complex C of left R-modules is n-injective if and only if C^+ is n-flat*;
(7) *A complex C of left R-modules is n-injective if and only if C^{++} is n-injective*;
(8) *A complex C of right R-modules is n-flat if and only if C^{++} is n-flat*;
(9) *For any ring S,* $\underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^n(A, B), D) \cong \overline{\mathrm{Tor}}_n(\underline{\mathrm{Hom}}(B, D), A)$ *for any n-presented complex A of left R-modules, any complex B of (R,S)-bimodules, any injective complex D of right S-modules.*

The paper is organized as follows:

In section 2 of this article, some notations are given.

In section 3, some isomorphisms are established which will be used to prove the main results of this paper.

In section 4, we firstly introduce and study n-injective and n-flat complexes for a fixed positive integer n. We give various equivalent conditions for a ring to be left n-coherent using n-injective and n-flat complexes.

2. Preliminaries. Throughout this paper, R denotes a ring with unity, R-Mod denotes the category of R-modules and $\mathscr{C}(R)$ denotes the abelian category of complexes of R-modules. A complex

$$\cdots \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} C_{-1} \xrightarrow{\delta_{-1}} \cdots$$

of R-modules will be denoted by (C, δ) or C.

We will use superscripts to distinguish complexes. So if $\{C^i\}_{i\in I}$ is a family of complexes, C^i will be

$$\cdots \xrightarrow{\delta_2} C_1^i \xrightarrow{\delta_1} C_0^i \xrightarrow{\delta_0} C_{-1}^i \xrightarrow{\delta_{-1}} \cdots.$$

Given a left R-module M, we use $D^m(M)$ to denote the complex

$$\cdots \longrightarrow 0 \longrightarrow M \xrightarrow{id} M \longrightarrow 0 \longrightarrow \cdots$$

with M in the mth and $(m-1)$th positions and set $\overline{M} = D^0(M)$. We also use $S^m(M)$ to denote the complex with M in the mth place and 0 in the other places and set $\underline{M} = S^0(M)$.

Given a complex C and an integer m, $\sum^m C$ denotes the complex such that $(\sum^m C)_l = C_{(l-m)}$, and whose boundary operators are $(-1)^m\delta_{l-m}$. The mth homology module of C

is the module $\mathrm{H}_m(C) = \mathrm{Z}_m(C)/\mathrm{B}_m(C)$ where $\mathrm{Z}_m(C) = \mathrm{Ker}(\delta_m^C)$ and $\mathrm{B}_m(C) = \mathrm{Im}(\delta_{m+1}^C)$. We set $\mathrm{H}_m(C) = \mathrm{H}^{-m}(C)$.

Let C be a complex of left R-modules (resp., of right R-modules), and let D be a complex of left R-modules. We will denote by $\mathrm{Hom}^{\cdot}(C, D)$ (resp., $C \otimes^{\cdot} D$) the usual homomorphism complex (resp., tensor product) of the complexes C and D.

Given two complexes C and D, let $\underline{\mathrm{Hom}}(C, D) = \mathrm{Z}(\mathrm{Hom}^{\cdot}(C, D))$. We then see that $\underline{\mathrm{Hom}}(C, D)$ can be made into a complex with $\mathrm{Hom}(C, D)_m$ the abelian group of morphisms from C to $\sum^{-m} D$ and with boundary operator given by $f \in \mathrm{Hom}(C, D)_m$, then $\delta_m(f) : C \to \sum^{-(m-1)} D$ with $\delta_m(f)_l = (-1)^m \delta^D f_l$ for any $l \in \mathbb{Z}$. For any complex C, $C^+ = \underline{\mathrm{Hom}}(C, \overline{\mathbb{Q}/\mathbb{Z}})$. Let C be a complex of right R-modules and D be a complex of left R-modules. We define $C \overline{\otimes} D$ to be $\frac{(C \otimes^{\cdot} D)}{\mathrm{B}(C \otimes^{\cdot} D)}$. Then with the maps

$$\frac{(C \otimes^{\cdot} D)_m}{\mathrm{B}_m(C \otimes^{\cdot} D)} \to \frac{(C \otimes^{\cdot} D)_{m-1}}{\mathrm{B}_{m-1}(C \otimes^{\cdot} D)}, \ x \otimes y \longmapsto \delta^C(x) \otimes y,$$

where $x \otimes y$ is used to denote the coset in $\frac{(C \otimes^{\cdot} D)_m}{\mathrm{B}_m(C \otimes^{\cdot} D)}$, we get a complex. We note that the new functor $\underline{\mathrm{Hom}}(C, D)$ will have right derived functors whose values will be complexes. These values should certainly be denoted $\underline{\mathrm{Ext}}^i(C, D)$. It is not hard to see that $\underline{\mathrm{Ext}}^i(C, D)$ is the complex

$$\cdots \to \mathrm{Ext}^i(C, \Sigma^{-(m+1)} D) \to \mathrm{Ext}^i(C, \Sigma^{-m} D) \to \mathrm{Ext}^i(C, \Sigma^{-(m-1)} D) \to \cdots$$

with boundary operator induced by the boundary operator of D. For a complex C of left R-modules we have two functors $-\overline{\otimes} C : \mathscr{C}_R \to \mathscr{C}_{\mathbb{Z}}$ and $\underline{\mathrm{Hom}}(C, -) : {}_R\mathscr{C} \to \mathscr{C}_{\mathbb{Z}}$, where $\mathscr{C}_R$ (resp., ${}_R\mathscr{C}$) denotes the category of complexes of right R-modules (resp., left R-modules). Since $-\overline{\otimes} C : \mathscr{C}_R \to \mathscr{C}_{\mathbb{Z}}$ is a right exact functor, we can construct left derived functors, which we denote by $\overline{\mathrm{Tor}}_1(-, C)$.

3. *n*-Presented complexes and some isomorphisms. In this section, we first introduce and study the concept of n-presented complexes. Moreover, some isomorphisms which are used to prove the following results are shown.

3.1. Definition ([[10]]). A complex C is called finitely generated if, in the case where we can write C $= \sum_{i \in I} D^i$ with $D^i \in \mathscr{C}(R)$ subcomplexes of C, there exists a finite subset $J \subseteq I$ such that C $= \sum_{i \in J} D^i$.

A complex C is called finitely presented if C is finitely generated and for every exact sequence of complexes $0 \to K \to L \to C \to 0$ with L finitely generated, K is also finitely generated.

3.2. Lemma ([[10]]). *A complex C is finitely generated if and only if C is bounded and C_m is finitely generated in R-*Mod *for all $m \in \mathbb{Z}$.*

*A complex C is finitely presented if and only if C is bounded and C_m is finitely presented in R-*Mod *for all $m \in \mathbb{Z}$.*

It is clear that we have the following results:

3.3. Lemma. *Let $0 \to A \to B \to C \to 0$ be a short exact sequence of complexes. Then the following statements hold:*
(1) *If A is finitely generated and B is finitely presented, then C is finitely presented*;
(2) *If A and C are finitely presented, then so is B*;
(3) *If R is left coherent ring, and B, C are finitely presented, then so is C.*

3.4. Lemma. *Let C be a complex. Then the following statements are equivalent:*
(1) *C is finitely presented*;
(2) *There exists an exact sequence $0 \to L \to P \to C \to 0$ of complexes, where P is finitely generated projective, and L is finitely generated*;

(3) *There exists an exact sequence* $P^1 \to P^0 \to C \to 0$ *of complexes, where* P^0, P^1 *are finitely generated projective, and* P^0_m, P^1_m *are free for all* $m \in \mathbb{Z}$.

An R-module M is called n-presented if it has a finite n-presentation, i.e., there is an exact sequence

$$F_n \to F_{n-1} \to \ldots \to F_1 \to F_0 \to M \to 0$$

in which each F_i is finitely generated free.

Now, we extend the notion of n-presented modules to that of complexes and characterize such complexes.

3.5. Definition. Let $n \geq 0$ be an integer. A complex C is said to be n-presented if there is an exact sequence $P^n \to P^{n-1} \to \cdots \to P^1 \to P^0 \to C \to 0$ of complexes, where P^i is finitely generated projective, and P^i_m is free for $i = 0, 1, \cdots, n$ and all $m \in \mathbb{Z}$.

3.6. Remark. (1) A complex C is n-presented if and only if C is bounded and C_m is n-presented in R-Mod for all $m \in \mathbb{Z}$;
(2) A complex C is n-presented if and only if there is an exact sequence of complexes

$$0 \to K^n \to P^{n-1} \to \cdots \to P^1 \to P^0 \to C \to 0$$

where P^i is finitely generated projective, P^i_m is free for $i = 0, 1, \cdots, n-1$ and all $m \in \mathbb{Z}$, K^n is finitely generated;
(3) A complex C is n-presented ($n \geq 1$) if and only if there is an exact sequence of complexes

$$0 \to K \to P \to C \to 0,$$

where K is $(n-1)$-presented and P is finitely generated projective.

3.7. Lemma. *Let* $n \geq 1$ *be an integer and* $0 \to K \to P \to C \to 0$ *an exact sequence of complexes. Then*

(1) *If* P *is n-presented and* K *is* $(n-1)$*-presented, then* C *is n-presented*;
(2) *If* K *and* C *are n-presented, then so is* P;
(3) *If* C *is n-presented and* P *is* $(n-1)$*-presented, then* K *is* $(n-1)$*-presented.*

Proof. It is similar to the proof of [[13], Theorem 2.1.2] by Remark 3.6 (1). □

Let I be a set. An R-module M is called I-graded if there exists a family $\{M_i\}_{i\in I}$ of submodules of M such that $M = \bigoplus_{i\in I} M_i$. A $\mathbb{Z}$-graded module is simply called a graded module. General background about graded modules can be found in [[6]].

3.8. Lemma. *Let* $\{C^i\}_{i\in I}$ *be a family of complexes,* D *a finitely generated complex. Then* $\underline{\mathrm{Hom}}(D, \bigoplus_{i\in I} C^i) \cong \bigoplus_{i\in I} \underline{\mathrm{Hom}}(D, C^i)$ *as complexes.*

Proof. Firstly,

$$\alpha : \bigoplus_{i\in I} \mathrm{Hom}^{\cdot}(D, C^i) \to \mathrm{Hom}^{\cdot}(D, \bigoplus_{i\in I} C^i)$$

is an isomorphism by $x = (x^i)_{i\in I} \longmapsto \sum_{i\in I} \mathrm{Hom}^{\cdot}(D, \varepsilon^i)(x^i) = \sum_{i\in I} \varepsilon^i x^i$, where $x = (x^i)_{i\in I} \in (\bigoplus_{i\in I} \mathrm{Hom}^{\cdot}(D, C^i))_l = \bigoplus_{i\in I}(\mathrm{Hom}^{\cdot}(D, C^i))_l$ with $x^i \in \mathrm{Hom}^{\cdot}(D, C^i)_l$ and $\varepsilon^j : C^j \longmapsto \bigoplus_{i\in I} C^i$ is the natural embedding (see [[6],Proposition 2.5.16]).

Secondly, we will show that $\underline{\mathrm{Hom}}(D, \bigoplus_{i\in I} C^i) \cong \bigoplus_{i\in I} \underline{\mathrm{Hom}}(D, C^i)$. We define a morphism

$$\gamma = \alpha|_{\bigoplus\limits_{i\in I}\underline{\mathrm{Hom}}(D,C^i)} : \bigoplus_{i\in I}\underline{\mathrm{Hom}}(D,C^i) \to \underline{\mathrm{Hom}}(D,\bigoplus_{i\in I}C^i).$$

Then γ is a graded isomorphism of graded modules with degree$\gamma = 0$. On the other hand, for any $(x^i)_{i\in I} \in (\bigoplus\limits_{i\in I}\underline{\mathrm{Hom}}(D,C^i))_l$,

$$\gamma\delta^{\bigoplus\limits_{i\in I}\underline{\mathrm{Hom}}(D,C^i)}(x^i)_{i\in I} = \alpha\delta^{\bigoplus\limits_{i\in I}\underline{\mathrm{Hom}}(D,C^i)}(x^i)_{i\in I} = \alpha(\delta^{\underline{\mathrm{Hom}}(D,C^i)}(x^i))_{i\in I} =$$

$$\sum_{i\in I}\mathrm{Hom}(D,\varepsilon^i)\delta^{\underline{\mathrm{Hom}}(D,C^i)}(x^i) = \sum_{i\in I}\varepsilon^i(-1)^l\delta^{C^i}(x^i) = (-1)^l\sum_{i\in I}\varepsilon^i\delta^{C^i}(x^i),$$

and

$$\delta^{\underline{\mathrm{Hom}}(D,\bigoplus\limits_{i\in I}C^i)}\gamma(x^i)_{i\in I} = \delta^{\underline{\mathrm{Hom}}(D,\bigoplus\limits_{i\in I}C^i)}\alpha(x^i)_{i\in I} = \delta^{\underline{\mathrm{Hom}}(D,\bigoplus\limits_{i\in I}C^i)}(\sum_{i\in I}\varepsilon^i x^i)$$

$$= (-1)^l\delta^{\bigoplus\limits_{i\in I}C^i}(\sum_{i\in I}\varepsilon^i x^i) = (-1)^l\sum_{i\in I}\delta^{\bigoplus\limits_{i\in I}C^i}\varepsilon^i x^i = (-1)^l\sum_{i\in I}\varepsilon^i\delta^{C^i}(x^i).$$

Thus γ is an isomorphism of complexes, and hence $\underline{\mathrm{Hom}}(D,\bigoplus\limits_{i\in I}C^i) \cong \bigoplus\limits_{i\in I}\underline{\mathrm{Hom}}(D,C^i)$. □

3.9. Lemma. *Let $\{C^i\}_{i\in I}$ be any direct system of complexes. Then a finitely generated complex D is finitely presented if and only if* $\mathrm{Hom}(D,\varinjlim C^i) \cong \varinjlim \mathrm{Hom}(D,C^i)$.

Proof. (⇒) It follows from Stenström [[15], Chap. V, Proposition 3.4].

(⇐) Let $\{M_i\}_{i\in I}$ be a family of R-modules. Then $\sum^{-n}\overline{M_i}$ is a complex for all $n\in\mathbb{Z}$ and $i\in I$. Hence $\mathrm{Hom}(D,\varinjlim\sum^{-n}\overline{M_i}) \cong \varinjlim \mathrm{Hom}(D,\sum^{-n}\overline{M_i})$ for all $n\in\mathbb{Z}$, which implies that

$$\underline{\mathrm{Hom}}(D,\varinjlim\overline{M_i}) \cong \varinjlim\underline{\mathrm{Hom}}(D,\overline{M_i}).$$

Since $\underline{\mathrm{Hom}}(D,\varinjlim\overline{M_i}) \cong \underline{\mathrm{Hom}}(D,\overline{\varinjlim M_i}) \cong \mathrm{Hom}_R(D,\varinjlim M_i)$ and $\underline{\mathrm{Hom}}(D,\overline{M_i}) \cong \mathrm{Hom}_R(D,M_i)$, we have that $\mathrm{Hom}_R(D^k,\varinjlim M_i) \cong \varinjlim\mathrm{Hom}_R(D^k,M_i)$ for all $k\in\mathbb{Z}$, then D^k is a finitely presented R-module. Therefore, D is finitely presented. □

3.10. Lemma. *Let $\{C^i\}_{i\in I}$ be a family of complexes, D a finitely presented complex. Then $D\overline{\otimes}\prod\limits_{i\in I}C^i \cong \prod\limits_{i\in I}(D\overline{\otimes}C^i)$ as complexes.*

Proof. Firstly,

$$\alpha : D\otimes^{\cdot}\prod_{i\in I}C^i \longrightarrow \prod_{i\in I}(D\otimes^{\cdot}C^i)$$

is an isomorphism by $x \longmapsto ((D\otimes^{\cdot}\pi^i)(x))_{i\in I}$, where $x = d\otimes c \in (D\otimes^{\cdot}\prod\limits_{i\in I}C^i)_l$ and $\pi^j : \prod\limits_{i\in I}C^i \longrightarrow C^j$ is the natural projection (see [[6], Proposition 2.5.17]).

Secondly, we will show that $D\overline{\otimes}\prod\limits_{i\in I}C^i \cong \prod\limits_{i\in I}(D\overline{\otimes}C^i)$. Since we have the following commutative diagram:

$$\begin{array}{ccccc}
(D\otimes^{\cdot}\prod_{i\in I}C^i)_l & \longrightarrow & \frac{(D\otimes^{\cdot}\prod_{i\in I}C^i)_l}{B_l(D\otimes^{\cdot}\prod_{i\in I}C^i)} & \longrightarrow & 0\\
\alpha_l\downarrow & & \beta_l\downarrow & & \\
(\prod_{i\in I}D\otimes^{\cdot}C^i)_l & \longrightarrow & \frac{(D\otimes^{\cdot}\prod_{i\in I}C^i)_l}{B_l(\prod_{i\in I}D\otimes^{\cdot}C^i)} & \longrightarrow & 0,
\end{array}$$

where $\beta : \frac{(D\otimes^{\cdot}\prod_{i\in I}C^i)_l}{B_l(D\otimes^{\cdot}\prod_{i\in I}C^i)} \longrightarrow \frac{(D\otimes^{\cdot}\prod_{i\in I}C^i)_l}{B_l(\prod_{i\in I}D\otimes^{\cdot}C^i)}$ is given by the assignment

$$d\otimes c + B(D\otimes^{\cdot}\prod_{i\in I}C^i) \longrightarrow \alpha(d\otimes c) + B(\prod_{i\in I}D\otimes^{\cdot}C^i)$$

for any $d\otimes c\in(D\otimes^{\cdot}\prod_{i\in I}C^i)_l$. Thus β is a graded isomorphism of graded modules with degree 0. Moreover,

$$\beta\delta^{D\overline{\otimes}\prod_{i\in I}C^i}(d\otimes c + B(D\otimes^{\cdot}\prod_{i\in I}C^i))$$

$$= \beta(\delta^D(d)\otimes c) = \alpha(\delta^D(d)\otimes c) = (\delta^D(d)\otimes\pi^i(c))_{i\in I}$$

and

$$\delta^{\prod_{i\in I}(D\overline{\otimes}C^i)}\beta(d\otimes c + B(D\otimes^{\cdot}\prod_{i\in I}C^i))$$

$$= \delta^{\prod_{i\in I}(D\overline{\otimes}C^i)}(\alpha(d\otimes c) + B(D\otimes^{\cdot}\prod_{i\in I}C^i))$$

$$= \delta^{\prod_{i\in I}(D\overline{\otimes}C^i)}\alpha(d\otimes c) = (\delta^{D\overline{\otimes}C^i}\alpha(d\otimes c))_{i\in I} = (\delta^D(d)\otimes\pi^i(c))_{i\in I}.$$

Therefore, β is an isomorphism of complexes. □

3.11. Lemma. *Let $n\geq 1$ be an integer, D an n-presented complex and $\{C^i\}_{i\in I}$ a direct system of complexes. Then* $\mathrm{Ext}^{n-1}(D,\varinjlim C^i)\cong\varinjlim\mathrm{Ext}^{n-1}(D,C^i)$.

Proof. We do an induction on n. If $n=1$, then the result follows from Lemma 3.9.

Let $n=2$ and D be an 2-presented complex. Then there exists an exact sequence of complexes $0\to L\to P\to D\to 0$ with P finitely generated projective and L finitely presented. Thus there is a commutative diagram with exact rows:

$$\begin{array}{ccccccc}
\mathrm{Hom}(P,\varinjlim C^i) & \longrightarrow & \mathrm{Hom}(L,\varinjlim C^i) & \longrightarrow & \mathrm{Ext}^1(D,\varinjlim C^i) & \longrightarrow & 0\\
\cong\downarrow & & \cong\downarrow & & \downarrow & & \\
\varinjlim\mathrm{Hom}(P,C^i) & \longrightarrow & \varinjlim\mathrm{Hom}(L,C^i) & \longrightarrow & \varinjlim\mathrm{Ext}^1(D,C^i) & \longrightarrow & 0.
\end{array}$$

Since $\mathrm{Hom}(P,\varinjlim C^i)\cong\varinjlim\mathrm{Hom}(P,C^i)$ and $\mathrm{Hom}(L,\varinjlim C^i)\cong\varinjlim\mathrm{Hom}(L,C^i)$ by Lemma 3.9, we have $\mathrm{Ext}^1(D,\varinjlim C^i)\cong\varinjlim\mathrm{Ext}^1(D,C^i)$.

If $n>2$, then it follows from the standard homological method. Therefore, $\mathrm{Ext}^{n-1}(D,\varinjlim C^i)\cong\varinjlim\mathrm{Ext}^{n-1}(D,C^i)$. □

3.12. Lemma. *Let $n \geq 1$ be an integer, D an n-presented complex and $\{N^\alpha\}_{\alpha\in I}$ a family of complexes. Then* $\overline{\mathrm{Tor}}_{n-1}(\prod_{\alpha\in I} N^\alpha, D) \cong \prod_{\alpha\in I} \overline{\mathrm{Tor}}_{n-1}(N^\alpha, D)$.

Proof. We do an induction on n. If $n = 1$, then the result follows from Lemma 3.10.

Let $n = 2$ and D be an 2-presented complex. Then there exists an exact sequence of complexes $0 \to L \to P \to D \to 0$ with P finitely generated projective and L finitely presented. Thus there is a commutative diagram with exact rows:

$$\begin{array}{ccccccc}
0 & \longrightarrow & \overline{\mathrm{Tor}}_1(\prod_{\alpha\in I} N^\alpha, D) & \longrightarrow & (\prod_{\alpha\in I} N^\alpha)\overline{\otimes}L & \longrightarrow & (\prod_{\alpha\in I} N^\alpha)\overline{\otimes}P \\
 & & \downarrow & & \cong\downarrow & & \cong\downarrow \\
0 & \longrightarrow & \prod_{\alpha\in I} \overline{\mathrm{Tor}}_1(N^\alpha, D) & \longrightarrow & \prod_{\alpha\in I}(N^\alpha\overline{\otimes}L) & \longrightarrow & \prod_{\alpha\in I}(N^\alpha\overline{\otimes}P).
\end{array}$$

Since $(\prod_{\alpha\in I} N^\alpha)\overline{\otimes}L \cong \prod_{\alpha\in I}(N^\alpha\overline{\otimes}L)$ and $(\prod_{\alpha\in I} N^\alpha)\overline{\otimes}P \cong \prod_{\alpha\in I}(N^\alpha\overline{\otimes}P)$ by Lemma 3.10, we have $\overline{\mathrm{Tor}}_1(\prod_{\alpha\in I} N^\alpha, D) \cong \prod_{\alpha\in I} \overline{\mathrm{Tor}}_1(N^\alpha, D)$.

If $n > 2$, then it follows from the standard homological method. Therefore, $\overline{\mathrm{Tor}}_{n-1}(\prod_{\alpha\in I} N^\alpha, D) \cong \prod_{\alpha\in I} \overline{\mathrm{Tor}}_{n-1}(N^\alpha, D)$. □

3.13. Lemma ([[12]])**.** *Let R and S be rings, L a complex of right S-modules, K a complex of (R,S)-bimodules and P a complex of left R-modules. Suppose that P is finitely presented and L is injective as complexes of right S-modules. Then $\underline{\mathrm{Hom}}(K, L)\overline{\otimes}P \cong \underline{\mathrm{Hom}}(\underline{\mathrm{Hom}}(P, K), L)$ as complexes. This isomorphism is functorial in P, K and L.*

3.14. Lemma. (1) *Let R and S be rings, n a fixed positive integer, A an n-presented complex of left R-modules, B a complex of (R,S)-bimodules, C an injective complex of right S-modules. Then $\underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^{n-1}(A, B), C) \cong \overline{\mathrm{Tor}}_{n-1}(\underline{\mathrm{Hom}}(B, C), A)$.*
(2) *Let R and S be rings, n a fixed positive integer, A a complex of left R-modules, B a complex of right (R,S)-bimodules, C an injective complex of right S-modules. Then $\underline{\mathrm{Ext}}^n(A, \underline{\mathrm{Hom}}(B, C) \cong \underline{\mathrm{Hom}}(\overline{\mathrm{Tor}}_n(B, A), C)$.*

Proof. (1) We do an induction on n. If $n = 1$, then the result follows from Lemma 3.13.

Let $n = 2$ and A be an 2-presented complex. Then there exists an exact sequence of complexes $0 \to K \to P \to A \to 0$ with P finitely generated projective and K finitely presented in $\mathscr{C}(R)$. Thus we have the commutative diagram with exact rows by Lemma 3.13:

$$\begin{array}{ccccccc}
0 & \longrightarrow & \underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^1(A, B), C) & \longrightarrow & \underline{\mathrm{Hom}}(\underline{\mathrm{Hom}}(K, B), C) & \longrightarrow & \underline{\mathrm{Hom}}(\underline{\mathrm{Hom}}(P, B), C) \\
 & & \downarrow & & \cong\downarrow & & \cong\downarrow \\
0 & \longrightarrow & \underline{\mathrm{Tor}}_1(\underline{\mathrm{Hom}}(B, C), A) & \longrightarrow & \underline{\mathrm{Hom}}(B, C) \otimes K & \longrightarrow & \underline{\mathrm{Hom}}(B, C) \otimes P.
\end{array}$$

Hence, $\underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^1(A, B), C) \cong \overline{\mathrm{Tor}}_1(\underline{\mathrm{Hom}}(B, C), A)$.

If $n > 2$, then it follows from the standard homological method. Therefore, $\underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^{n-1}(A, B), C) \cong \overline{\mathrm{Tor}}_{n-1}(\underline{\mathrm{Hom}}(B, C), A)$.

(2) It follows by similar arguments since $\underline{\mathrm{Hom}}(A\overline{\otimes}B, C) \cong \underline{\mathrm{Hom}}(A, \underline{\mathrm{Hom}}(B, C))$ for any complex A, B and C. □

3.15. Remark. It is not hard to see that

$$\underline{\mathrm{Hom}}(D, \prod_{i\in I} C^i) \cong \prod_{i\in I} \underline{\mathrm{Hom}}(D, C^i),$$

$$D\overline{\otimes}\bigoplus_{i\in I}C^i \cong \bigoplus_{i\in I}(D\overline{\otimes}C^i),$$

$$\underline{\mathrm{Ext}}^n(D,\prod_{i\in I}C^i)\cong\prod_{i\in I}\underline{\mathrm{Ext}}^n(D,C^i),$$

and

$$\overline{\mathrm{Tor}}_n(\bigoplus_{\alpha\in I}N^\alpha,D)\cong\bigoplus_{\alpha\in I}\overline{\mathrm{Tor}}_n(N^\alpha,D)$$

for a fixed positive integer n, any complex D and any family $\{C^i\}_{i\in I}$ of complexes by analogy with the proof of the results above.

4. n-Injective complexes and n-flat complexes. In what follows, if A is n-presented, i.e., there is a finite n-presentation $F^n \to F^{n-1} \to \cdots \to F^1 \to F^0 \to A \to 0$, we will write $K^0 = A, K^1$=Ker$(F^0 \to A)$, K^i=Ker$(F^{i-1} \to F^{i-2})$ for $2 \le i \le n$. Clearly, each K^i is $(n-i)$-presented for $0 \le i < n$.

A complex E is said to be FP-injective if $\mathrm{Ext}^1(P,C) = 0$ for any finitely presented complex P, if and only if $\underline{\mathrm{Ext}}^1(P,C) = 0$ for any finitely presented complex P [[17]]. A complex F is flat if and only if $\overline{\mathrm{Tor}}_1(F,C) = 0$ ($\overline{\mathrm{Tor}}_i(F,C) = 0$ for any $i \ge 1$) for any complex C if and only if $\overline{\mathrm{Tor}}_1(F,P) = 0$ ($\overline{\mathrm{Tor}}_i(F,P) = 0$ for any $i \ge 1$) for any finitely presented complex P [[9]].

To characterize left n-coherent rings for a fixed positive integer n, we introduce the following definitions.

4.1. Definition. (1) A complex C is called n-injective if $\underline{\mathrm{Ext}}^n(D,C) = 0$ for any n-presented complex D;
(2) A complex C is called n-flat if $\overline{\mathrm{Tor}}_n(C,D) = 0$ for any n-presented complex D.

4.2. Remark. (1) It is obvious that a complex D is 1-injective (resp. 1-flat) if and only if D is FP-injective (resp. flat); and any n-injective (resp. n-flat) complex is $n+1$-injective (resp. $n+1$-flat). However, the converse is not true in general (see Example 4.12).
(2) It is clear that the class of all n-injective complexes and the class of all n-flat complexes are closed under extensions and summands.
(3) A complex C is n-injective if and only if Ext^n(D,C)=0 for any n-presented complex D.
(4) If R is a left coherent ring and C is an n-flat (resp. n-injective) complex, then $\overline{\mathrm{Tor}}_i(C,F) = 0$ (resp. $\underline{\mathrm{Ext}}^i(F,C) = 0$) for each n-presented complex F and $i \ge 1$.

4.3. Proposition. *Let $\{C^i\}_{i\in I}$ be a family of complexes of R-modules. Then*

(1) $\prod_{i\in I} C^i$ *is n-injective if and only if each C^i is n-injective*;
(2) $\bigoplus_{i\in I} B^i$ *is n-flat if and only if each B^i is n-flat.*

Proof. (1) It follows from the isomorphism $\underline{\mathrm{Ext}}^n(N,\prod_{i\in I}C^i)\cong\prod_{i\in I}\underline{\mathrm{Ext}}^n(N,C^i)$, where N is a complex of R-modules.

(2) It follows from the isomorphism $\overline{\mathrm{Tor}}_n(\bigoplus_{i\in I}B^i,N)\cong\bigoplus_{i\in I}\overline{\mathrm{Tor}}_n(B^i,N)$, where N is a complex of R-modules. □

4.4. Proposition. *Let C be a complex of right R-modules and n a fixed positive integer. Then C is n-flat if and only if C^+ is n-injective.*

Proof. It follows from the isomorphism $\underline{\mathrm{Ext}}^n(D, C^+) \cong \overline{\mathrm{Tor}}_n(C, D)^+$ for any complex D. $\square$

4.5. Lemma. *A complex C is n-injective if and only if, for every n-presentation $F^n \to F^{n-1} \to \cdots \to F^1 \to F^0 \to A \to 0$ of a complex A, every $f : K^n \to C$ can be extended a map $g : F^{n-1} \to C$.*

Proof. We have an exact sequence of complexes $0 \to K^n \to F^{n-1} \to K^{n-1} \to 0$, and an isomorphism $\underline{\mathrm{Ext}}^n(A, C) \cong \underline{\mathrm{Ext}}^1(K^{n-1}, C)$ for any complex C. Therefore, the result follows by definition of n-injective complexes. $\square$

4.6. Lemma. *Consider the commutative diagram with exact rows in $\mathscr{C}(R)$:*

$$\begin{array}{ccccccc}
 & & M_1 & \xrightarrow{f_1} & M_2 & \xrightarrow{f_2} & M_3 & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \varphi_1} & & \downarrow{\scriptstyle \varphi_2} & & \downarrow{\scriptstyle \varphi_3} & & \\
0 & \longrightarrow & N_1 & \xrightarrow{g_1} & N_2 & \xrightarrow{g_2} & N_3. & &
\end{array}$$

Then the following assertions are equivalent:

(a) there exists $\alpha : M_3 \to N_2$ with $\alpha g_2 = \varphi_3$;

(b) there exists $\beta : M_2 \to N_1$ with $f_1\beta = \varphi_1$.

Proof. (b)⇒(a) If $\beta : M_2 \to N_1$ has the given propery, then $g_1\beta f_1 = g_1\varphi_1 = \varphi_2 f_1$, i.e. $(\varphi_2 g_1\beta)f_1 = 0$. Since f_2 is the cokernel of f_1, there exists $\alpha : M_3 \to N_2$ with $\alpha f_2 = \varphi_2 - g_1\beta$. This implies $g_2\alpha f_2 = g_2\varphi_2 - g_2 g_1\beta = g_2\varphi_2 = \varphi_3 f_2$. f_2 being epic we conclude $g_2\alpha = \varphi_3$.

(a)⇒(b) is obtained similarly. $\square$

4.7. Proposition. *The class of all n-injective complexes and the class of all n-flat complexes are closed under pure subcomplexes.*

Proof. Let C_1 be a pure subcomplex of an n-injective complex C. For any finite n-presentation $F^n \to F^{n-1} \to \cdots \to F^1 \to F^0 \to A \to 0$ of A and any map $f : K^n \to C_1$, by Lemma 4.5 and Lemma 4.6, we get the following diagram with exact rows:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & K^n & \xrightarrow{i} & F^{n-1} & \xrightarrow{p} & K^{n-1} & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle f} & \overset{g}{\swarrow} & \downarrow{\scriptstyle k} & \overset{h}{\swarrow} & \downarrow{\scriptstyle l} & & \\
0 & \longrightarrow & C_1 & \xrightarrow{j} & C & \xrightarrow{q} & C/C_1 & \longrightarrow & 0
\end{array}$$

where i and j are inclusion maps. So C_1 is n-injective by Lemma 4.5 again.

Let S be a pure subcomplex of an n-flat complex C. Then the pure exact sequence $0 \to S \to C \to C/S \to 0$ induces the split exact sequence $0 \to (C/S)^+ \to C^+ \to S^+ \to 0$. Thus S^+ is n-injective since C^+ is n-injective by Proposition 4.4. So S is n-flat by Proposition 4.4 again. $\square$

4.8. Lemma. *Let $\{C^i\}_{i\in I}$ be a family of complexes. Then*

(1) $\bigoplus_{i\in I} C^i$ *is a pure subcomplex of* $\prod_{i\in I} C^i$;

(2) $\prod_{i\in I} C^i$ *is a pure subcomplex of* $\prod_{i\in I} (C^i)^{++}$.

Proof. (1) Since for any finitely presented complex P, we have $(\prod_{i\in I} C^i)\overline{\otimes}P \cong \prod_{i\in I}(C^i\overline{\otimes}P)$ by Lemma 3.10. Thus we get the following commutative diagram:

$$\begin{array}{ccccc} & & (\bigoplus_{i\in I} C^i)\overline{\otimes}P & \longrightarrow & (\prod_{i\in I} C^i)\overline{\otimes}P \\ & & \cong\downarrow & & \downarrow\cong \\ 0 & \longrightarrow & \bigoplus_{i\in I}(C^i\overline{\otimes}P) & \longrightarrow & \prod_{i\in I}(C^i\overline{\otimes}P). \end{array}$$

Hence, $\bigoplus_{i\in I} C^i$ is a pure subcomplex of $\prod_{i\in I} C^i$.

(2) It is similar to the proof of (1) since C^i is a pure subcomplex of $(C^i)^{++}$ for each $i \in I$. □

4.9. Lemma. *The following are equivalent for a bounded complex C of right R-modules*:
(1) *C is finitely generated*;
(2) $C\overline{\otimes}\prod_\Lambda A^\lambda \to \prod_\Lambda(C\overline{\otimes}A^\lambda)$ *is an epimorphism for every family $\{A^\lambda\}_\Lambda$ of complexes of left R-modules.*

Proof. (1) ⇒ (2) Let $0 \to K \to F \to C \to 0$ be an exact sequence of complexes with F finitely generated projective. Then we have the following commutative diagram:

$$\begin{array}{ccccccc} K\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & F\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & C\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & 0 \\ \tau_K\downarrow & & \tau_F\downarrow & & \tau_C\downarrow & & \\ \prod_\Lambda(K\overline{\otimes}A^\lambda) & \longrightarrow & \prod_\Lambda(F\overline{\otimes}A^\lambda) & \longrightarrow & \prod_\Lambda(C\overline{\otimes}A^\lambda) & \longrightarrow & 0 \end{array}$$

with exact rows. But τ_F is isomorphism by Lemma 3.10. So τ_C is onto.

(2) ⇒ (1) Since C is bounded, we can assume that C has the following form:

$$\cdots \longrightarrow 0 \longrightarrow C_m \longrightarrow C_{m-1} \longrightarrow \cdots \longrightarrow C_1 \longrightarrow C_0 \longrightarrow 0 \longrightarrow \cdots.$$

It is enough to prove that C_j is finitely generated in R-Mod for $j = 1, \cdots, m$. Let $\{M_i\}_{i\in I}$ be a family of left R-modules. Then

$$(C \otimes^{\cdot} \prod_{i\in I} \underline{M_i})_m = \bigoplus_{t\in\mathbb{Z}} C_t \otimes (\prod_{i\in I} \underline{M_i})_{m-t} = C_m \otimes \prod_{i\in I} M_i$$

and

$$(\prod_{i\in I} C \otimes^{\cdot} \underline{M_i})_m = \prod_{i\in I}\bigoplus_{t\in\mathbb{Z}} C_t \otimes (\underline{M_i})_{m-t} = \prod_{i\in I} C_m \otimes M_i.$$

$$C \otimes^{\cdot} \prod_{i\in I} \underline{M_i} : \quad \cdots \to 0 \xrightarrow{\delta_{m+1}} C_m \otimes \prod_{i\in I} M_i \xrightarrow{\delta_m} \cdots \xrightarrow{\delta_2} C_1 \otimes \prod_{i\in I} M_i \xrightarrow{\delta_1} 0 \to \cdots.$$

$$\prod_{i\in I} C \otimes^{\cdot} \underline{M_i} : \quad \cdots \to 0 \xrightarrow{\sigma_{m+1}} \prod_{i\in I} C_m \otimes M_i \xrightarrow{\sigma_m} \cdots \xrightarrow{\sigma_2} \prod_{i\in I} C_1 \otimes M_i \xrightarrow{\sigma_1} 0 \to \cdots.$$

Hence $C\overline{\otimes}\prod_{i\in I}\underline{M_i}$ and $\prod_{i\in I}(C\overline{\otimes}\underline{M_i})$ have the following form:

$C\overline{\otimes}\prod_{i\in I}\underline{M_i}$:

$$\cdots \to 0 \longrightarrow C_m \otimes \prod_{i\in I} M_i \longrightarrow \frac{C_{m-1}\otimes\prod_{i\in I} M_i}{\mathrm{Im}\delta_m} \longrightarrow \cdots \longrightarrow \frac{C_1\otimes\prod_{i\in I} M_i}{\mathrm{Im}\delta_2} \longrightarrow 0 \to \cdots.$$

$\prod_{i\in I}(C\overline{\otimes}\underline{M_i})$:

$$\cdots \to 0 \longrightarrow \prod_{i\in I}(C_m\otimes M_i) \longrightarrow \frac{\prod_{i\in I}(C_{m-1}\otimes M_i)}{\mathrm{Im}\sigma_m} \longrightarrow \cdots \longrightarrow \frac{\prod_{i\in I}(C_1\otimes M_i)}{\mathrm{Im}\sigma_2} \longrightarrow 0 \to \cdots.$$

Since $C\overline{\otimes}\prod_{i\in I}\underline{M_i} \to \prod_{i\in I}(C\overline{\otimes}\underline{M_i})$ is epic, $C_m\otimes\prod_{i\in I}M_i \to \prod_{i\in I}(C_m\otimes M_i)$ is epic, then C_m is finitely generated in R-Mod by [[11], Lemma 3.2.21]. If we replace the complex $\underline{M_i}$ with $\overline{M_i}$, we have C_{m-1} is finitely generated in R-Mod. If we replace $\underline{M_i}$ with $\cdots \to 0 \to M_0 \to M_1 \to M_2 \to 0 \to \cdots$, we have C_{m-2} is finitely generated in R-Mod. We continue the process, we can get C_j is finitely generated in R-Mod for $j = 1, \cdots, m$ by [[11], Lemma 3.2.21]. □

4.10. Lemma. *The following are equivalent for a bounded complex C of right R-modules*:
(1) *C is finitely presented*;
(2) $C\overline{\otimes}\prod_\Lambda A^\lambda \to \prod_\Lambda(C\overline{\otimes}A^\lambda)$ *is an isomorphism for every family* $\{A^\lambda\}_\Lambda$ *of complexes of left R-modules.*

Proof. (1) ⇒ (2) It follows by Lemma 3.10.

(2) ⇒ (1) C is finitely generated by the Lemma 4.9 above. So let $0 \to K \to F \to C \to 0$ be exact with F finitely generated projective. It now suffices to show that K is finitely generated. But for any Λ, we have a commutative diagram:

$$\begin{array}{ccccccc} K\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & F\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & C\overline{\otimes}\prod_\Lambda A^\lambda & \longrightarrow & 0 \\ \downarrow{\scriptstyle\tau_K} & & \downarrow{\scriptstyle\tau_F} & & \downarrow{\scriptstyle\tau_C} & & \\ \prod_\Lambda(K\overline{\otimes}A^\lambda) & \longrightarrow & \prod_\Lambda(F\overline{\otimes}A^\lambda) & \longrightarrow & \prod_\Lambda(C\overline{\otimes}A^\lambda) & \longrightarrow & 0 \end{array}$$

with exact rows where τ_F and τ_C are isomorphisms. So τ_K is onto and hence K is finitely generated by Lemma 4.9.

A ring R is left coherent if and only if the direct limit of FP-injective complexes of left R-modules is FP-injective [[17]]. Now we will give some characterizations of n-coherent rings using the results above.

4.11. Theorem. *Let R be a ring and n a fixed positive integer. Then the following are equivalent*:
(1) *R is left n-coherent*;
(2) *Every direct product of n-flat complexes of right R-modules is n-flat*;
(3) *Every direct limit of n-injective complexes of left R-modules is n-injective*;
(4) $\mathrm{Ext}^n(A, \varinjlim C^i) \cong \varinjlim \mathrm{Ext}^n(A, C^i)$ *for every n-presented complex A of left R-modules and direct system* $\{C^i\}_{i\in I}$ *of complexes of left R-modules*;
(5) $\overline{\mathrm{Tor}}_n(\prod_{\alpha\in I} D^\alpha, A) \cong \prod_{\alpha\in I}\overline{\mathrm{Tor}}_n(D^\alpha, A)$ *for any family* $\{D^\alpha\}_{\alpha\in\Lambda}$ *of complexes and any n-presented complex A of left R-modules*;
(6) *A complex C of left R-modules is n-injective if and only if C^+ is n-flat*;
(7) *A complex C of left R-modules is n-injective if and only if C^{++} is n-injective*;
(8) *A complex C of right R-modules is n-flat if and only if C^{++} is n-flat*;
(9) *For any ring S,* $\underline{\mathrm{Hom}}(\underline{\mathrm{Ext}}^n(A,B),D) \cong \overline{\mathrm{Tor}}_n(\underline{\mathrm{Hom}}(B,D),A)$ *for any n-presented complex A of left R-modules, any complex B of (R,S)-bimodules, any injective complex D of right S-modules.*

Proof. (1) ⇒ (4) It follows by Lemma 3.11.

(4) ⇒ (3) It is trivial.

(3) ⇒ (1) Let A be an n-presented complex of left R-modules. It is sufficient to show that K^n is finitely presented. Let $\{C^i\}_{i\in I}$ be a family of n-injective complexes of

left R-modules, where I is a directed set. Then $\varinjlim C^i$ is n-injective by (3), and hence $\mathrm{Ext}^1(K^{n-1}, \varinjlim C^i) = \mathrm{Ext}^n(A, \varinjlim C^i) = 0$.

Thus there is a commutative diagram with exact rows:

$$\begin{array}{ccc}
\mathrm{Hom}(K^{n-1}, \varinjlim C^i) & \xrightarrow{f_1} & \varinjlim \mathrm{Hom}(K^{n-1}, C^i) \\
\downarrow & & \downarrow \\
\mathrm{Hom}(F^{n-1}, \varinjlim C^i) & \xrightarrow{f_2} & \varinjlim \mathrm{Hom}(F^{n-1}, C^i) \\
\downarrow & & \downarrow \\
\mathrm{Hom}(K^{n}, \varinjlim C^i) & \xrightarrow{f_3} & \varinjlim \mathrm{Hom}(K^{n}, C^i) \\
\downarrow & & \downarrow \\
0 & & 0
\end{array}.$$

Since both K^{n-1} and F^{n-1} are finitely presented, f_1 and f_2 are isomorphisms by Lemma 3.9. Hence f_3 is an isomorphism. K^n is finitely generated, so K^n is finitely presented by Lemma 3.9. Thus A is $(n+1)$-presented. Therefore, R is left n-coherent.

(1) $\Rightarrow$ (5) It holds by Lemma 3.12.

(5) $\Rightarrow$ (2) It is obvious.

(2) $\Rightarrow$ (1) Let A be an n-presented complex of left R-modules. We will show that K^{n-1} is 2-presented. For any family $\{A^i\}_{i\in I}$ of n-flat complexes of right R-modules, $\prod_{i\in I} A^i$ is an n-flat complex. Thus the exact sequence of complexes $0 \to K^n \to F^{n-1} \to K^{n-1} \to 0$ gives rise to the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & (\prod_{i\in I} A^i)\overline{\otimes} K^n & \longrightarrow & (\prod_{i\in I} A^i)\overline{\otimes} F^{n-1} & \longrightarrow & (\prod_{i\in I} A^i)\overline{\otimes} K^{n-1} & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \phi_1} & & \downarrow{\scriptstyle \phi_2} & & \downarrow{\scriptstyle \phi_3} & & \\
0 & \longrightarrow & \prod_{i\in I}(A^i\overline{\otimes} K^n) & \longrightarrow & \prod_{i\in I}(A^i\overline{\otimes} F^{n-1}) & \longrightarrow & \prod_{i\in I}(A^i\overline{\otimes} K^{n-1}) & \longrightarrow & 0.
\end{array}$$

By Lemma 4.10, ϕ_2 and ϕ_3 are isomorphisms, and hence ϕ_1 is an isomorphism. Thus K^n is finitely presented, and so K^{n-1} is 2-presented, hence A is $n+1$-presented.

(6) $\Rightarrow$ (7) Let C be a complex of left R-modules. If C is n-injective, then C^+ is n-flat by (6), and so C^{++} is n-injective by Proposition 4.4. Conversely, if C^{++} is n-injective, then C is a pure subcomplex of C^{++} (see [[12], Proposition 5.1.4]). So C is n-injective by Proposition 4.7.

(7) $\Rightarrow$ (8) If C is an n-flat complex of right R-modules, then C^+ is an n-injective complex of left R-modules by Proposition 4.4. Hence C^{+++} is n-injective by (7). Thus C^{++} is n-flat by Proposition 4.4. Conversely, if C^{++} is n-flat, then C is n-flat by Proposition 4.7.

(8) $\Rightarrow$ (2) Let $\{C^i\}_{i\in I}$ be a family of n-flat complexes of right R-modules. By Proposition 4.3, $\bigoplus_{i\in I} C^i$ is n-flat, so $(\bigoplus_{i\in I} C^i)^{++} \cong (\prod_{i\in I} C^{i+})^+$ is n-flat by (8). But $\bigoplus_{i\in I}(C^i)^+$ is a pure subcomplex of $\prod_{i\in I}(C^i)^+$ by Lemma 4.8, and so $(\prod_{i\in I}(C^i)^+)^+ \to (\bigoplus_{i\in I}(C^i)^+)^+ \to 0$ splits. Thus $\prod_{i\in I}(C^i)^{++} \cong (\bigoplus_{i\in I}(C^i)^+)^+$ is n-flat. Since $\prod_{i\in I} C^i$ is a pure subcomplex of $\prod_{i\in I}(C^i)^{++}$ by Lemma 4.8, $\prod_{i\in I} C^i$ is n-flat by Proposition 4.7.

(1) $\Rightarrow$ (9) It follows from Lemma 3.14.

(9) ⇒ (6) Let $S = \mathbb{Z}$, $D = \overline{\mathbb{Q}/\mathbb{Z}}$ and $B = C$. Then

$$\overline{\mathrm{Tor}}_n(C^+, A) \cong \underline{\mathrm{Ext}}^n(A, C)^+$$

for all n-presented complexes A of left R-modules by (9), and hence (6) holds. □

4.12. Example. If R is $n+1$-coherent but not n-coherent, then we can form a direct limit $\lim\limits_{\rightarrow} C^i$ of n-injective complexes $\{C^i\}_{i\in I}$, which is not n-injective but is necessary $n+1$-injective; we can also form a direct product $\prod\limits_{\alpha\in I} C^i$ of n-flat complexes $\{C^i\}_{i\in I}$, which is not n-flat but is necessary $n+1$-flat.

Acknowledgements

The authors would like to thank the referees for useful comments and suggestions that improved the paper.

References

[1] D. Bennis, n-$\mathscr{X}$-coherent rings, Int. Electron. J. Algebra. 7 (2010) 128-139.
[2] N. Bourbaki, Algèbre Commutative, Chapitre 1-2, Masson, Paris, 1961.
[3] S. U. Chase, Direct products of modules, Trans. Amer. Math. Soc. 97 (1960) 457-473.
[4] T. J. Cheatham, D. R. Stone, Flat and projective character modules, Proc. Amer. Math. Soc. 81 (1981) 175-177.
[5] J. L. Chen and N. Q. Ding, On n-coherent rings, Comm. Algebra. 24(10) (1996) 3211-3216.
[6] L. W. Christensen, H. B. Foxby and H. Holm, Derived Category Methods in Commutative Algebra, 2011.
[7] D. L. Costa, Parameterizing families of non-noetherian rings, Comm. Algebra 22 (1994) 3997-4011.
[8] N. Q. Ding and L. X. Mao, On a new characterization of coherent rings, Publ. Math. (Debreen), 71 (2007) 67-82
[9] E. E. Enochs, J. R. Garcβa Rozas, Tensor products of complexes, Math. J. Okayama Univ. 39 (1997) 17-39.
[10] E. E. Enochs, J. R. Garcβa Rozas, Flat covers of complexes, J. Algebra 210 (1998) 86-102.
[11] E. E. Enochs and O. M. G. Jenda, Relative Homological Algebra, (Walter de Gruyter, 2000).
[12] J. R. Garcβa Rozas, Covers and envelopes in the category of complexes of modules, Boca Raton London New York Washington, D.C. (1999).
[13] S. Glaz, Commutative Coherent Rings, Lect. Notes Math. 1371, Springer-Verlag, Berlin, 1989. Publishing. (1993) 172-181.
[14] B. Stenström, Coherent rings and FP-injective modules, J. London Math. Soc. 2 (1970) 323-329.
[15] S. Stenström, Rings of Quotients, New York: Springer-Verlag, 1975.
[16] W. V. Vasconcelos, The Rings of Dimension Two, Marcel Dekker, New York, 1976.
[17] Z. P. Wang and Z. K. Liu, FP-injective complexes and FP-injective dimension of complexes, J. Aust. Math. Soc. (2011) 1-25.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 535 – 546

Integral representations and new generating functions of Chebyshev polynomials

Clemente Cesarano *

Abstract

In this paper we use the two-variable Hermite polynomials and their operational rules to derive integral representations of Chebyshev polynomials. The concepts and the formalism of the Hermite polynomials $H_n(x, y)$ are a powerful tool to obtain most of the properties of the Chebyshev polynomials. By using these results, we also show how it is possible to introduce relevant generalizations of these classes of polynomials and we derive for them new identities and integral representations. In particular we state new generating functions for the first and second kind Chebyshev polynomials.

2000 AMS Classification: 33C45, 33D45.

Keywords: Chebyshev polynomials, Integral representations, Generating functions.

Received 27/09/2012 : Accepted 23/12/2013 Doi : 10.15672/HJMS.20154610029

1. Introduction

The Hermite polynomials [1] can be introduced by using the concept and the formalism of the generating function and related operational rules. In the following we recall the main definitions and properties.

1.1. Definition. The two-variable Hermite Polynomials $H_m^{(2)}(x, y)$ of Kampé de Fériet form [2, 3] are defined by the following formula

$$H_m^{(2)}(x, y) = \sum_{n=0}^{\left[\frac{m}{2}\right]} \frac{m!}{n!(m-2n)!} y^n x^{m-2n} \tag{1.1}$$

*Faculty of Engineering, International Telematic University UNINETTUNO Corso Vittorio Emanuele II, 39 – 00186 – Roma (Italy)
Email: c.cesarano@uninettunouniversity.net

We will indicate the two-variable Hermite polynomials of Kampé de Fériet form by using the symbol $H_m(x,y)$ instead than $H_m^{(2)}(x,y)$.

The two-variable Hermite polynomials $H_m(x,y)$ are linked to the ordinary Hermite polynomials by the following relations

$$H_m\left(x,-\frac{1}{2}\right)=He_m(x),$$

where

$$He_m(x)=m!\sum_{r=0}^{\left[\frac{m}{2}\right]}\frac{(-1)^r x^{n-2r}}{r!(n-2r)!2^r}$$

and

$$H_m\left(2x,-1\right)=H_m(x),$$

where

$$H_m(x)=m!\sum_{r=0}^{\left[\frac{m}{2}\right]}\frac{(-1)^r (2x)^{n-2r}}{r!(n-2r)!}$$

and it is also important to note that the Hermite polynomials $H_m(x,y)$ satisfy the relation

$$H_m(x,0)=x^m. \tag{1.2}$$

1.2. Proposition. *The polynomials $H_m(x,y)$ solve the following partial differential equation:*

$$\frac{\partial^2}{\partial x^2}H_m(x,y)=\frac{\partial}{\partial y}H_m(x,y). \tag{1.3}$$

Proof. By deriving, separately with respect to x and to y, in the (1), we obtain

$$\begin{aligned}\frac{\partial}{\partial x}H_m(x,y) &= mH_{m-1}(x,y)\\ \frac{\partial}{\partial y}H_m(x,y) &= H_{m-2}(x,y).\end{aligned}$$

From the first of the above relation, by deriving again with respect to x and by noting the second identity, we end up with the (7). □

The *Proposition 1* help us to derive an important operational rule for the Hermite polynomials $H_m(x,y)$. In fact, by considering the differential equation (7) as linear ordinary in the variable y and by remanding the (6) we can immediately state the following relation:

$$H_m(x,y)=e^{y\frac{\partial^2}{\partial x^2}}x^m. \tag{1.4}$$

The generating function of the above Hermite polynomials can be state in many ways, we have in fact:

1.3. Proposition. *The polynomials $H_m(x,y)$ satisfy the following differential difference equation:*

$$\begin{aligned} \frac{d}{dz} Y_n(z) &= a n Y_{n-1}(z) + b n(n-1) Y_{n-2}(z) \\ Y_n(0) &= \delta_{n,0} \end{aligned} \tag{1.5}$$

where a and b are real numbers.

Proof. By using the generating function method, by putting:

$$G(z;t) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} Y_n(z),$$

with t continuous variable,we can rewrite the (9) in the form

$$\begin{aligned} \frac{d}{dz} G(z;t) &= \left(at + bt^2\right) G(z;t) \\ G(0;t) &= 1 \end{aligned}$$

that is a linear ordinary differential equation and then its solution reads

$$G(z;t) = \exp\left(xt + yt^2\right)$$

where we have putted $az = x$ and $bz = y$. Finally, by exploiting the r.h.s of the previous relation we find the thesis and also the relation linking the Hermite polynomials and their generating function

$$\exp\left(xt + yt^2\right) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} H_n(x,y). \tag{1.6}$$

□

The use of operational identities, may significantly simplify the study of Hermite generating functions and the discovery of new relations, hardly achievable by conventional means.

By remanding that the following identity

$$e^{-\frac{1}{4}\frac{d^2}{dx^2}} (2x)^n = \left(2x - \frac{d}{dx}\right)^n (1) \tag{1.7}$$

is linked to the standard Burchnall identity [4], we can immediately state the following relation.

1.4. Proposition. *The operational definition of the polynomials $H_n(x)$ reads:*

$$e^{-\frac{1}{4}\frac{d^2}{dx^2}} (2x)^n = H_n(x). \tag{1.8}$$

Proof. By exploiting the r.h.s of the (13), we immediately obtain the Burchnall identity

$$\left(2x - \frac{d}{dx}\right)^n = n! \sum_{s=0}^{n} (-1)^s \frac{1}{(n-s)!s!} H_{n-s}(x) \frac{d^s}{dx^s} \tag{1.9}$$

after using the decoupling Weyl identity [4, 5, 6], since the commutator of the operators of l.h.s. is not zero. The derivative operator of the (15) gives a not trivial contribution only in the case $s = 0$ and then we can conclude with

$$\left(2x - \frac{d}{dx}\right)^n (1) = H_n(x)$$

which prove the statement. □

The Burchnall identity can be also inverted to give another important relation for the Hermite polynomials $H_n(x)$. We find in fact:

1.5. Proposition. *The polynomials $H_n(x)$ satisfy the following operational identity:*

$$(1.10) \qquad H_n\left(x + \frac{1}{2}\frac{d}{dx}\right) = \sum_{s=0}^{n} \binom{n}{s} (2x)^{n-s} \frac{d^s}{dx^s}.$$

Proof. By multiplying the l.h.s. of the above relation by $\frac{t^n}{n!}$ and then summing up, we obtain:

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!} H_n\left(x + \frac{1}{2}\frac{d}{dx}\right) = e^{2\left(x+\frac{1}{2}\right)\left(\frac{d}{dx}\right)t - t^2}.$$

By using the Weyl identity, the r.h.s. of the above equation reads:

$$e^{2\left(x+\frac{1}{2}\right)\left(\frac{d}{dx}\right)t - t^2} = e^{2xt} e^{t\frac{d}{dx}}$$

and from which (17) immediately follows, after expanding the r.h.s and by equating the like $t-$powers. □

The previous results can be used to derive some addition and multiplication relations for the Hermite polynomials.

1.6. Proposition. *The polynomials $H_n(x)$ satisfy the following identity, $\forall\ n, m \in N$:*

$$(1.11) \qquad H_{n+m}(x) = \sum_{s=0}^{\min(n,m)} (-2)^s \binom{n}{s} \binom{m}{s} s! H_{n-s}(x) H_{m-s}(x).$$

Proof. By using the *Proposition 3*, we can write:

$$H_{n+m}(x) = \left(2x - \frac{d}{dx}\right)^n \left(2x - \frac{d}{dx}\right)^m = \left(2x - \frac{d}{dx}\right)^n H_m(x)$$

and by exploiting the r.h.s. of the above relation, we find:

$$H_{n+m}(x) = \sum_{s=0}^{n} (-1)^s \binom{n}{s} H_{n-s}(x) \frac{d^s}{dx^s} H_m(x).$$

After noting that the following operational identity holds:

$$\frac{d^s}{dx^s} H_m(x) = \frac{2^s m!}{(m-s)!} H_{m-s}(x)$$

we obtain immediately the statement. □

From the above proposition we can immediately derive as a particular case, the following identity:

$$(1.12) \quad H_{2n}(x) = (-1)^n 2^n (n!)^2 \sum_{s=0}^{n} \frac{(-1)^s \left[H_s(x)\right]^2}{2^s (s!)^2 (n-s)!}.$$

The use of the identity (17), stated in *Proposition 4*, can be exploited to obtain the inverse of relation contained in (24). We have indeed:

1.7. Proposition. *Given the Hermite polynomial $H_n(x)$, the square $\left[H_n(x)\right]^2$ can be written as:*

$$(1.13) \quad H_n(x)H_n(x) = \left[H_n(x)\right]^2 = 2^n (n!)^2 \sum_{s=0}^{n} \frac{H_{2n}(x)}{2^s (s!)^2 (n-s)!}.$$

Proof. We can write:

$$\left[H_n(x)\right]^2 = e^{-\frac{1}{4}\frac{d^2}{dx^2}} \left[H_n\left(x + \frac{1}{2}\frac{d}{dx}\right) H_n\left(x + \frac{1}{2}\frac{d}{dx}\right)\right],$$

by using the relation (17), we find, after manipulating the r.h.s.:

$$\left[H_n(x)\right]^2 = e^{-\frac{1}{4}\frac{d^2}{dx^2}} \left[2^n (n!)^2 \sum_{s=0}^{n} \frac{(2x)^{2n}}{2^s (s!)^2 (n-s)!}\right]$$

and then, from the Burchnall identity (16), the thesis. □

A generalization of the identities stated for the one variable Hermite polynomials can be easily done for the polynomials $H_n(x,y)$.

We have in fact:

1.8. Proposition. *The following identity holds*

$$(1.14) \quad \left(x + 2y\frac{\partial}{\partial x}\right)^n (1) = \sum_{s=0}^{n} (2y)^s \binom{n}{s} H_n(x,y) \frac{\partial^s}{\partial x^s}(1).$$

Proof. By multiplying the l.h.s. of the above equation by $\frac{t^n}{n!}$ and then summing up, we find

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!} \left(x + 2y\frac{\partial}{\partial x}\right)^n = e^{t\left(x + 2y\frac{\partial}{\partial x}\right)}(1).$$

By noting that the commutator of the two operators of the r.h.s. is

$$\left[tx, t2y\frac{\partial}{\partial x}\right] = -2t^2 y$$

we obtain

$$(1.15) \quad \sum_{n=0}^{+\infty} \frac{t^n}{n!} \left(x + 2y\frac{\partial}{\partial x}\right)^n = e^{xt + yt^2} e^{2ty\frac{\partial}{\partial x}}(1).$$

After expanding and manipulating the r.h.s. of the previous relation and by equating the like t powers we find immediately the (28). □

By using the *Proposition 7* and the definition of polynomials $H_n(x,y)$, we can derive a generalization of the Burchnall-type identity

$$(1.16)\qquad e^{y\frac{\partial^2}{\partial x^2}}x^n = \left(x + 2y\frac{\partial}{\partial x}\right)^n$$

and the related inverse

$$(1.17)\qquad H_n\left(x - 2y\frac{\partial}{\partial x}, y\right) = \sum_{s=0}^{n} (-2y)^s \binom{n}{s} x^{n-s}\frac{\partial^s}{\partial x^s}.$$

We can also generalize the multiplication rules obtained for the Hermite polynomials $H_n(x)$, stated in *Proposition 5.*

1.9. Proposition. *Given the Kampé de Fériet Hermite polynomials $H_n(x,y)$, we have*

$$(1.18)\qquad H_{n+m}(x,y) = m!n! \sum_{s=0}^{\min(n,m)} (2y)^s \frac{H_{n-s}(x,y)H_{m-s}(x,y)}{(n-s)!(m-s)!s!}.$$

Proof. By using the relations stated in the (28) and (32), we can write

$$H_{n+m}(x,y) = \left(x + 2y\frac{\partial}{\partial x}\right)^n H_m(x,y)$$

and then

$$(1.19)\qquad H_{n+m}(x,y) = \sum_{s=0}^{n} (2y)^s \binom{n}{s} H_n(x,y)\frac{\partial^s}{\partial x^s} H_m(x,y).$$

By noting that

$$\frac{\partial^s}{\partial x^s}x^m = \frac{m!}{(m-2s)!}x^{m-2s}$$

we obtain

$$\frac{\partial^s}{\partial x^s}H_m(x,y) = \frac{m!}{(m-s)!}H_{m-s}(x,y).$$

After substituting the above relation in the (36) and rearranging the terms we immediately obtain the thesis. □

From the previous results, it also immediately follows:

$$(1.20)\qquad H_n(x,y)H_m(x,y) = n!m! \sum_{s=0}^{\min(n,m)} (-2y)^s \frac{H_{n+m-2s}(x,y)}{(n-s)!(m-s)!s!}.$$

The previous identity and the equation (34) can be easily used to derive the particular case for $n = m$. We have in fact

$$(1.21)\qquad H_{2n}(x,y) = 2^n (n!)^2 \sum_{s=0}^{n} \frac{[H_s(x,y)]^2}{(s)!^2 (n-s)! 2^s}$$

$$(1.22)\qquad [H_n(x,y)]^2 = (-2y)^n (n!)^2 \sum_{s=0}^{n} \frac{(-1)^s H_{2s}(x,y)}{(n-s)!(s!)^2 2^s}.$$

Before concluding this section we want prove two other important relations satisfied by the Hermite polynomials $H_n(x,y)$.

1.10. Proposition. *The Hermite polynomials $H_n(x,y)$ solve the following differential equation:*

$$(1.23)\qquad 2y\frac{\partial^2}{\partial x^2} H_n(x,y) + x\frac{\partial}{\partial x} H_n(x,y) = nH_n(x,y)$$

Proof. By using the results derived from the *Proposition 7*, we can easily write that:

$$\left(x + 2y\frac{\partial}{\partial x}\right) H_n(x,y) = H_{n+1}(x,y)$$

and from the previous recurrence relations:

$$\frac{\partial}{\partial x} H_n(x,y) = nH_{n-1}(x,y)$$

we have

$$\left(x + 2y\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial x}\right) H_n(x,y) = nH_n(x,y)$$

which is the thesis.

□

From this statement can be also derived an important recurrence relation. In fact, by noting that:

$$(1.24)\qquad H_{n+1}(x,y) = xH_n(x,y) + 2y\frac{\partial}{\partial x} H_n(x,y)$$

and then we can conclude with:

$$(1.25)\qquad H_{n+1}(x,y) = xH_n(x,y) + 2nyH_{n-1}(x,y).$$

2. Integral representations of Chebyshev polynomials

In this section we will introduce new representations of Chebyshev polynomials [7, 8, 9, 10, 11], by using the Hermite polynomials and the method of the generating function. Since the second kind Chebyshev polynomials $U_n(x)$ reads

$$(2.1)\qquad U_n(x) = \frac{\sin\left[(n+1)\arccos(x)\right]}{\sqrt{1-x^2}},$$

by exploiting the right hand side of the above relation, we can immediately get the following explicit form

$$(2.2)\qquad U_n(x)=\sum_{k=0}^{\left[\frac{n}{2}\right]}\frac{(-1)^k(n-k)!(2x)^{n-2k}}{k!(n-2k)!}.$$

2.1. Proposition. *The second kind Chebyshev polynomials satisfy the following integral representation* [9]*:*

$$(2.3)\qquad U_n(x)=\frac{1}{n!}\int_0^{+\infty}e^{-t}t^nH_n\left(2x,-\frac{1}{t}\right)dt.$$

Proof. By noting that

$$n!=\int_0^{+\infty}e^{-t}t^n dt$$

we can write

$$(2.4)\qquad (n-k)!=\int_0^{+\infty}e^{-t}t^{n-k}dt.$$

From the explicit form of the Chebyshev polynomials $U_n(x)$, given in the (49), and by recalling the standard form of the two-variable Hermite polynomials:

$$H_n(x,y)=n!\sum_{k=0}^{\left[\frac{n}{2}\right]}\frac{y^kx^{n-2k}}{k!(n-2k)!}$$

we can immediately write:

$$U_n(x)=\int_0^{+\infty}e^{-t}t^n\sum_{k=0}^{\left[\frac{n}{2}\right]}\frac{(-1)^kt^{-k}(2x)^{n-2k}}{k!(n-2k)!}dt$$

and then the thesis. □

By following the same procedure, we can also obtain an analogous integral representation for the Chebyshev polynomials of first kind $T_n(x)$. Since their explicit form is given by:

$$(2.5)\qquad T_n(x)=\frac{n}{2}\sum_{k=0}^{\left[\frac{n}{2}\right]}\frac{(-1)^k(n-k-1)!(2x)^{n-2k}}{k!(n-2k)!},$$

by using the same relations written in the previous proposition, we easily obtain:

$$(2.6)\qquad T_n(x)=\frac{1}{2(n-1)!}\int_0^{+\infty}e^{-t}t^{n-1}H_n\left(2x,-\frac{1}{t}\right)dt.$$

These results can be useful in several physics and engineering problems, for instance in electromagnetic field problems and particle accelerators analysis [12, 13, 14] In the previous Section we have stated some useful operational results regarding the two-variable Hermite polynomials; in particular we have derived their fundamental recurrence relations. These relations can be used to state important results linking the Chebyshev polynomials of the first and second kind [7, 9].

2.2. Theorem. *The Chebyshev polynomials $T_n(x)$ and $U_n(x)$ satisfy the following recurrence relations:*

$$\begin{aligned} \frac{d}{dx}U_n(x) &= nW_{n-1}(x) \\ U_{n+1}(x) &= xW_n(x) - \frac{n}{n+1}W_{n-1}(x) \end{aligned} \tag{2.7}$$

and

$$T_{n+1}(x) = xU_n(x) - U_{n-1}(x) \tag{2.8}$$

where

$$W_n(x) = \frac{2}{(n+1)!}\int_0^{+\infty} e^{-t}t^{n+1}H_n\left(2x, -\frac{1}{t}\right)dt.$$

Proof. The recurrence relations for the standard Hermite polynomials $H_n(x, y)$ stated in the first Section, can be costumed in the form

$$\begin{aligned} \left[(2x) + \left(-\frac{1}{t}\right)\frac{\partial}{\partial x}\right]H_n\left(2x, -\frac{1}{t}\right) &= H_{n+1}\left(2x, -\frac{1}{t}\right) \\ \frac{1}{2}\frac{\partial}{\partial x}H_n\left(2x, -\frac{1}{t}\right) &= nH_{n-1}\left(2x, -\frac{1}{t}\right). \end{aligned} \tag{2.9}$$

From the integral representations stated in the relations (50) and (53), relevant to the Chebyshev polynomials of the first and second kind, and by using the second of the identities written above, we obtain

$$\frac{d}{dx}U_n(x) = \frac{2n}{n!}\int_0^{+\infty} e^{-t}t^nH_{n-1}\left(2x, -\frac{1}{t}\right)dt \tag{2.10}$$

and

$$\frac{d}{dx}T_n(x) = \frac{n}{(n-1)!}\int_0^{+\infty} e^{-t}t^{n-1}H_{n-1}\left(2x, -\frac{1}{t}\right)dt. \tag{2.11}$$

It is easy to note that the above relation gives a link between the polynomials $T_n(x)$ and $U_n(x)$; in fact, since:

$$U_{n-1}(x) = \frac{1}{(n-1)!}\int_0^{+\infty} e^{-t}t^{n-1}H_{n-1}\left(2x, -\frac{1}{t}\right)dt$$

we immediately get:

$$\frac{d}{dx}T_n(x) = nU_{n-1}(x). \tag{2.12}$$

By applying the multiplication operator to the second kind Chebyshev polynomials, stated in the first of the identities (56), we can write

$$U_{n+1}(x) = \frac{1}{(n+1)!}\int_0^{+\infty} e^{-t}t^{n+1}\left[(2x) + \left(-\frac{1}{t}\right)\frac{\partial}{\partial x}\right]H_n\left(2x, -\frac{1}{t}\right)dt$$

that is

$$U_{n+1}(x) = \tag{2.13}$$

$$= x\frac{2}{(n+1)!}\int_0^{+\infty} e^{-t}t^{n+1}H_n\left(2x,-\frac{1}{t}\right)dt - \frac{n}{n+1}\frac{2}{n!}\int_0^{+\infty} e^{-t}t^{n}H_{n-1}\left(2x,-\frac{1}{t}\right)dt.$$

The second member of the r.h.s. of the above relation suggests us to introduce the following polynomials:

$$(2.14)\quad W_n(x) = \frac{2}{(n+1)!}\int_0^{+\infty} e^{-t}t^{n+1}H_n\left(2x,-\frac{1}{t}\right)dt$$

recognized as belonging to the families of the Chebyshev polynomials. Thus, from the relation (57), we have:

$$(2.15)\quad \frac{d}{dx}U_n(x) = nW_{n-1}(x)$$

and, from the identity (60), we get

$$(2.16)\quad U_{n+1}(x) = xW_n(x) - \frac{n}{n+1}W_{n-1}(x).$$

Finally, by using the multiplication operator for the first kind Chebyshev polynomials, we can write

$$(2.17)\quad T_{n+1}(x) = \frac{1}{2n!}\int_0^{+\infty} e^{-t}t^{n}\left[(2x)+\left(-\frac{1}{t}\right)\frac{\partial}{\partial x}\right]H_n\left(2x,-\frac{1}{t}\right)dt$$

and then, after exploiting the r.h.s. of the above relation, we can find

$$(2.18)\quad T_{n+1}(x) = xU_n(x) - U_{n-1}(x)$$

which completely prove the theorem.

□

3. Generating functions

By using the integral representations and the related recurrence relations, stated in the previous Section, for the Chebyshev polynomials of the first and second kind, it is possible to derive a slight different relations linking these polynomials and their generating functions [1, 2, 3, 4, 5, 7, 8, 9, 15].

We note indeed, for the Chebyshev polynomials $U_n(x)$, that by multiplying both sides of equation (50) by ξ^n, $|\xi| < 1$ and by summing up over n, it follows that

$$(3.1)\quad \sum_{n=0}^{+\infty}\xi^n U_n(x) = \int_0^{+\infty} e^{-t}\sum_{n=0}^{+\infty}\frac{(t\xi)^n}{n!}H_n\left(2x,-\frac{1}{t}\right)dt.$$

By recalling the generating function of the polynomials $H_n(x,y)$ stated in the relation (12) and by integrating over t, we end up with

$$(3.2)\quad \sum_{n=0}^{+\infty}\xi^n U_n(x) = \frac{1}{1-2\xi x+\xi^2}.$$

We can now state the related generating function for the first kind Chebyshev polynomials $T_n(x)$ and for the polynomials $W_n(x)$, by using the results proved in the previous theorem.

3.1. Corollary. *Let $x, \xi \in \mathbf{R}$, such that $|x| < 1, |\xi| < 1$; the generating functions of the polynomials $T_n(x)$ and $W_n(x)$ read*

$$\sum_{n=0}^{+\infty} \xi^n T_{n+1}(x) = \frac{x - \xi}{1 - 2\xi x + \xi^2} \tag{3.3}$$

and

$$\sum_{n=0}^{+\infty} (n+1)(n+2\xi^n W_{n+1}(x) = \frac{8(x-\xi)}{(1 - 2\xi x + \xi^2)^3}. \tag{3.4}$$

Proof. By multiplying both sides of the relation (2.8) by ξ^n and by summing up over n, we obtain

$$\sum_{n=0}^{+\infty} \xi^n T_{n+1}(x) = x \sum_{n=0}^{+\infty} \xi^n U_n(x) - \sum_{n=0}^{+\infty} \xi^n U_{n-1}(x)$$

that is

$$\sum_{n=0}^{+\infty} \xi^n T_{n+1}(x) = \frac{x}{1 - 2\xi x + \xi^2} - \frac{\xi}{1 - 2\xi x + \xi^2}$$

which gives the (68).

In the same way, by multiplying both sides of the second relation stated in the (54) by ξ^n and by summing up over n, we get

$$\sum_{n=0}^{+\infty} \xi^n U_{n+1}(x) = x \sum_{n=0}^{+\infty} \xi^n W_n(x) - \sum_{n=0}^{+\infty} \frac{n}{n+1} \xi^n W_{n-1}(x)$$

and then the thesis.

These results allows us to note that the use of integral representations relating Chebyshev and Hermite polynomials are a fairly important tool of analysis allowing the derivation of a wealth of relations between first and second kind Chebyshev polynomials and the Chebyshev-like polynomials $W_n(x)$. In a forthcoming paper, we will deeper investigate other generalizations for these families of polynomials, recognized as Chebyshev polynomials, by using the instruments of integral representations.

□

References

[1] H. W. Srivastava and H.L. Manocha, *A treatise on Generating Functions*, Ellis Horwood Series: Mathematics and its Applications. Ellis Horwood Ltd., Chichester; Halsted Press [John Wiley & Sons, Inc.], New York, 1984.

[2] G. Dattoli, C. Cesarano, *A new family of Hermite polynomials associated to parabolic cylinder functions*, Applied Mathematics and Computation, vol. **141** (1) (2003), pp. 143-149.

[3] G. Dattoli, S. Lorenzutta and C. Cesarano, *Bernstein polynomials and operational methods*, Journal of Computation Analysis and Applications, vol. **8** (4) (2006), pp. 369-377.

[4] C. Cesarano, G.M. Cennamo and L. Placidi, *Operational methods for Hermite polynomials with applications*, WSEAS Transactions on Mathematics, vol. **13** (2014), pp. 925-931.

[5] G. Dattoli, D. Sacchetti and C. Cesarano, *A note on truncated polynomials*, Applied Mathematics and Computation, vol. **134** (2-3) (2002), pp. 595-605.

[6] C. Cesarano, C. Fornaro and L. Vazquez, *A note on a special class of Hermite polynomials*, International Journal of Pure and Applied Mathematics, vol. **98** (2) (2015), pp. 261-273.

[7] C. Cesarano, *Identities and genarating functions on Chebyshev polynomials*, Georgian Mathematical Journal, vol. **19** (2012), pp. 427-440.

[8] C. Cesarano, *Generalization of two-variable Chebyshev and Gegenbauer polynomials*, International Journal of Applied Mathematics & Statistics, vol. **53** (2015), pp. 1-7.

[9] C. Cesarano and C. Fornaro, *Operational identities on generalized two-variable Chebyshev polynomials*, International Journal of Pure and Applied Mathematics, vol. **100** (1) (2015), pp. 59-74.
[10] C. Cesarano, *Generalized Chebyshev polynomials*, Hacettepe Journal of Mathematics and Statistics, vol. **43** (5) (2014), pp. 731-740.
[11] G. Dattoli, D. Sacchetti and C. Cesarano, *A note on Chebyshev polynomials*, Annali dell'Universitá di Ferrara, vol. **47** (1) (2001), pp. 107-115.
[12] C. Cesarano and D. Assante, *A note on Generalized Bessel Functions*, International Journal of Mathematical Models and Methods in Applied Sciences (2013), vol. **7** (6), 625-629.
[13] D. Assante, L. Verolino, *Efficient evaluation of the longitudinal coupling impedance of a plane strip*, Progress In Electromagnetics Research M, **26** (2012), pp. 251-265.
[14] D. Assante, D. Davino, S. Falco, F. Schettino, L. Verolino, *Coupling impedance of a charge traveling in a drift tube*, IEEE Transactions on Magnetics, vol. **41** (5) (2005), pp. 1924-1927.
[15] M. Abramowitz; I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York: Dover, 1965.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 547 – 558

Common fixed point theorems for two pairs of non-self mappings in cone metric spaces

Xianjiu Huang*†, Chuanxi Zhu‡ and Xi Wen§

Abstract

Some common fixed point theorems for two pairs of non-self mappings defined on a closed subset of a metrically convex cone metric space (over the cone which is not necessarily normal) are obtained which generalize earlier results due to Imdad et al. and Jankovic et al.

2000 AMS Classification: 47H10, 54H25.

Keywords: Cone metric spaces, Common fixed point, Non-self mappings.

Received 14/05/2009 : Accepted 28/02/2012 Doi : 10.15672/HJMS.2015449104

1. Introduction and preliminaries

Recently, Huang and Zhang ([14]) generalized the concept of a metric space, replacing the set of real numbers by ordered Banach space and obtained some fixed point theorems for mappings satisfying different contractive conditions. Subsequently, the study of fixed point theorems in such spaces is followed by some other mathematicians, see [1]-[5], [7]-[12], [15]-[18], [20]-[23]. The aim of this paper is to prove some common fixed point theorems for two pairs of non-self mappings on cone metric spaces in which the cone need not be normal. This result generalizes the result of Jankovic et al.([18]).

Consistent with Huang and Zhang ([14]), the following definitions and results will be needed in the sequel.

Let E be a real Banach space. A subset P of E is called a cone if and only if:

(a) P is closed, nonempty and $P \neq \{\theta\}$;

(b) $a, b \in R, a, b \geq 0, x, y \in P$ implies $ax + by \in P$;

(c) $P \cap (-P) = \{\theta\}$.

*Department of Mathematics, Nanchang University, Nanchang, 330031, Jiangxi, P.R.China, Email: xjhuangxwen@163.com

†Corresponding Author.

‡Department of Mathematics, Nanchang University, Nanchang, 330031, Jiangxi, P.R.China, Email: chuanxizhu@126.com

§Department of Computer Sciences, Nanchang University, Nanchang, 330031, Jiangxi, P.R.China, Email: ncuxwen@163.com

Given a cone $P \subset E$, we define a partial ordering $\leq$ with respect to P by $x \leq y$ if and only if $y - x \in P$. A cone P is called normal if there is a number $K > 0$ such that for all $x, y \in E$,

$$\theta \leq x \leq y \text{ implies } \| x \| \leq K \| y \|.$$

The least positive number K satisfying the above inequality is called the normal constant of P, while $x \ll y$ stands for $y - x \in intP$ (interior of P).

1.1. Definition ([14])**.** Let X be a nonempty set. Suppose that the mapping $d : X \times X \to E$ satisfies:

(d1) $\theta \leq d(x, y)$ for all $x, y \in X$ and $d(x, y) = \theta$ if and only if $x = y$;
(d2) $d(x, y) = d(y, x)$ for all $x, y \in X$;
(d3) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.
Then d is called a cone metric on X and (X, d) is called a cone metric space.

The concept of a cone metric space is more general than that of a metric space.

1.2. Definition ([14])**.** Let (X, d) be a cone metric space. We say that $\{x_n\}$ is:

(e) a Cauchy sequence if for every $c \in E$ with $\theta \ll c$, there is an N such that for all $n, m > N, d(x_n, x_m) \ll c$;

(f) a convergent sequence if for every $c \in E$ with $\theta \ll c$, there is an N such that for all $n > N, d(x_n, x) \ll c$ for some fixed $x \in X$.

A cone metric space X is said to be complete if for every Cauchy sequence in X is convergent in X. It is known that if P is normal, then $\{x_n\}$ converges to $x \in X$ if and only if $d(x_n, x) \to \theta$ as $n \to \infty$. It is a Cauchy sequence if and only if $d(x_n, x_m) \to \theta (n, m \to \infty)$.

1.3. Remark ([24])**.** Let E be an ordered Banach (normed) space. Then c is an interior point of P, if and only if $[-c, c]$ is a neighborhood of θ.

1.4. Corollary ([19])**.** *(1) If $a \leq b$ and $b \ll c$, then $a \ll c$.*
Indeed, $c - a = (c - b) + (b - a) \geq c - b$ implies $[-(c - a), c - a] \supseteq [-(c - b), c - b]$.
(2) If $a \ll b$ and $b \ll c$, then $a \ll c$.
Indeed, $c - a = (c - b) + (b - a) \geq c - b$ implies $[-(c - a), c - a] \supseteq [-(c - b), c - b]$.
(3) If $\theta \leq u \ll c$ for each $c \in intP$ then $u = \theta$.

1.5. Remark ([18])**.** If $c \in intP, \theta \leq a_n$ and $a_n \to \theta$, then there exists an n_0 such that for all $n > n_0$ we have $a_n \ll c$.

1.6. Remark ([18])**.** If E is a real Banach space with cone P and if $a \leq ka$ where $a \in P$ and $0 < k < 1$, then $a = \theta$.

We find it convenient to introduce the following definition.

1.7. Definition ([18])**.** Let (X, d) be a complete cone metric space and C be a nonempty closed subset of X, and $f, g : C \to X$. Denote, for$x, y \in C$,

$$M_1^{f,g} = \{d(gx, gy), d(fx, gx), d(fy, gy), \frac{d(fx, gy) + d(fy, gx)}{2}\}. \tag{1.1}$$

Then f is called a generalized g_{M_1}−contractive mapping of C into X if for some $\lambda \in (0, \sqrt{2} - 1)$ there exists $u(x, y) \in M_1^{f,g}$ such that for all $x, y \in C$

$$d(fx, fy) \leq \lambda u(x, y). \tag{1.2}$$

1.8. Definition ([2])**.** Let f and g be self maps of a set X (i.e., $f, g : X \to X$). If $w = fx = gx$ for some x in X, then x is called a coincidence point of f and g, and w is called a point of coincidence of f and g. Self maps f and g are said to be weakly compatible if they commute at their coincidence point; i.e., if $fx = gx$ for some $x \in X$, then $fgx = gfx$.

2. Main results

Recently, Jankovic et al.([18]) proved some fixed point theorems for a pair of non-self mappings defined on a nonempty closed subset of complete metrically convex cone metric spaces with new contractive conditions.

2.1. Theorem ([18])**.** *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ (the boundary of C) such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Suppose that $f, g : C \to X$ are such that f is a generalized $g_{M_1}-$contractive mapping of C into X, and

(i) $\partial C \subseteq gC, fC \cap C \subseteq gC$,
(ii) $gx \in \partial C$ implies that $fx \in C$,
(iii) gC is closed in X.

Then the pair (f, g) has a coincidence point. Moreover, if pair (f, g) is weakly compatible, then f and g have a unique common fixed point.

The purpose of this paper is to extend above theorem for two pairs of non-self mappings in cone metric spaces. We begin with the following definition.

2.2. Definition. Let (X, d) be a complete cone metric space and C be a nonempty closed subset of X, and $F, G, S, T : C \to X$. Denote, for $x, y \in C$,

$$M_1^{F,G,S,T} = \{d(Tx, Sy), d(Tx, Fx), d(Sy, Gy), \frac{d(Tx, Gy) + d(Fx, Sy)}{2}\}. \tag{2.1}$$

Then (F, G) is called a generalized $(T, S)_{M_1}-$contractive mappings pair of C into X if for some $\lambda \in (0, 1)$ there exists $u(x, y) \in M_1^{F,G,S,T}$ such that for all $x, y \in C$

$$d(Fx, Gy) \leq \lambda u(x, y). \tag{2.2}$$

Notice that by setting $G = F = f$ and $T = S = g$ in (2.1), one deduces a slightly generalized form of (1.1).

We state and prove our main result as follows.

2.3. Theorem. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Suppose that $F, G, S, T : C \to X$ are such that (F, G) is a generalized $(T, S)_{M_1}-$ contractive mappings pair of C into X, and

(I) $\partial C \subseteq SC \cap TC, FC \cap C \subseteq SC, GC \cap C \subseteq TC$,
(I) $Tx \in \partial C$ implies that $Fx \in C, Sx \in \partial C$ implies that $Gx \in C$,
(III) SC and TC (or FC and GC) are closed in X.

Then

(IV) (F, T) has a point of coincidence,
(V) (G, S) has a point of coincidence.

Moreover, if (F, T) and (G, S) are weakly compatible pairs, then F, G, S and T have a unique common fixed point.

Proof. Firstly, we proceed to construct two sequences $\{x_n\}$ and $\{y_n\}$ in the following way.

Let $x \in \partial C$ be arbitrary. Then (due to $\partial C \subseteq TC$) there exists a point $x_0 \in C$ such that $x = Tx_0$. Since $Tx \in \partial C \Rightarrow Fx \in C$, one concludes that $Fx_0 \in FC \cap C \subseteq SC$.

Thus, there exist $x_1 \in C$ such that $y_1 = Sx_1 = Fx_0 \in C$. Since $y_1 = Fx_0$ there exists a point $y_2 = Gx_1$ such that

$$d(y_1, y_2) = d(Fx_0, Gx_1).$$

Suppose $y_2 \in C$. Then $y_2 \in GC \cap C \subseteq TC$ which implies that there exists a point $x_2 \in C$ such that $y_2 = Tx_2$. otherwise, if $y_2 \notin C$, then there exists a point $p \in \partial C$ such that

$$d(Sx_1, p) + d(p, y_2) = d(Sx_1, y_2).$$

Since $p \in \partial C \subseteq TC$ there exists a point $x_2 \in C$ with $p = Tx_2$ so that

$$d(Sx_1, Tx_2) + d(Tx_2, y_2) = d(Sx_1, y_2).$$

Let $y_3 = Fx_2$ be such that $d(y_2, y_3) = d(Gx_1, Fx_2)$. Thus, repeating the foregoing arguments, one obtains two sequences $\{x_n\}$ and $\{y_n\}$ such that

(a) $y_{2n} = Gx_{2n-1}, y_{2n+1} = Fx_{2n}$,

(b) $y_{2n} \in C \Rightarrow y_{2n} = Tx_{2n}$ or $y_{2n} \notin C \Rightarrow Tx_{2n} \in \partial C$ and

$$d(Sx_{2n-1}, Tx_{2n}) + d(Tx_{2n}, y_{2n}) = d(Sx_{2n-1}, y_{2n}).$$

(c) $y_{2n+1} \in C \Rightarrow y_{2n+1} = Sx_{2n+1}$ or $y_{2n+1} \notin C \Rightarrow Sx_{2n+1} \in \partial C$ and

$$d(Tx_{2n}, Sx_{2n+1}) + d(Sx_{2n+1}, y_{2n+1}) = d(Tx_{2n}, y_{2n+1}).$$

We denote

$$P_0 = \{Tx_{2i} \in \{Tx_{2n}\} : Tx_{2i} = y_{2i}\},$$
$$P_1 = \{Tx_{2i} \in \{Tx_{2n}\} : Tx_{2i} \neq y_{2i}\},$$
$$Q_0 = \{Sx_{2i+1} \in \{Sx_{2n+1}\} : Sx_{2i+1} = y_{2i+1}\},$$
$$Q_1 = \{Sx_{2i+1} \in \{Sx_{2n+1}\} : Sx_{2i+1} \neq y_{2i+1}\}.$$

Note that $(Tx_{2n}, Sx_{2n+1}) \notin P_1 \times Q_1$, as if $Tx_{2n} \in P_1$, then $y_{2n} \neq Tx_{2n}$ and one infers that $Tx_{2n} \in \partial C$ which implies that $y_{2n+1} = Fx_{2n} \in C$. Hence $y_{2n+1} = Sx_{2n+1} \in Q_0$. Similarly, one can argue that $(Sx_{2n-1}, Tx_{2n}) \notin Q_1 \times P_1$.

Now, we distinguish the following three cases.

Case 1. If $(Tx_{2n}, Sx_{2n+1}) \in P_0 \times Q_0$, then from (2.2)

$$d(Tx_{2n}, Sx_{2n+1}) = d(Fx_{2n}, Gx_{2n-1}) \leq \lambda u_{2n-1},$$

where

$$\begin{aligned} u_{2n-1} \in \{&d(Sx_{2n-1}, Tx_{2n}), d(Sx_{2n-1}, Gx_{2n-1}), d(Tx_{2n}, Fx_{2n}), \\ &\frac{d(Tx_{2n}, Gx_{2n-1}) + d(Sx_{2n-1}, Fx_{2n})}{2}\} \\ = \{&d(y_{2n-1}, y_{2n}), d(y_{2n}, y_{2n+1}), \frac{d(y_{2n-1}, y_{2n+1})}{2}\}. \end{aligned}$$

Clearly, there are infinitely many n such that at least one of the following three cases holds:

(1) $d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(y_{2n-1}, y_{2n}) = \lambda d(Sx_{2n-1}, Tx_{2n})$;

(2) $d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(y_{2n}, y_{2n+1}) \Rightarrow d(Tx_{2n}, Sx_{2n+1}) = \theta \leq \lambda d(Sx_{2n-1}, Tx_{2n})$;

(3) $d(Tx_{2n}, Sx_{2n+1}) \leq \lambda \frac{d(y_{2n-1}, y_{2n+1})}{2} \leq \frac{\lambda}{2} d(y_{2n-1}, y_{2n}) + \frac{1}{2} d(y_{2n}, y_{2n+1}) \Rightarrow d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n})$.

From (1), (2), (3) it follows that

$$d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n}). \tag{2.3}$$

Similarly, if $(Sx_{2n+1}, Tx_{2n+2}) \in Q_0 \times P_0$, we have

$$d(Sx_{2n+1}, Tx_{2n+2}) = d(Fx_{2n}, Gx_{2n+1}) \leq \lambda d(Tx_{2n}, Sx_{2n+1}). \tag{2.4}$$

If $(Sx_{2n-1}, Tx_{2n}) \in Q_0 \times P_0$, we have

$$d(Sx_{2n-1}, Tx_{2n}) = d(Fx_{2n-2}, Gx_{2n-1}) \leq \lambda d(Tx_{2n-2}, Sx_{2n-1}). \tag{2.5}$$

Case 2. If $(Tx_{2n}, Sx_{2n+1}) \in P_0 \times Q_1$, then $Sx_{2n+1} \in Q_1$ and

$$d(Tx_{2n}, Sx_{2n+1}) + d(Sx_{2n+1}, y_{2n+1}) = d(Tx_{2n}, y_{2n+1}) \tag{2.6}$$

which in turn yields

$$d(Tx_{2n}, Sx_{2n+1}) \leq d(Tx_{2n}, y_{2n+1}) = d(y_{2n}, y_{2n+1}) \tag{2.7}$$

and hence

$$d(Tx_{2n}, Sx_{2n+1}) \leq d(y_{2n}, y_{2n+1}) = d(Fx_{2n}, Gx_{2n-1}). \tag{2.8}$$

Now, proceeding as in Case 1, we have that (2.3) holds.

If $(Sx_{2n+1}, Tx_{2n+2}) \in Q_1 \times P_0$, then $Tx_{2n} \in P_0$. We show that

$$d(Sx_{2n+1}, Tx_{2n+2}) \leq \lambda d(Tx_{2n}, Sx_{2n-1}). \tag{2.9}$$

Using (2.6), we get

$$\begin{aligned} d(Sx_{2n+1}, Tx_{2n+2}) &\leq d(Sx_{2n+1}, y_{2n+1}) + d(y_{2n+1}, Tx_{2n+2}) \\ &= d(Tx_{2n}, y_{2n+1}) - d(Tx_{2n}, Sx_{2n+1}) + d(y_{2n+1}, Tx_{2n+2}). \end{aligned} \tag{2.10}$$

By noting that $Tx_{2n+2}, Tx_{2n} \in P_0$, one can conclude that

$$d(y_{2n+1}, Tx_{2n+2}) = d(y_{2n+1}, y_{2n+2}) = d(Fx_{2n}, Gx_{2n+1}) \leq \lambda d(Tx_{2n}, Sx_{2n+1}), \tag{2.11}$$

and

$$d(Tx_{2n}, y_{2n+1}) = d(y_{2n}, y_{2n+1}) = d(Fx_{2n}, Gx_{2n-1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n}), \tag{2.12}$$

in view of Case 1.

Thus,

$$d(Sx_{2n+1}, Tx_{2n+2}) \leq \lambda d(Sx_{2n-1}, Tx_{2n}) - (1-\lambda) d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n}),$$

and we proved (2.9).

Case 3. If $(Tx_{2n}, Sx_{2n+1}) \in P_1 \times Q_0$, then $Sx_{2n-1} \in Q_0$. We show that

$$d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n-2}). \tag{2.13}$$

Since $Tx_{2n} \in P_1$, then

$$d(Sx_{2n-1}, Tx_{2n}) + d(Tx_{2n}, y_{2n}) = d(Sx_{2n-1}, y_{2n}). \tag{2.14}$$

From this, we get

$$\begin{aligned} d(Tx_{2n}, Sx_{2n+1}) &\leq d(Tx_{2n}, y_{2n}) + d(y_{2n}, Sx_{2n+1}) \\ &= d(Sx_{2n-1}, y_{2n}) - d(Sx_{2n-1}, Tx_{2n}) + d(y_{2n}, Sx_{2n+1}). \end{aligned} \tag{2.15}$$

By noting that $Sx_{2n+1}, Sx_{2n-1} \in Q_0$, one can conclude that

$$d(y_{2n}, Sx_{2n+1}) = d(y_{2n}, y_{2n+1}) = d(Fx_{2n}, Gx_{2n-1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n}), \tag{2.16}$$

and

$$d(Sx_{2n-1}, y_{2n}) = d(y_{2n-1}, y_{2n}) = d(Fx_{2n-2}, Gx_{2n-1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n-2}), \tag{2.17}$$

in view of Case 1.

Thus,

$$d(Tx_{2n}, Sx_{2n+1}) \leq \lambda d(Sx_{2n-1}, Tx_{2n-2}) - (1-\lambda) d(Sx_{2n-1}, Tx_{2n}) \leq \lambda d(Sx_{2n-1}, Tx_{2n-2}),$$

and we proved (2.13).

Similarly, If $(Sx_{2n+1}, Tx_{2n+2}) \in Q_0 \times P_1$, then $Tx_{2n+2} \in P_1$, and

$$d(Sx_{2n+1}, Tx_{2n+2}) + d(Tx_{2n+2}, y_{2n+2}) = d(Sx_{2n+1}, y_{2n+2}).$$

From this, we have

$$\begin{aligned} d(Sx_{2n+1}, Tx_{2n+2}) &\le d(Sx_{2n+1}, y_{2n+2}) + d(y_{2n+2}, Tx_{2n+2}) \\ &\le d(Sx_{2n+1}, y_{2n+2}) + d(Sx_{2n+1}, y_{2n+2}) - d(Sx_{2n+1}, Tx_{2n+2}) \\ &= 2d(Sx_{2n+1}, y_{2n+2}) - d(Sx_{2n+1}, Tx_{2n+2}) \\ &\Rightarrow d(Sx_{2n+1}, Tx_{2n+2}) \le d(Sx_{2n+1}, y_{2n+2}). \end{aligned}$$

By noting that $Sx_{2n+1} \in Q_0$, one can conclude that

$$d(Sx_{2n+1}, Tx_{2n+2}) \le d(Sx_{2n+1}, y_{2n+2}) = d(Fx_{2n}, Gx_{2n+1}) \le \lambda d(Tx_{2n}, Sx_{2n+1}), \tag{2.18}$$

in view of Case 1.

Thus, in all the cases 1-3, there exists $w_{2n} \in \{d(Sx_{2n-1}, Tx_{2n}), d(Tx_{2n-2}, Sx_{2n-1})\}$ such that

$$d(Tx_{2n}, Sx_{2n+1}) \le \lambda w_{2n}$$

and exists $w_{2n+1} \in \{d(Sx_{2n-1}, Tx_{2n}), d(Tx_{2n}, Sx_{2n+1})\}$ such that

$$d(Sx_{2n+1}, Tx_{2n+2}) \le \lambda w_{2n+1}.$$

Following the procedure of Assad and Kirk ([6]), it can easily be shown by induction that, for $n \ge 1$, there exists $w_2 \in \{d(Tx_0, Sx_1), d(Sx_1, Tx_2)\}$ such that

$$d(Tx_{2n}, Sx_{2n+1}) \le \lambda^{n-\frac{1}{2}} w_2 \quad \text{and} \quad d(Sx_{2n+1}, Tx_{2n+2}) \le \lambda^n w_2. \tag{2.19}$$

From (2.19) and by the triangle inequality, for $n > m$ we have

$$\begin{aligned} d(Tx_{2n}, Sx_{2m+1}) &\le d(Tx_{2n}, Sx_{2n-1}) + d(Sx_{2n-1}, Tx_{2n-2}) + \cdots + d(Tx_{2m+2}, Sx_{2m+1}) \\ &\le (\lambda^m + \lambda^{m+\frac{1}{2}} + \cdots + \lambda^{n-1}) w_2 \le \frac{\lambda^m}{1-\sqrt{\lambda}} w_2 \to \theta, \quad \text{as} \quad m \to \infty. \end{aligned}$$

From Remark 1.3 and Corollary 1.4 (1) $d(Tx_{2n}, Sx_{2m+1}) \ll c$.

Thus, the sequence $\{Tx_0, Sx_1, Tx_2, Sx_3, \cdots, Sx_{2n-1}, Tx_{2n}, Sx_{2n-1}, \cdots\}$ is a Cauchy sequence. Then, as noted in [13], there exists at least one subsequence $\{Tx_{2n_k}\}$ or $\{Sx_{2n_k+1}\}$ which is contained in P_0 or Q_0 respectively and finds its limit $z \in C$. Furthermore, subsequences $\{Tx_{2n_k}\}$ and $\{Sx_{2n_k+1}\}$ both converge to $z \in C$ as C is a closed subset of complete cone metric space (X, d). We assume that there exists a subsequence $\{Tx_{2n_k}\} \subseteq P_0$ for each $k \in N$, then $Tx_{2n_k} = y_{2n_k} = Gx_{2n_k-1} \in C \cap GC \subseteq TC$ Since TC as well as SC are closed in X and $\{Tx_{2n_k}\}$ is Cauchy sequence in TC, it converges to a point $z \in TC$. Let $w \in T^{-1}z$, then $Tw = z$. Similarly, $\{Sx_{2n_k+1}\}$ being a subsequence of Cauchy sequence $\{Tx_0, Sx_1, Tx_2, Sx_3, \cdots, Sx_{2n-1}, Tx_{2n}, Sx_{2n-1}, \cdots\}$ also converges to z as SC is closed. Using (2.2), one can write

$$d(Fw, z) \le d(Fw, Gx_{2n_k-1}) + d(Gx_{2n_k-1}, z) \le \lambda u_{2n_k-1} + d(Gx_{2n_k-1}, z),$$

where

$$\begin{aligned} u_{2n_k-1} \in \Big\{&d(Tw, Sx_{2n_k-1}), d(Tw, Fw), d(Sx_{2n_k-1}, Gx_{2n_k-1}), \\ &\frac{d(Tw, Gx_{2n_k-1}) + d(Fw, Sx_{2n_k-1})}{2}\Big\} \\ = \Big\{&d(z, Sx_{2n_k-1}), d(z, Fw), d(Sx_{2n_k-1}, Gx_{2n_k-1}), \\ &\frac{d(z, Gx_{2n_k-1}) + d(Fw, Sx_{2n_k-1})}{2}\Big\}. \end{aligned}$$

Let $\theta \ll c$. Clearly at least one of the following four cases holds for infinitely many n.

(1) $d(Fw, z) \le \lambda d(z, Sx_{2n_k-1}) + d(Gx_{2n_k-1}, z) \ll \lambda\frac{c}{2\lambda} + \frac{c}{2} = c$;

(2) $d(Fw, z) \le \lambda d(z, Fw) + d(Gx_{2n_k-1}, z) \Rightarrow d(Fw, z) \le \frac{1}{1-\lambda} d(Gx_{2n_k-1}, z) \ll \frac{1}{1-\lambda}(1-\lambda)c = c$;

(3) $d(Fw,z) \leq \lambda d(Sx_{2n_k-1}, Gx_{2n_k-1}) + d(Gx_{2n_k-1}, z) \leq \lambda(d(Sx_{2n_k-1}, z) + d(z, Gx_{2n_k-1})) + d(Gx_{2n_k-1}, z)$
$\leq (\lambda+1)d(Gx_{2n_k-1}, z) + \lambda d(Sx_{2n_k-1}, z) \ll (\lambda+1)\frac{c}{2(\lambda+1)} + \lambda\frac{c}{2\lambda} = c$;

(4) $d(Fw,z) \leq \lambda\frac{d(z,Gx_{2n_k-1})+d(Fw,Sx_{2n_k-1})}{2} + d(Gx_{2n_k-1}, z)$
$\leq \lambda\frac{d(z,Gx_{2n_k-1})+d(z,Sx_{2n_k-1})}{2} + \frac{1}{2}d(Fw,z) + d(Gx_{2n_k-1}, z)$
$\Rightarrow d(Fw,z) \leq (2+\lambda)d(Gx_{2n_k-1}, z) + \lambda d(z, Sx_{2n_k-1}) \ll (2+\lambda)\frac{c}{2(2+\lambda)} + \lambda\frac{c}{2\lambda} = c$

In all the cases we obtain $d(Fw,z) \ll c$ for each $c \in intP$, using Corollary 1.4 (3) it follows that $d(Fw,z) = \theta$ or $Fw = z$. Thus, $Fw = z = Tw$, that is z is a coincidence point of F, T.

Further, since Cauchy sequence $\{Tx_0, Sx_1, Tx_2, Sx_3, \cdots, Sx_{2n-1}, Tx_{2n}, Sx_{2n-1}, \cdots\}$ converges to $z \in C$ and $z = Fw, z \in FC \cap C \subseteq SC$, there exists $v \in C$ such that $Sv = z$. Again using (2.2), we get

$$d(Sv, Gv) = d(z, Gv) = d(Fw, Gv) \leq \lambda u,$$

where

$$u \in \{d(Tw, Sv), d(Tw, Fw), d(Sv, Gv), \frac{d(Tw, Gv) + d(Fw, Sv)}{2}\}$$
$$= \{\theta, \theta, d(Sv, Gv), \frac{d(z, Gv) + \theta}{2}\} = \{\theta, d(Sv, Gv), \frac{d(Sv, Gv)}{2}\}.$$

Hence, we get the following cases:

$$d(Sv, Gv) \leq \lambda\theta = \theta, d(Sv, Gv) \leq \lambda d(Sv, Gv)$$

and

$$d(Sv, Gv) \leq \frac{\lambda}{2}d(Sv, Gv) \leq \lambda d(Sv, Gv).$$

Using Remark 1.3 and Corollary 1.4 (3), it follows that $Sv = Gv$, therefore, $Sv = z = Gv$, that is z is a coincidence point of (G, S).

In case FC and GC are closed in X, then $z \in FC \cap C \subseteq SC$ or $z \in GC \cap C \subseteq TC$. The analogous arguments establish (IV) and (V). If we assume that there exists a subsequence $\{Sx_{2n_k+1}\} \subseteq Q_0$ with TC as well SC are closed in X, then noting that $\{Sx_{2n_k+1}\}$ is a Cauchy sequence in SC, foregoing arguments establish (IV) and (V).

Suppose now that (F, T) and (G, S) are weakly compatible pairs, then

$$z = Fw = Tw \Rightarrow Fz = FTw = TFw = Tz$$

and

$$z = Gv = Sv \Rightarrow Gz = GSv = SGv = Sz.$$

Then, from (2.2),

$$d(Fz, z) = d(Fz, Gv) \leq \lambda u,$$

where

$$u \in \{d(Sv, Tz), d(Tz, Fz), d(Sv, Gv), \frac{d(Tz, Gv) + d(Sv, Fz)}{2}\}$$
$$= \{d(z, Fz), d(z, z), \frac{d(Fz, z) + d(z, Fz)}{2}\}$$
$$= \{d(z, Fz), \theta\}.$$

Hence, we get the following cases:

$$d(Fz, z) \leq \lambda d(z, Fz) \Rightarrow d(Fz, z) = 0,$$
$$d(Fz, z) \leq \lambda\theta = \theta \Rightarrow d(Fz, z) = 0.$$

Using Remark 1.3 and Corollary 1.1 (3), it follows that $Fz = z$. Thus, $Fz = z = Tz$

Similarly, we can prove $Gz = z = Sz$. Therefore $z = Fz = Gz = Sz = Tz$, that is, z is a common fixed point of F, G, S and T.

Uniqueness of the common fixed point follows easily from (2.2). □

2.4. Remark. 1. Theorem 2.2 in [18] is a special case of Theorem 2.3 with $G = F = f, T = S = g$ and $\lambda \in (0, \sqrt{2} - 1)$.

2. Setting $G = F = f$ and $T = S = I_X$ (the identity mapping on X) in Theorem 2.3, we obtain the following result:

2.5. Corollary. *Let (X, d) be a complete cone metric space, and C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Suppose that $f : C \to X$ satisfying the condition

$$d(fx, fy) \leq \lambda u(x, y),$$

where

$$u(x, y) \in \{d(x, y), d(x, fx), d(y, fy), \frac{d(x, fy) + d(y, fx)}{2}\}$$

for all $x, y \in C, 0 < \lambda < 1$ and f has the additional property that for each $x \in \partial C, fx \in C$. Then f has a unique fixed point.

2.6. Remark. The following definition is a special case of Definition 2.2 when (X, d) is a metric space. But when (X, d) is a cone metric space, which is not a metric space, this is not true. Indeed, there may exist $x, y \in X$ such that the vectors $d(Tx, Fx), d(Sy, Gy)$ and $\frac{d(Tx,Fx)+d(Sy,Gy)}{2}$ are incomparable. For the same reason Theorems 2.3 and 2.8 (given below) are incomparable.

2.7. Definition. Let (X, d) be a complete cone metric space and C be a nonempty closed subset of X, and $F, G, S, T : C \to X$. Denote, for$x, y \in C$,

$$M_2^{F,G,S,T} = \{d(Tx, Sy), \frac{d(Tx, Fx) + d(Sy, Gy)}{2}, \frac{d(Tx, Gy) + d(Fx, Sy)}{2}\}. \tag{2.20}$$

Then (F, G) is called a generalized $(T, S)_{M_2}-$contractive mapping of C into X if for some $\lambda \in (0, 1)$ there exists $u(x, y) \in M_2^{F,G,S,T}$ such that for all $x, y \in C$

$$d(Fx, Gy) \leq \lambda u(x, y). \tag{2.21}$$

Our next result is the following.

2.8. Theorem. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Suppose that $F, G, S, T : C \to X$ are such that (F, G) is a generalized $(T, S)_{M_2}-$contractive mappings pair of C into X, and

(I) $\partial C \subseteq SC \cap TC, FC \cap C \subseteq SC, GC \cap C \subseteq TC$,

(II) $Tx \in \partial C$ implies that $Fx \in C, Sx \in \partial C$ implies that $Gx \in C$,

(III) SC and TC (or FC and GC) are closed in X.

Then

(IV) (F, T) has a point of coincidence,

(V) (G, S) has a point of coincidence.

Moreover, if (F, T) and (G, S) are weakly compatible pairs, then F, G, S and T have a unique common fixed point.

The proof of this theorem is very similar to the proof of Theorem 2.3 and it is omitted.

We now list some corollaries of Theorems 2.3 and 2.8.

2.9. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $F, G, S, T : C \to X$ be such that

$$d(Fx, Gy) \le \lambda d(Tx, Sy), \tag{2.22}$$

for some $\lambda \in (0, 1)$ and for all $x, y \in C$.

Suppose, further, that F, G, S, T and C satisfy the following conditions:
(I) $\partial C \subseteq SC \cap TC, FC \cap C \subseteq SC, GC \cap C \subseteq TC$,
(II) $Tx \in \partial C$ implies that $Fx \in C, Sx \in \partial C$ implies that $Gx \in C$,
(III) SC and TC (or FC and GC) are closed in X.
Then
(IV) (F, T) has a point of coincidence,
(V) (G, S) has a point of coincidence.

Moreover, if (F, T) and (G, S) are weakly compatible pairs, then F, G, S and T have a unique common fixed point.

2.10. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $F, G, S, T : C \to X$ be such that

$$d(Fx, Gy) \le \lambda(d(Tx, Fx) + d(Sy, Gy)), \tag{2.23}$$

for some $\lambda \in (0, 1/2)$ and for all $x, y \in C$.

Suppose, further, that F, G, S, T and C satisfy the following conditions:
(I) $\partial C \subseteq SC \cap TC, FC \cap C \subseteq SC, GC \cap C \subseteq TC$,
(II) $Tx \in \partial C$ implies that $Fx \in C, Sx \in \partial C$ implies that $Gx \in C$,
(III) SC and TC (or FC and GC) are closed in X.
Then
(IV) (F, T) has a point of coincidence,
(V) (G, S) has a point of coincidence.

Moreover, if (F, T) and (G, S) are weakly compatible pairs, then F, G, S and T have a unique common fixed point.

2.11. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $F, G, S, T : C \to X$ be such that

$$d(Fx, Gy) \le \lambda(d(Tx, Gy) + d(Fx, Sy)), \tag{2.24}$$

for some $\lambda \in (0, 1/2)$ and for all $x, y \in C$.

Suppose, further, that F, G, S, T and C satisfy the following conditions:
(I) $\partial C \subseteq SC \cap TC, FC \cap C \subseteq SC, GC \cap C \subseteq TC$,
(II) $Tx \in \partial C$ implies that $Fx \in C, Sx \in \partial C$ implies that $Gx \in C$,
(III) SC and TC (or FC and GC) are closed in X.
Then
(IV) (F, T) has a point of coincidence,
(V) (G, S) has a point of coincidence.

Moreover, if (F, T) and (G, S) are weakly compatible pairs, then F, G, S and T have a unique common fixed point.

2.12. Remark. Setting $G = F = f$ and $T = S = g$ (the identity mapping on X) in Corollary 2.9-2.11, we obtain the following result:

2.13. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $f, g : C \to X$ be such that

$$d(fx, fy) \leq \lambda d(gx, gy), \tag{2.25}$$

for some $\lambda \in (0, 1)$ and for all $x, y \in C$. Suppose, further, that f, g and C satisfy the following conditions:

(I) $\partial C \subseteq gC, fC \cap C \subseteq gC$,
(II) $gx \in \partial C$ implies that $fx \in C$,
(III) gC is closed in X.

Then there exists a coincidence point z of f, g in C. Moreover, if (f, g) are weakly compatible, then z is the unique common fixed point of f and g.

2.14. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $f, g : C \to X$ be such that

$$d(fx, fy) \leq \lambda(d(fx, gx) + d(fy, gy)), \tag{2.26}$$

for some $\lambda \in (0, 1/2)$ and for all $x, y \in C$. Suppose, further, that f, g and C satisfy the following conditions:

(I) $\partial C \subseteq gC, fC \cap C \subseteq gC$,
(II) $gx \in \partial C$ implies that $fx \in C$,
(III) gC is closed in X.

Then there exists a coincidence point z of f, g in C. Moreover, if (f, g) are weakly compatible, then z is the unique common fixed point of f and g.

2.15. Corollary. *Let (X, d) be a complete cone metric space, C a nonempty closed subset of X such that for each $x \in C$ and $y \notin C$ there exists a point $z \in \partial C$ such that*

$$d(x, z) + d(z, y) = d(x, y).$$

Let $f, g : C \to X$ be such that

$$d(fx, fy) \leq \lambda(d(fx, gy) + d(fy, gx)), \tag{2.27}$$

for some $\lambda \in (0, 1/2)$ and for all $x, y \in C$. Suppose, further, that f, g and C satisfy the following conditions:

(I) $\partial C \subseteq gC, fC \cap C \subseteq gC$,
(II) $gx \in \partial C$ implies that $fx \in C$,
(III) gC is closed in X.

Then there exists a coincidence point z of f, g in C. Moreover, if (f, g) are weakly compatible, then z is the unique common fixed point of f and g.

2.16. Remark. Corollaries 2.13-2.15 are the corresponding theorems of Abbas and Jungck from [2] in the case that f, g are non-self mappings.

Acknowledgements The authors would like to thank the referees for their constructive suggestions and kind comments. This work is supported by the National Natural Science Foundation of China(11461043, 11361042 and 11326099)) and supported partly by the Provincial Natural Science Foundation of Jiangxi, China (20114BAB201003 and

20142BAB201005) and the Science and Technology Project of Educational Commission of Jiangxi Province, China (GJJ11294).

References

[1] Abbas, M. Cho, Y. J. and Nazir, T. *Common fixed point theorems for four mappings in TVS-valued cone metric spaces*, J. Math. Inequal. 5 (2009), 287-299.

[2] Abbas, M. and Jungck, G. *Common fixed point results for noncommuting mappings without continuity in cone metric spaces*, J. Math. Anal. Appl. 341(1) (2008), 416-420.

[3] Abbas, M. Rhoades, B.E. *Fixed and periodic point results in cone metric spaces*, Appl. Math. Lett. 22 (2009), 511-515.

[4] Arshad, M. Azam, A. and Beg, I. *Common fixed points of two maps in cone metric spaces*, Rend. Circ. Mat. Palermo, 57 (2008), 433-411.

[5] Arshad, M. Azam, A. and Vetro, P. *Some common fixed point results in cone metric spaces*, Fixed Point Theory Appl. Volume 2009, Article ID 493965, 11 pages.

[6] Assad, N. A. and Kirk, W. A. *Fixed point theorems for set valued mappings of contractive type*, Pacific J. Math. 43(3): (1972), 553-562.

[7] Cho, Y. J. Rhoades, B. E. Saadati, R. Samet, B. and Shantawi, W. *Nonlinear coupled fixed point theorems in ordered generalized metric spaces with integral type*, Fixed Point Theory Appl. 2012, 2012:8.

[8] Cho, Y. J. Saadati, R. and Wang, S. *Common fixed point theorems on generalized distance in order cone metric spaces*, Comput. Math. Appl. 61(2011), 1254-1260.

[9] Cho, Y. J. Shah, M. H. Hussain, N. *Coupled fixed points of weakly F- contractive mappings in topological spaces*, Appl. Math. Lett. 24(2011), 1185-1190.

[10] Di Bari, C. and Vetro, P. *φ-pairs and common fixed points in cone metric spaces*, Rend. Circ. Mat. Palermo, 57 (2008), 279-285.

[11] Di Bari, C. and Vetro, P. *Weakly φ-pairs and common fxed points in cone metric spaces*, Rend. Circ. Mat. Palermo, 58 (2009), 125-132.

[12] Graily, E. Vaezpour, S. M. Saadati, R. and Cho, Y. J. *Generalization of fixed point theorems in ordered metric spaces concerning generalized distance*, Fixed Point Theory Appl. 2011, 2011:30.

[13] Hadzic, O. and Gajic, Lj. *Coincidence points for set-valued mappings in convex metric spaces*, Univ. U. Novom. Sadu. Zb. Rad. Prirod. Mat. Fak. Ser. Mat. 16(1) (1986), 13-25.

[14] Huang, L. G. and Zhang, X. *Cone metric spaces and fixed point theorems of contractive mappings*, J. Math. Anal. Appl. 332 (2007) 1468-1476.

[15] Huang, X. J. Zhu, C. X. and Wen, X. *Common fixed point theorem for four non-self mappings in cone metric spaces.* Fixed Point Theory Appl. Volume 2010, Article ID 983802, 14 pages.

[16] Ilic, D. and Rakocevic, V. *Common fixed points for maps on cone metric space*, J. Math. Anal. Appl. 341(2) (2008), 876-882.

[17] Jungck, G. Radenovic, S. Radojevic, S. and Rakocevic, V. *Common fixed point theorems for weakly compatible pairs on cone metric spaces*, Fixed Point Theory and Appl. Volume 2009 (2009), Article ID 643840, 13 pages.

[18] Jankovic, S. Kadelburg, Z. Radenovic, S. and Rhoades, B.E. *Assad-Kirk-Type fixed point theorems for a pair of nonself mappings on cone metric spaces*, Fixed Point Theory and Appl. Volume 2009, Article ID 761086, 16 pages.

[19] Rezapour, Sh. *A review on topological properties of cone metric spaces*, Analysis, Topology and Applications 2008, Vrnjacka Banja, Serbia, from May 30 to June 4, 2008.

[20] Rezapour, Sh. and Hamlbarani, R. *Some notes on the paper "Cone metric spaces and fixed point theorems of contractive mappings"*, J. Math. Anal. Appl. 345: (2008) 719-724.

[21] Vetro, P. *Common fixed points in cone metric spaces*, Rendiconti del circolo matematico di Palermo, Serie II, Tomo LVI (2007) 464-468.

[22] Sintunavarat, W. Y. Cho, J. and Kumam, P. *Common fixed point theorems for c-distance in ordered cone metric spaces*, Comput. Math. Appl. 62(2011), 1969-1978.

[23] Wardowski, D. *Endpoints and fixed points of set-valued contractions in cone metric space*, Nonlinear Anal. 71 (2009), 512-516.

[24] Wong, Y. C. and Ng, K. F. *Partially ordered topological vector spaces*, Clarendon Press-Oxford, 1973.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 559 – 568

Soft separation axioms in soft topological spaces

Sabir Hussain * and Bashir Ahmad †

Abstract

Shabir et. al [27] and D. N. Georgiou et. al [7], defined and studied some soft separation axioms, soft θ-continuity and soft connectedness in soft spaces using (ordinary) points of a topological space X. In this paper, we redefine and explore several properties of soft T_i, $i = 0, 1, 2$, soft regular, soft T_3, soft normal and soft T_4 axioms using soft points defined by I. Zorlutuna [30]. We also discuss some soft invariance properties namely soft topological property and soft hereditary property. We hope that these results will be useful for the future study on soft topology to carry out general framework for the practical applications and to solve the complicated problems containing uncertainties in economics, engineering, medical, environment and in general man-machine systems of various types.

Keywords: Soft topology, Soft open(closed) sets, Soft interior(closure), Soft T_i; ($i = 0, 1, 2, 3, 4$) spaces, Soft regular spaces, Soft normal spaces and Invariance properties.

Received 03/02/2014 : Accepted 12/03/2014 Doi : 10.15672/HJMS.2015449426

1. Introduction

In 1999, Molodtsov [22] initiated the theory of soft sets as a new mathematical tool to deal with uncertainties while modelling the problems with incomplete information in engineering, physics, computer science, economics, social sciences and medical sciences. Soft set theory does not require the specification of parameters. Instead, it accommodates approximate description of an object as its starting point which makes it a natural mathematical formalism for approximate reasoning. So the application of soft set theory in other disciplines and real life problems are now catching momentum. In [23], Molodtsov applied soft sets successfully in directions such as smoothness of functions,

*Department of Mathematics, College of Science, Qassim University, P. O. Box 6644, Buraydah 51482, Saudi Arabia.
Email: sabiriub@yahoo.com, sh.hussain@qu.edu.sa
†Department of Mathematics, King Abdul Aziz University, P. O. Box 80203, Jeddah 21589, Saudi Arabia.
Email:drbashir9@gmail.com

game theory, operations research, Riemann-integration, Perron integration, probability and theory of measurement. Maji et. al [20] applied soft sets in a multicriteria decision making problems. It is based on the notion of knowledge reduction of rough sets. They applied the technique of knowledge reduction to the information table induced by the soft set. In [21], they defined and studied several basic notions of soft set theory. In 2005, Pei and Miao [25] and Chen [6] improved the work of Maji et.al [20-21]. A. Kharal and B. Ahmad [19] defined and discussed the several properties of soft images and soft inverse images of soft sets. They also applied these notions to the problem of medical diagnosis in medical systems. Many researchers have contributed towards the algebraic structure of soft set theory [1-2],[5], [7], [9-19], [24], [27-28].

In 2011, Shabir and Naz [27] initiated the study of soft topological spaces. Also in 2011, S. Hussain and B. Ahmad [9] continued investigating the properties of soft open(closed), soft neighbourhood and soft closure. They also defined and discussed the properties of soft interior, soft exterior and soft boundary.

Shabir et. al [27] and D. N. Georgiou et. al [7], defined and studied some soft separation axioms, soft θ-continuity and soft connectedness in soft spaces using (ordinary) points of a topological space X. In this paper, we redefine and explore several properties of soft T_i, $i = 0, 1, 2$, soft regular, soft T_3, soft normal and soft T_4 axioms using soft points defined by I. Zorlutuna [30]. We also discuss some soft invariance properties namely soft topological property and soft hereditary property. We hope that these results will be useful for the future study on soft topology to carry out general framework for the practical applications and to solve the complicated problems containing uncertainties in economics, engineering, medical, environment and in general man-machine systems of various types.

2. Preliminaries

For the definitions and results on soft set theory, we refer to [1-2],[5], [7], [9-19], [24], [27-28]. However, we recall some definitions and results on soft set theory and soft topology.

Definition 1 [22]. Let X be an initial universe and E a set of parameters. Let $P(X)$ denote the power set of X and A a non-empty subset of E. A pair (F, A) is called a soft set over X, where F is a mapping given by $F : A \to P(X)$. In other words, a soft set over X is a parameterized family of subsets of the universe X. For $e \in A$, $F(e)$ may be considered as the set of e-approximate elements of the soft set (F, A). Clearly, a soft set is not a set.

Definition 2 [22]. The complement of a soft set (F, A) is denoted by $(F, A)^c$ and is defined by $(F, A)^c = (F^c, A)$ where, $F^c : A \to P(U)$ is a mapping given by $F^c(\alpha) = U - F(\alpha)$, for all $\alpha \in A$.

Let us call F^c to be the soft complement function of F. Clearly $(F^c)^c$ is the same as F and $((F, A)^c)^c = (F, A)$.

Definition 3 [30]. A soft set (F, A) over X is said to be a null soft set, denoted by Φ_A , if for all $e \in A$, $F(e) = \phi$. Clearly, $(\Phi_A^c)^c = \Phi_A$.

Definition 4 [30]. A soft set (F, A) over X is said to be an absolute soft set, denoted by X_A, if for all $e \in A$, $F(e) = X$. Clearly, $X_A^c = \Phi_A$.

Definition 5 [27]. Let τ be the collection of soft sets over X with the fixed set of parameters A. Then τ is said to be a soft topology on X, if

(1) Φ_A, X_A belong to τ,

(2) the union of any number of soft sets in τ belongs to τ,

(3) the intersection of any two soft sets in τ belongs to τ.

The triplet (X, τ, A) is called a soft topological space over X. The members of τ are called soft open sets. The soft complement of a soft open set A is called the soft closed set in (X, τ, A). If (F, A) belongs to τ, we write $(F, A)\tilde{\in}\tau$.
Proposition 1 [27]. Let (X, τ, A) be a soft topological space over X. Then the collection $\tau_\alpha = \{F(e) : (F, A)\tilde{\in}\tau\}$, for each $e \in A$ defines a topology on X.
Remark 1. It is known that the intersection of two soft topologies over the same universe X is a soft topology, whereas the union may or may not be a soft topology as given in [27].

Hereafter, $SS(X)_A$ denotes the family of soft sets over X with the set of parameters A.
Definition 6 [30]. The soft set $(F, A)\tilde{\in}SS(X)_A$ is called a soft point in X_A, denoted by e_F, if for the element $e \in A$, $F(e) \neq \phi$ and $F(e^{'}) = \phi$, for all $e^{'} \in A - \{e\}$.
Definition 7 [30]. The soft point e_F is said to be in the soft set (G, A), denoted by $e_F\tilde{\in}(G, A)$, if for the element $e \in A$, $F(e) \subseteq G(e)$.
Proposition 2 [30]. Let $e_F\tilde{\in}X_A$ and $(G, A)\tilde{\in}SS(X)_A$. If $e_F\tilde{\in}(G, A)$, then $e_F\tilde{\notin}(G, A)^c$.
Definition 8 [30]. Let (X, τ, A) be a soft topological space over X and (F, A) a soft set in $SS(X)_A$. The soft point $e_F\tilde{\in}X_A$ is called a soft interior point of a soft set (F, A), if there exists a soft open set (H, A) such that $e_F\tilde{\in}(H, A)\tilde{\subseteq}(F, A)$. The soft interior of a soft set (F, A) is denoted by $(F, A)^\circ$ and is defined as the union of all soft open sets contained in (F, A). Clearly $(F, A)^\circ$ is the largest soft open set contained in (F, A).
Definition 9 [30]. Let (X, τ, A) be a soft topological space. Then a soft set (G, A) in $SS(X)_A$ is called a soft neighborhood (briefly: soft nbd) of the soft point $e_F\tilde{\in}X_A$, if there exists a soft open set (H, A) such that $e_F\tilde{\in}(H, A)\tilde{\subseteq}(G, A)$.
The soft neighborhood system of a soft point e_F, denoted by $N_\tau(e_F)$, is the family of all its soft neighborhoods.
Definition 10 [27]. Let (X, τ, A) be a soft topological space over X. Then a soft set (G, A) in $SS(X)_A$ is called a soft neighborhood (briefly: soft nbd) of the soft set (F, A), if there exists a soft open set (H, A) such that $(F, A)\tilde{\subseteq}(H, A)\tilde{\subseteq}$
(G, A).
Definition 11 [27]. Let (X, τ, A) be a soft topological space over X and (F, A) a soft set over X. Then the soft closure of (F, A), denoted by $\overline{(F, A)}$, is the intersection of all soft closed supersets of (F, A). Clearly $\overline{(F, A)}$ is the smallest soft closed set in (X, τ, A) which contains (F, A).
Definition 12 [27]. Let (X, τ, A) be a soft topological space over X and $Y \subseteq X$. Then $\tau_Y = \{(F_Y, A) = Y_A\tilde{\cap}(F, A)|(F, A)\tilde{\in}\tau\}$ is said to be the soft relative topology on Y, where $F_Y(e) = Y \cap F(e)$, for all $e \in A$. (Y, τ_Y, A) is called a soft subspace of (X, τ, A). We can easily verify that τ_Y is, in fact, a soft topology on Y.
Proposition 3 [27]. Let (X, τ, A) be a soft topological space over X and $Y \subseteq X$. Then (Y, τ_{eY}) is a subspace of (X, τ_e), for each $e \in A$.
Proposition 4 [27]. Let (Y, τ_Y, A) be a soft subspace of a soft topological space (X, τ, A) and (F, A) a soft open set in (Y, τ_Y, A). If $Y_A\tilde{\in}\tau$, then $(F, A)\tilde{\in}\tau$.
Theorem 1 [27]. Let (Y, τ_Y, A) be a soft subspace of a soft topological space (X, τ, A) and (F, A) a soft set over X. Then
(1) (F, A) is soft open in (Y, τ_Y, A) if and only if $(F, A) = Y_A\tilde{\cap}(G, A)$, for some soft open set (G, A) in (X, τ, A).
(2) (F, A) is soft closed in (Y, τ_Y, A) if and only if $(F, A) = Y_A\tilde{\cap}(G, A)$, for some soft closed set (G, A) in (X, τ, A).

3. Soft T_i; $(i = 0, 1, 2)$ Spaces

In this section, we redefine soft separation axioms namely soft T_i axioms, for $(i = 0, 1, 2)$ using soft points and discuss several properties and their relationship with the help of examples. Note that some authors ([7],[27]) defined soft separation axioms using ordinary points of a topological space.
Now we define:
Definition 13. Two soft sets (G, A), (H, A) in $SS(X)_A$ are said to be soft disjoint, written $(G, A)\tilde{\cap}(H, A) = \Phi_A$, if $G(e) \cap H(e) = \phi$, for all $e \in A$.

Definition 14. Two soft points e_G, e_H in X_A are distinct, written $e_G \neq e_H$, if there corresponding soft sets (G, A) and (H, A) are disjoint.

Definition 15. Let (X, τ, A) be a soft topological space over X and $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$. If there exist at least one soft open set (F_1, A) or (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\notin} (F_1, A)$ or $e_H \tilde{\in} (F_2, A)$, $e_G \tilde{\notin} (F_2, A)$, then (X, τ, A) is called a soft T_0-space.

Definition 16. Let (X, τ, A) be a soft topological space over X and $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$. If there exist soft open sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\notin} (F_1, A)$ and $e_H \tilde{\in} (F_2, A)$, $e_G \tilde{\notin} (F_2, A)$, then (X, τ, A) is called a soft T_1-space.

Definition 17. Let (X, τ, A) be a soft topological space over X and $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$. If there exist soft open sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\in} (F_2, A)$ and $(F_1, A)\tilde{\cap}(F_2, A) = \Phi_A$, then (X, τ, A) is called a soft T_2-space.

Proposition 5. (1) Every soft T_1-space is a soft T_0-space.
(2) Every soft T_2-space is a soft T_1-space.
Proof. (1) Obvious.
(2) If (X, τ, A) is a soft T_2-space, then by definition of soft T_2-space, for $e_G, e_H \tilde{\in} X_A$, $e_G \neq e_H$, there exist soft open sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\in} (F_2, A)$ and $(F_1, A)\tilde{\cap}(F_2, A) = \Phi_A$. Since $(F_1, A)\tilde{\cap}(F_2, A) = \Phi_A$, $e_G \tilde{\notin} (F_2, A)$ and $e_H \tilde{\notin} (F_1, A)$. Thus it follows that (X, τ, A) is a soft T_1-space. □

Note that every soft T_1-space is a soft T_0-space. Every soft T_2-space is a soft T_1-space. The converses do not hold in general.
Example 1. Let $X = \{x_1, x_2\}$, $A = \{e_1, e_2\}$ and $\tau = \{\Phi_A, X_A, (F, A)\}$, where $(F, A) = \{(e_1, \{x_1\}), (e_2, \{x_2\})\}$. Then (X, τ, A) is a soft topological space over X. There are two pairs of soft points namely $e_{1(G_1)} = (e_1, \{x_2\})$, $e_{1(H_1)} = (e_1, \{x_1\})$ and $e_{2(G_2)} = (e_2, \{x_1\})$, $e_{2(H_2)} = (e_2, \{x_2\})$. Since $e_{1(G_1)} \neq e_{1(H_1)}$, then there is soft open set (F, A) such that $e_{1(G_1)} \tilde{\notin} (F, A)$, $e_{1(H_1)} \tilde{\in} (F, A)$. Similarly for the pair $e_{2(G_2)} \neq e_{2(H_2)}$, there is soft open set (F, A) such that $e_{2(H_2)} \tilde{\in} (F, A)$, $e_{2(G_2)} \tilde{\notin} (F, A)$. This shows that (X, τ, A) is a soft T_0-space. Clearly (X, τ, A) is not a soft T_1-space.

Example 2. Let $X = \{x_1, x_2\}$, $A = \{e_1, e_2\}$ and $\tau = \{\Phi_A, X_A, (F_1, A), (F_2, A), (F_3, A), (F_4, A)\}$, where $(F_1, A) = \{(e_1, \{x_2\}), (e_2, \{x_1\})\}$, $(F_2, A) = \{(e_1, \{x_1\}), (e_2, \{x_2\})\}$, $(F_3, A) = \{(e_1, \{x_1\})\}$, $(F_4, A) = \{(e_1, X), (e_2, \{x_1\})\}$ Then (X, τ, A) is a soft topological space over X. Note that $\tau_{e_1} = \{\phi, X, \{x_1\}, \{x_2\}\}$ and $\tau_{e_2} = \{\phi, X, \{x_1\}, \{x_2\}\}$ are topologies on X. Clearly (X, τ_{e_1}) and (X, τ_{e_2}) are T_i-spaces (for $i = 0, 1$). There are two pairs of distinct soft points namely, $e_{1(G_1)} = (e_1, \{x_2\})$, $e_{1(H_1)} = (e_1, \{x_1\})$ and $e_{2(G_2)} = (e_2, \{x_1\})$, $e_{2(H_2)} = (e_2, \{x_2\})$. Then for the soft pair $e_{1(G_1)} \neq e_{1(H_1)}$ of points, there are soft open sets (F_1, A) and (F_2, A) such that $e_{1(G_1)}\tilde{\in}(F_1, A)$, $e_{1(H_1)}\tilde{\notin}(F_1, A)$ and $e_{1(H_1)}\tilde{\in}(F_2, A)$, $e_{1(G_1)}\tilde{\notin}(F_2, A)$. Similarly for the pair $e_{2(G_2)} \neq e_{2(H_2)}$, there are soft open sets (F_2, A) and (F_1, A) such that $e_{2(G_2)}\tilde{\notin}(F_2, A)$, $e_{2(H_2)}\tilde{\in}(F_2, A)$ and $e_{2(H_2)}\tilde{\notin}(F_1, A)$, $e_{2(G_2)}\tilde{\in}(F_1, A)$. This shows that (X, τ, A) is a soft T_1-space and hence a soft T_0-space. Note that (X, τ, A) is a soft T_2-space.

Example 3. Let $X = \{x_1, x_2\}$, $A = \{e_1, e_2\}$ and $\tau = \{\Phi_A, X_A, (F_1, A),$ $(F_2, A), (F_3, A)\}$, where $(F_1, A) = \{(e_1, \{x_1\})\}$, $(F_2, A) = \{(e_2, \{x_2\})\}$, $(F_3, A) = \{(e_1, \{x_1\}), (e_2, \{x_2\})\}$. Then (X, τ, A) is a soft topological space over X. There are two pairs of distinct soft points namely, $e_{1(G_1)} = (e_1, \{x_2\})$, $e_{1(H_1)} = (e_1, \{x_1\})$ and $e_{2(G_2)} = (e_2, \{x_1\})$, $e_{2(H_2)} = (e_2, \{x_2\})$. Then for the soft pair $e_{1(G_1)} \neq e_{1(H_1)}$ of points, there does not exist soft disjoint soft open sets (F, A) and (G, A) such that $e_{1(G_1)}\tilde{\in}(F, A)$, $e_{1(H_1)}\tilde{\notin}(F, A)$ and $e_{1(H_1)}\tilde{\in}(G, A)$, $e_{1(G_1)}\tilde{\notin}(G, A)$. Thus (X, τ, A) is not a soft T_2-space. Clearly (X, τ, A) is a soft T_1-space and hence a soft T_0-space.

Theorem 2. Let (X, τ, A) be a soft topological space over X. Then each soft point is soft closed if and only if (X, τ, A) is a soft T_1-space.

Proof. Suppose soft points $e_F = (F, A)$, $e_G = (G, A)$ are soft closed and $e_F \neq e_G$. Then $(F, A)^c$ and $(G, A)^c$ are soft open in (X, τ, A). Then by definition $(F, A)^c = (F^c, A)$, where $F^c(e) = X - F(e)$ and $(G, A)^c = (G^c, A)$, where $G^c(e) = X - G(e)$. Since $F(e) \cap G(e) = \Phi$. This implies $F(e) \subseteq X - G(e) = G^c(e)$, for all e. This implies $e_F = (F, A)\tilde{\in}(G, A)^c$. Similarly $e_G = (G, A)\tilde{\in}(F, A)^c$. Thus we have $e_F\tilde{\in}(G, A)^c$, $e_G\tilde{\notin}(G, A)^c$ and $e_F\tilde{\notin}(F, A)^c$, $e_G\tilde{\in}(F, A)^c$. This proves that (X, τ, A) is soft T_1-space.

Conversely, let (X, τ, A) is soft T_1-space. To prove that $e_F = (F, A) \in \tilde{X}_A$ is soft closed, we show that $(F, A)^c$ is soft open in (X, τ, A). Let $e_G = (G, A)\tilde{\in}(F, A)^c$. Then $e_F \notin e_G$. Since (X, τ, A) is soft T_1-space, there exists a soft open set (H, A) such that $e_G\tilde{\in}(H, A)$ and $e_F\tilde{\notin}(H, A)$. Thus $e_G\tilde{\in}(H, A)\tilde{\subseteq}(F, A)^c$ and hence $\tilde{\bigcup}_{e_G}\{(H, A), e_G\tilde{\in}(F, A)^c\} = (F, A)^c$. This proves that $(F, A)^c$ is soft open in (X, τ, A), that is $e_F = (F, A)$ is soft closed in (X, τ, A). $\square$

Remark 2. In general, if (X, τ, A) is a soft T_1-space, then (X, τ_e) is not necessarily a T_1-space for $e \in A$. The following propositions give conditions for (X, τ_e) to be a T_1 space. We use Proposition 3 to prove this.

Proposition 6. Let (X, τ, A) be a soft topological space over X and $e_G, e_H\tilde{\in}X_A$ such that $e_G \neq e_H$. If there exist soft open sets (F_1, A) and (F_2, A) such that $e_G\tilde{\in}(F_1, A)$ and $e_H\tilde{\in}(F_1, A)^c$ or $e_H\tilde{\in}(F_2, A)$ and $e_G\tilde{\in}(F_2, A)^c$, then (X, τ, A) is a soft T_0-space and (X, τ_e) is a T_0-space, for each $e \in A$.

Proof. Clearly $e_H\tilde{\in}(F_1, A)^c = (F_1^c, A)$ implies $e_H\tilde{\notin}(F_1, A)$. Similarly $e_G\tilde{\in}(F_2, A)^c = (F_2^c, A)$ implies $e_G\tilde{\notin}(F_2, A)$. Thus we have $e_G\tilde{\in}(F_1, A)$, $e_H\tilde{\notin}(F_1, A)$ or $e_H\tilde{\in}(F_2, A)$, $e_G\tilde{\notin}(F_2, A)$. This proves (X, τ, A) is a soft T_0-space. Now for any $e \in A$, (X, τ_e) is a topological space and $e_G\tilde{\in}(F_1, A)$ and $e_H\tilde{\in}(F_1, A)^c$ or $e_H\tilde{\in}(F_2, A)$ and $e_G\tilde{\notin}(F_2, A)^c$. So that $G(e)\tilde{\in}F_1(e)$, $H(e)\tilde{\notin}F_1(e)$ or $H(e)\tilde{\in}F_2(e)$, $G(e)\tilde{\notin}F_2(e)$. Thus (X, τ_e) is a T_0-space. $\square$

Proposition 7. Let (X, τ, A) be a soft topological space over X and e_G, e_H $\tilde{\in}X_A$ such that $e_G \neq e_H$. If there exist soft open sets (F_1, A) and (F_2, A) such that $e_G\tilde{\in}(F_1, A)$, $e_H\tilde{\in}(F_1, A)^c$ and $e_H\tilde{\in}(F_2, A)$, $e_G\tilde{\in}(F_2, A)^c$ then (X, τ, A) is a soft T_1-space and (X, τ_e) is a T_1-space, for each $e \in A$.

Proof. The proof is similar to the proof of Proposition 6. $\square$

The following propositions 8, 9 and 11 show that the each soft T_i, $(i = 0, 1)$ property is a soft hereditary property.

Proposition 8. Let (X, τ, A) be a soft topological space over X and $Y \subseteq X$. If (X, τ, A) is a soft T_0-space, then (Y, τ_Y, A) is a soft T_0-space.

Proof. Let $e_G, e_H \tilde{\in} Y_A$ be such that $e_G \neq e_H$. Then $e_G, e_H \tilde{\in} X_A$. Since (X, τ, A) is a soft T_0 space, therefore there exist soft open sets (F, A) and (G, A) in (X, τ, A) such that $e_G \tilde{\in} (F, A)$ and $e_H \tilde{\notin} (F, A)$ or $e_H \tilde{\in} (G, A)$ and $e_G \tilde{\notin} (G, A)$. Therefore $e_G \tilde{\in} Y_A \tilde{\cap} (F, A) = (F_Y, A)$. Similarly it can be proved that if $e_H \tilde{\in} (G, A)$ and $e_G \tilde{\notin} (G, A)$, then $e_H \tilde{\in} (G_Y, A)$ and $e_G \tilde{\notin} (G_Y, A)$. Thus (Y, τ_Y, A) is a soft T_0-space. □

Proposition 9. Let (X, τ, A) be a soft topological space over X and Y be a non-empty subset of X. If (X, τ, A) is a soft T_1-space, then (Y, τ_Y, A) is a soft T_1-space.

Proof. The proof is similar to the proof of Proposition 8. □

Proposition 10. Let (X, τ, A) be a soft topological space over X. If (X, τ, A) is a soft T_2-space over X, then (X, τ_e) is a T_2-space, for each $e \in A$.

Proof. Let (X, τ, A) be a soft topological space over X. For any $e \in A$, $\tau_e = \{F(e) : (F, A) \tilde{\in} \tau\}$ is a topology on X. Let $x, y \in X$ such that $x \neq y$. Since (X, τ, A) is a soft-T_2 space, therefore soft points $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$ and $x \in G(e)$, $y \in H(e)$, there exist soft open sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\in} (F_2, A)$ and $(F_1, A) \tilde{\cap} (F_2, A) = \Phi_A$. This imply that $x \in G(e) \subseteq F_1(e)$, $y \in H(e) \subseteq F_2(e)$ and $F_1(e) \cap F_2(e) = \phi$. This proves that (X, τ_e) is a T_2-space. □

Proposition 11. Let (X, τ, A) be a soft topological space over X and $Y \subseteq X$. If (X, τ, A) is a soft T_2-space, then (Y, τ_Y, A) is a soft T_2-space and (X, τ_e) is a T_2-space, for each $e \in A$.

Proof. Let $e_G, e_H \tilde{\in} Y_A$ such that $e_G \neq e_H$. Then $e_G, e_H \tilde{\in} X_A$. Since (X, τ, A) is a soft-T_2-space, therefore there exist soft open sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\in} (F_2, A)$ and $(F_1, A) \tilde{\cap} (F_2, A) = \Phi_A$. Thus $e_G \tilde{\in} Y_A \tilde{\cap} (F_1, A) = (F_{1_Y}, A)$, $e_H \tilde{\in} Y_A \tilde{\cap} (F_2, A) = (F_{2_Y}, A)$ and $(F_{1_Y}, A) \tilde{\cap} (F_{2_Y}, A) = \Phi_A$. This proves that (Y, τ_Y, A) is a soft T_2-space. □

Theorem 3. Let (X, τ, A) be a soft topological space over X. If (X, τ, A) is a soft T_2-space and for any $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$, then there exist soft closed sets (F_1, A) and (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $e_H \tilde{\notin} (F_1, A)$ and $e_G \tilde{\notin} (F_2, A)$, $e_H \tilde{\in} (F_2, A)$, and $(F_1, A) \tilde{\cup} (F_2, A) = X_A$.

Proof. Since (X, τ, A) is a soft T_2-space and $e_G, e_H \tilde{\in} X_A$ such that $e_G \neq e_H$, there exist soft open sets (G_1, A) and (G_2, A) such that $e_G \tilde{\in} (G_1, A)$ and $e_H \tilde{\in} (G_2, A)$ and $(G_1, A) \tilde{\cap} (G_2, A) = \Phi_A$. Clearly $(G_1, A) \tilde{\subseteq} (G_2, A)^c$ and $(G_2, A) \tilde{\subseteq}$ $(G_1, A)^c$. Hence $e_G \tilde{\in} (G_2, A)^c$. Put $(G_2, A)^c = (F_1, A)$. This gives $e_G \tilde{\in} (F_1, A)$ and $e_H \tilde{\notin} (F_1, A)$. Also $e_H \tilde{\in} (G_1, A)^c$. Put $(G_1, A)^c = (F_2, A)$. Therefore $e_G \tilde{\in} (F_1, A)$ and $e_H \tilde{\in} (F_2, A)$. Moreover
$(F_1, A) \tilde{\cup} (F_2, A) = (G_2, A)^c \tilde{\cup} (G_1, A)^c = X_A$. □

4. Soft Regular, Soft Normal and Soft T_i; $(i = 4, 3)$ Spaces

In this section, we redefine soft regular and soft T_3 spaces using soft points and characterize soft regular and soft normal spaces. Moreover, we prove that soft regular and soft T_3 properties are soft hereditary, whereas soft normal and soft T_4 are soft closed

hereditary properties.
Now we define soft regular space as:
Definition 18. Let (X, τ, A) be a soft topological space over X, (G, A) a soft closed set in (X, τ, A) and $e_F \tilde{\in} X_A$ such that $e_F \tilde{\notin} (G, A)$. If there exist soft open sets (F_1, A) and (F_2, A) such that $e_F \tilde{\in} (F_1, A)$, $(G, A) \tilde{\subseteq} (F_2, A)$ and $(F_1, A) \tilde{\cap} (F_2, A) = \Phi_A$, then (X, τ, A) is called a soft regular space. □

In the following theorem, we give the characterizations of soft regular spaces.
Theorem 4. Let (X, τ, A) be a soft topological space over X. Then the following statements are equivalent:
(1) (X, τ, A) is soft regular.
(2) For any soft open set (F, A) in (X, τ, A) and $e_G \tilde{\in} (F, A)$, there is a soft open set (G, A) containing e_G such that $e_G \tilde{\in} \overline{(G, A)} \tilde{\subseteq} (F, A)$.
(3) Each soft point in (X, τ, A) has a soft nbd base consisting of soft closed sets.
Proof. (1) $\Rightarrow$ (2) Let (F, A) be a soft open set in (X, τ, A) and $e_G \tilde{\in} (F, A)$. Then $(F, A)^c$ is a soft closed set such that $e_G \tilde{\notin} (F, A)^c$. By the soft regularity of (X, τ, A), there are soft open sets (F_1, A), (F_2, A) such that $e_G \tilde{\in} (F_1, A)$, $(F, A)^c \tilde{\subseteq} (F_2, A)$ and $(F_1, A) \tilde{\cap} (F_2, A) = \Phi_A$. Clearly $(F_2, A)^c$ is a soft closed set contained in (F, A). Thus $(F_1, A) \tilde{\subseteq} (F_2, A)^c \tilde{\subseteq} (F, A)$. This gives $\overline{(F_1, A)} \tilde{\subseteq} (F_2, A)^c$
$\tilde{\subseteq} (F, A)$. Put $(F_1, A) = (G, A)$. Consequently, $e_G \tilde{\in} (G, A)$ and $\overline{(G, A)} \tilde{\subseteq} (F, A)$. This proves (2).
(2) $\Rightarrow$ (3) Let $e_G \tilde{\in} X_A$. For soft open set (F, A) in (X, τ, A), there is a soft open set (G, A) containing e_G such that $e_G \tilde{\in} (G, A)$, $\overline{(G, A)} \tilde{\subseteq} (F, A)$. Thus for each $e_G \tilde{\in} X_A$, the sets $\overline{(G, A)}$ form a soft nbd base consisting of soft closed sets of (X, τ, A). This proves (3).
(3) $\Rightarrow$ (1). Let (F, A) be a soft closed set such that $e_G \tilde{\notin} (F, A)$. Then $(F, A)^c$ is a soft open nbd of e_G. By (3), there is a soft closed set (F_1, A) which contains e_G and is a soft nbd of e_G with $(F_1, A) \tilde{\subseteq} (F, A)^c$. Then $e_G \tilde{\notin} (F_1, A)^c$, $(F, A) \tilde{\subseteq} (F_1, A)^c = (F_2, A)$ and $(F_1, A) \tilde{\cap} (F_2, A) = \Phi_A$. Therefore (X, τ, A) is soft regular. □.

The following theorem shows that soft regularity is a soft hereditary property:
Theorem 5. Let (X, τ, A) be a soft regular space over X. Then every soft subspace of (X, τ, A) is soft regular.
Proof . Let (Y, τ_Y, A) be a soft subspace of a soft regular space (X, τ, A). Suppose (H, A) is a soft closed set in (Y, τ_Y, A) and $e_F \tilde{\in} Y_A$ such that $e_F \tilde{\notin} (H, A)$. Then $(H, A) = (G, A) \tilde{\cap} Y_A$, where (G, A) is soft closed in (X, τ, A). Then $e_F \tilde{\notin} (G, A)$. Since (X, τ, A) is soft regular, there exist soft disjoint soft open sets (F_1, A), (F_2, A) in (X, τ, A) such that $e_F \tilde{\in} (F_1, A)$, $(G, A) \tilde{\subseteq} (F_2, A)$. Clearly $e_F \tilde{\in} (F_1, A) \tilde{\cap} Y_A = (F_{1_Y}, A)$ and $(H, A) \tilde{\subseteq} (F_2, A) \tilde{\cap} Y_A = (F_{2_Y}, A)$ such that $(F_{1_Y}, A) \tilde{\cap} (F_{2_Y}, A) = \Phi_A$. This proves that (Y, τ_Y, A) is a soft regular subspace of (X, τ, A). □

Next we give another characterization of soft regular spaces.:
Theorem 6. Let (X, τ, A) be a soft topological space over X. A space (X, τ, A) is soft regular if and only if for each $e_H \tilde{\in} X_A$ and a soft closed set (F, A) in (X, τ, A) such that $e_H \tilde{\notin} (F, A)$, there exist soft open sets (F_1, A), (F_2, A) in (X, τ, A) such that $e_H \tilde{\in} (F_1, A)$ and $(F, A) \tilde{\subseteq} (F_2, A)$ and $\overline{(F_1, A)} \tilde{\cap} \overline{(F_2, A)} = \Phi_A$.
Proof. For each $e_H \tilde{\in} X_A$ and a soft closed set (F, A) such that $e_H \tilde{\notin} (F, A)$, by Theorem 4(2), there is a soft open set (G, A) such that $e_H \tilde{\in} (G, A)$, $\overline{(G, A)} \tilde{\subseteq} (F, A)^c$. Again by Theorem 4(2), there is a soft open set (F_1, A) containing e_H such that $\overline{(F_1, A)} \tilde{\subseteq} (G, A)$. Let

$(F_2,A) = (\overline{(G,A)})^c$. Then $\overline{(F_1,A)}\tilde{\subseteq}(G,A)\tilde{\subseteq}\overline{(G,A)}\tilde{\subseteq}(F,A)^c$ implies $(F,A)\tilde{\subseteq}(\overline{(G,A)})^c = \overline{(F_2,A)}$ or $(F,A)\tilde{\subseteq}(F_2,A)$. Also $\overline{(F_1,A)}\tilde{\cap}\overline{(F_2,A)} = \overline{(F_1,A)}\tilde{\cap}\overline{(\overline{(G,A)})^c}\tilde{\subseteq}(G,A)\tilde{\cap}\overline{(\overline{(G,A)})^c}\tilde{\subseteq}\overline{(G,A)\tilde{\cap}(\overline{(G,A)})^c} = \overline{\Phi_A} = \Phi_A$.
Thus (F_1,A), (F_2,A) are the required soft open sets in (X,τ,A). This proves the necessity. The sufficiency is immediate. □

Definition 19. Let (X,τ,A) be a soft topological space over X. Then (X,τ,A) is said to be a soft T_3-space, if it is a soft regular and a soft T_1-space.

Remark 3. (1) A soft T_3-space may not be a soft T_2-space.
(2) If (X,τ,A) is a soft T_3-space, then (X,τ_e) may not be a T_3-space for each parameter $e\in A$.

The following proposition follows form Proposition 9 and Theorem 5.
Proposition 12. Let (X,τ,A) be a soft topological space over X and $Y\subseteq X$. If (X,τ,A) is a soft T_3-space then (Y,τ_Y,A) is a soft T_3-space.

The notions of soft normal and soft T_4 spaces have been introduced in [25]as:
Definition 20[25]. Let (X,τ,A) be a soft topological space over X, (F,A) and (G,A) soft closed sets over X such that $(F,A)\tilde{\cap}(G,A) = \Phi_A$. If there exist soft open sets (F_1,A) and (F_2,A) such that $(F,A)\tilde{\subseteq}(F_1,A)$, $(G,A)\tilde{\subseteq}(F_2,A)$ and $(F_1,A)\tilde{\cap}(F_2,A) = \Phi_A$, then (X,τ,A) is called a soft normal space.

Definition 21[25]. Let (X,τ,A) be a soft topological space over X. Then (X,τ,A) is said to be a soft T_4-space, if it is soft normal and soft T_1-space.

Now we prove the following theorem which characterizes soft normal spaces.:
Theorem 7. A soft topological space (X,τ,A) is soft normal if and only if for any soft closed set (F,A) and soft open set (G,A) such that $(F,A)\tilde{\subset}(G,A)$, there exists at least one soft open set (H,A)containing (F,A) such that

$$(F,A)\tilde{\subseteq}(H,A)\tilde{\subseteq}\overline{(H,A)}\tilde{\subseteq}(G,A).$$

Proof. Suppose that (X,τ,A) is a soft normal space and (F,A) is any soft closed subset of (X,τ,A) and (G,A) a soft open set such that $(F,A)\tilde{\subseteq}(G,A)$. Then $(G,A)^c$ is soft closed and $(F,A)\tilde{\cap}(G,A)^c = \Phi_A$. So by supposition, there are soft open sets (H,A) and (K,A) such that $(F,A)\tilde{\subseteq}(H,A)$, $(G,A)^c\tilde{\subseteq}(K,A)$ and $(H,A)\tilde{\cap}(K,A)$
$= \Phi_A$. Since $(H,A)\tilde{\cap}(K,A) = \Phi_A$, $(H,A)\tilde{\subseteq}(K,A)^c$. But $(K,A)^c$ is soft closed, so that $(F,A)\tilde{\subseteq}(H,A)\tilde{\subseteq}\overline{(H,A)}\tilde{\subseteq}(K,A)^c\tilde{\subseteq}(G,A)$. Hence $(F,A)\tilde{\subseteq}(H,A)\tilde{\subseteq}\overline{(H,A)}$
$\tilde{\subseteq}(G,A)$.
Conversely, suppose that for every soft closed set (F,A) and a soft open set (G,A) such that $(F,A)\tilde{\subseteq}(G,A)$, there is a soft open set (H,A) such that $(F,A)\tilde{\subseteq}(H,A)\tilde{\subseteq}\overline{(H,A)}\tilde{\subseteq}(G,A)$.
Let (F_1,A), (F_2,A) be any two soft disjoint soft closed sets. Then $(F_1,A)\tilde{\subseteq}(F_2,A)^c$, where $(F_2,A)^c$ is soft open. Hence there is a soft open set (H,A) such that $(F_1,A)\tilde{\subseteq}(H,A)\tilde{\subseteq}\overline{(H,A)}\tilde{\subseteq}(F_2,A)^c$.
But then $(F_2,A)\tilde{\subseteq}(\overline{(H,A)})^c$ and $(H,A)\tilde{\cap}(\overline{(H,A)})^c \neq \Phi$. Hence $(F_1,A)\tilde{\subseteq}(H,A)$, $(F_2,A)\tilde{\subseteq}(\overline{(H,A)})^c$ with $(H,A)\tilde{\cap}(\overline{(H,A)})^c = \Phi_A$. Hence (X,τ,A) is soft normal. This completes the proof.

The following proposition is easy to proof.
Proposition 13. Let (Y,τ_Y,A) be a soft subspace of a soft topological space (X,τ,A) and (F,A) be a soft open (closed) in (Y,τ_Y,A). If Y_A is soft open(closed) in (X,τ,A), then (F,A) is soft open(closed) in (X,τ,A).

Soft normality is a soft closed hereditary property as is proved in the following:
Theorem 8. A soft closed subspace of a soft normal space is soft normal.
Proof. Let (Y, τ_Y, A) be soft subspace of soft normal space (X, τ, A) such that $Y_A \in \tau^c$. Let (F_1, A), (F_2, A) be two disjoint soft closed subsets of (Y, τ_Y, A). Then there exists soft closed sets (F, A), (G, A) in (X, τ, A) such that $(F_1, A) = Y_A \tilde{\cap} (F, A)$ and $(F_2, A) = Y_A \tilde{\cap} (G, A)$. Since Y_A is soft closed in (X, τ, A), therefore (F_1, A), (F_2, A) are soft disjoint soft closed in (X, τ, A). Then (X, τ, A) is soft normal implies that there exist soft open sets (F_3, A), (F_4, A) in (X, τ, A) such that $(F_1, A) \tilde{\subseteq} (F_3, A)$, $(F_2, A) \tilde{\subseteq} (F_4, A)$ and $(F_3, A) \tilde{\cap} (F_4, A) = \Phi_A$. But then $(F_1, A) \tilde{\subseteq} Y_A \tilde{\cap} (F_3, A)$, $(F_2, A) \tilde{\subseteq} Y_A \tilde{\cap} (F_4, A)$, where $Y_A \tilde{\cap} (F_3, A)$, $Y_A \tilde{\cap} (F_4, A)$ are soft disjoint soft open subsets of (Y, τ_Y, A). This proves that (Y, τ_Y, A) is soft normal. Hence the proof.

The following corollary directly follows from proposition 9 and Theorem 8. **Corollary 1.** Every soft closed subspace of a soft T_4-space is a soft T_4-space.

Conclusion : The study of soft sets and soft topology is very important in the study of possible applications in classical and non classical logic. We redefined and explored soft separation axioms, namely soft T_i, $i = 0, 1, 2$, soft regular, soft T_3, soft normal and soft soft T_4 axioms using soft point defined by I. Zorlutuna [30]. We also discussed some soft invariance properties namely soft topological property and soft hereditary property . These soft separation axioms would be useful for the development of the theory of soft topology to solve the complicated problems containing uncertainties in economics, engineering, medical, environment and in general man-machine systems of various types. These findings are the addition for strengthening the toolbox of soft topology.

References

[1] U. Acar, F. Koyuncu, B. Tanay: *Soft sets and soft rings, Computers and Mathematics with Applications* **59** (2010) 3458-3463.

[2] H. Aktas, N. Cagman: *Soft sets and soft groups, Information Sciences* **1(77)** (2007) 2726-2735.

[3] K. Atanassov: *Intuitionistic fuzzy sets, Fuzzy Sets and Systems* **20**(1986) 87-96.

[4] K. Atanassov: *Operators over interval valued intuitionistic fuzzy sets, Fuzzy Sets and Systems* **64** (1994) 159-174.

[5] N. Cagman, S. Karatas, S. Enginoglu: *Soft Topology, Computers and Mathematics with Applications* **62** (2011) 351-358.

[6] D. Chen: *The parametrization reduction of soft sets and its applications, Computers and Mathematics with Applications* **49** (2005) 757-763.

[7] D. N. Georgiou, A. C. Megaritis, V. I. Petropoulos: *On Soft Topological Spaces, Applied Mathematics and Information Sciences* **7(2)** (2013) 1889-1901.

[8] M. B. Gorzalzany: *A method of inference in approximate reasoning based on interval-valued fuzzy sets, Fuzzy Sets and Systems* **21** (1987) 1-17.

[9] S. Hussain, B. Ahmad : *Some Properties of Soft Topological Spaces, Computers and Mathematics with Applications* **62** (2011) 4058-4067.

[10] Y. B. Jun: *Soft BCK/BCI-algebras, Computers and Mathematics with Applications* **56** (2008) 1408-1413.

[11] Y. B. Jun, C. H. Park: *Applications of soft sets in ideal theory of BCK/BCI-algebras, Information Sciences* **178** (2008) 2466-2475.

[12] Y. B. Jun, C. H. Park: *Applications of soft sets in Hilbert algebras, Iranian Journal Fuzzy Systems* **6(2)** (2009) 55-86.

[13] Y. B. Jun, H. S. Kim, J. Neggers: *Pseudo d-algebras, Information Sciences* **179** (2009) 1751-1759.

[14] Y. B. Jun, K. J. Lee, A. Khan: *Soft ordered semigroups, Mathematical Logic Quarterly* **56(1)** (2010) 42-50.
[15] Y. B. Jun, K. J. Lee, C. H. Park: *Soft set theory applied to commutative ideals, in BCK-algebras, Journal of Applied Mathematics Informatics* **26(3-4)** (2008) 707-720.
[16] Y. B. Jun, K. J. Lee, C. H. Park: *Soft set theory applied to ideals in d-algebras, Computers and Mathematics with Applications* **57** (2009) 367-378.
[17] Y. B. Jun, K. J. Lee, C. H. Park: *Fuzzy soft set theory applied to BCK/BCI-algebras , Computers and Mathematics with Applications* **59** (2010) 3180-3192.
[18] Y. B. Jun, C. H. Park: *Applications of soft sets in ideal theory of BCK/BCI-Algebras, Inform. Sci.* **178** (2008) 2466-2475.
[19] A. Kharral , B. Ahmad : *Mappings on soft classes, New Mathematics and Natural Computations,* **7(3)**(2011), 471–481.
[20] P. K. Maji, R. Biswas, R. Roy: *An application of soft sets in a decision making problem, Computers and Mathematics with Applications* **44** (2002) 1077-1083.
[21] P. K. Maji R. Biswas, R. Roy: *Soft set theory, Computers and Mathematics with Applications* **45** (2003) 555-562.
[22] D. Molodtsov: *Soft set theory - first results, Computers and Mathematics with Applications* **37** (1999)19-31.
[23] D. Molodtsov, V. Y. Leonov, D. V. Kovkov: *Soft sets technique and its application, Nechetkie Sistemy i Myagkie Vychisleniya* **1(1)** (2006) 8-39.
[24] Z. Pawlak: *Rough sets, Int. J. Comput. Sci.* **11** (1982) 341-356.
[25] D. Pie, D. Miao: *From soft sets to information systems, Granular Computing, 2005 IEEE Inter. Conf.* **2** , 617-621.
[26] C. H. Park, Y. B. Jun, M. A. Ozturk: *Soft WS-algebras, Communications of the Korean Mathematical Society* **23(3)** (2008) 313-324.
[27] M. Shabir, M. Naz: *On Soft Topological Spaces, Computers and Mathematics with Applications* **61** (2011) 1786-1799.
[28] Q. M. Sun, Z. L. Zhang, J. Liu: *Soft sets and soft modules, Proceedings of Rough Sets and Knowledge Technology, Third International Conference, RSKT 2008, 17-19 May, Chengdu, China,* pp. 403-409, 2008.
[29] L. A. Zadeh: *Fuzzy sets, Information and Control* **8** (1965) 338-353.
[30] I. Zorlutana, N. Akdag, W. K. Min, S. Atmaca: *Remarks on Soft Topological Spaces, Annals of Fuzzy Mathematics and Informatics* **3(2)** (2012) 171-185.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 569 – 586

Fuzzy ideals in right regular LA-semigroups

Madad Khan [*] , Young Bae Jun [†] and Faisal Yousafzai [‡]

Abstract

In this paper, we have discussed fuzzy left (right, two-sided) ideals, fuzzy (generalized) bi-ideals, fuzzy interior ideals, fuzzy $(1, 2)$-ideals and fuzzy quasi-ideals of a right regular LA-semigroup. Moreover we have characterized a right regular LA-semigroup in terms of their fuzzy left and fuzzy right ideals.

2000 AMS Classification: 20M10, 20N99.

Keywords: Right regular LA-semigroups, left invertive law and fuzzy ideals.

Received 14/10/2011 : Accepted 21/02/2014 Doi : 10.15672/HJMS.2015449419

1. Introduction

A fuzzy subset (fuzzy set) of a non-empty set $\mathcal{S}$ is an arbitrary mapping $f : \mathcal{S} \to [0, 1]$, where $[0, 1]$ is the unit segment of the real line. A fuzzy subset is a class of objects with a grades of membership. This important concept of fuzzy sets was first proposed by Zadeh [13] in 1965. Since then, many papers on fuzzy sets appeared which shows its importance and applications to set theory, group theory, groupoids, real analysis, measure theory and topology etc. In one of the recent paper, Zadeh introduced a new idea to explore the relationship between probabilities and fuzzy sets [14].

Rosenfeld [12] was the first who consider the case when $\mathcal{S}$ is a groupoid. He gave the definition of fuzzy subgroupoid and the fuzzy left (right, two-sided) ideal of $\mathcal{S}$ and justified these definitions by showing that a subset $\mathcal{A}$ of a groupoid $\mathcal{S}$ is a subgroupoid or a left (right, two-sided) ideal of $\mathcal{S}$ if the characteristic function of $\mathcal{A}$, that is

$$C_{\mathcal{A}}(x) = \begin{cases} 1, \text{ if } x \in \mathcal{A} \\ 0, \text{ if } x \notin \mathcal{A} \end{cases}$$

[*]Department of Mathematics COMSATS Institute of Information Technology Abbottabad, Pakistan
Email: madadmath@yahoo.com

[†]Department of Mathematics Education Gyeongsang National University Chinju, Korea
Email:skywine@gmail.com

[‡]Department of Mathematics COMSATS Institute of Information Technology Abbottabad, Pakistan
Email:yousafzaimath@yahoo.com

is a fuzzy subgroupoid or a fuzzy left (right, two-sided) ideal of $\mathcal{S}$.

Kuroki and Mordeson have widely explored fuzzy semigroups in [5] and [6].

Fuzzy algebra is going popular day by day due to wide applications of fuzzification in almost every field. Our aim in this paper is to develop some characterizations for a new non-associative algebraic structure known as a left almost semigroup (LA-semigroup in short) which is the generalization of a commutative semigroup (see [2]). An LA-semigroup is an algebraic structure mid way between a groupoid and a commutative semigroup. An LA-semigroup has wide range of applications in theory of flocks (see [9]).

The concept of a left almost semigroup [2] was first introduced by M. A. Kazim and M. Naseeruddin in 1972. A groupoid $\mathcal{S}$ is called an LA-semigroup if it satisfy the following left invertive law

(1) $\quad (ab)c = (cb)a$, for all $a,\ b,\ c \in \mathcal{S}$.

An LA-semigroup is also known as an Abel-Grassmann's groupoid (AG-groupoid) [10]. P. Holgate called it left invertive groupoid [1].

In an LA-semigroup $\mathcal{S}$, the medial law [2] holds

(2) $\quad (ab)(cd) = (ac)(bd)$, for all $a,\ b,\ c,\ d \in \mathcal{S}$.

The left identity in an LA-semigroup if exists is unique [7]. Every LA-semigroup with left identity satisfy the following laws

(3) $\quad (ab)(cd) = (dc)(ba),\ $ for all $a, b, c, d \in \mathcal{S}$.

(4) $\quad a(bc) = b(ac),\ $ for all $a, b, c \in \mathcal{S}$.

If an LA-semigroup S contains left identity e then $S = eS \subseteq S^2$. Therefore $S = S^2$.

An LA-semigroup is closely related with a commutative semigroup, because if it contains a right identity, then it becomes a commutative semigroup [7].

Define the binary operation "$\bullet$" on a commutative inverse semigroup S as

$$a \bullet b = ba^{-1},\ \text{for all } a, b \in S,$$

then $(S, \bullet)$ becomes an LA-semigroup [8].

An LA-semigroup $(S, \cdot)$ becomes a semigroup S under new binary operation "$\circ$" defined in [11] as

$$x \circ y = (xa)y, \text{ for all } x, y \in S.$$

It is easy to show that "$\circ$" is associative

$$\begin{aligned}(x \circ y) \circ z &= (((xa)y)a)z = (za)((xa)y) = (xa)((za)y) \\ &= (xa)((ya)z) = x \circ (y \circ z).\end{aligned}$$

Connections discussed above make this non-associative structure interesting and useful.

Here we have given some examples of LA-semigroups in terms of abelian groups.

1.1. Example. Let us consider the abelian group $(\mathbb{R}, +)$ of all real numbers under the binary operation of addition. If we define

$$a * b = b - a - r,\ \text{where } a, b, r \in \mathbb{R},$$

then $(\mathbb{R}, *)$ becomes an LA-semigroup. Indeed

$$(a * b) * c = c - (a * b) - r = c - (b - a - r) - r = c - b + a + r - r = c - b + a,$$

and

$$(c * b) * a = a - (c * b) - r = a - (b - c - r) - r = a - b + c + r - r = a - b + c.$$

Since $(\mathbb{R}, +)$ is commutative, so $(a * b) * c = (c * b) * a$ and therefore $(\mathbb{R}, *)$ satisfies a left invertive law. It is easy to observe that $(\mathbb{R}, *)$ is non-commutative and non-associative.

The same is hold for set of integers and rationals. Thus $(\mathbb{R}, *)$ is an LA-semigroup which is the generalization of an LA-semigroup given in [8].

1.2. Example. Consider the abelian group $(\mathbb{R}\backslash\{0\}, .)$ of all real numbers except zero under the binary operation of multiplication. If we define

$$a * b = ba^{-1}r^{-1}, \text{ where } a, b, r \in \mathbb{R},$$

then $(\mathbb{R}\backslash\{0\}, *)$ becomes an LA-semigroup. Indeed

$$(a * b) * c = ba^{-1}r^{-1} * c = c(ba^{-1}r^{-1})^{-1}r^{-1} = crab^{-1}r^{-1} = cab^{-1},$$

and

$$(c * b) * a = bc^{-1}r^{-1} * a = a(bc^{-1}r^{-1})^{-1}r^{-1} = arcb^{-1}r^{-1} = acb^{-1}.$$

As $(\mathbb{R}\backslash\{0\}, .)$ is commutative, therefore $(a * b) * c = (c * b) * a$ and thus $(\mathbb{R}, *)$ satisfies a left invertive law. Clearly $(\mathbb{R}, *)$ is non-commutative and non-associative. The same is hold for set of integers and rationals. This LA-semigroup is also the generalization of an LA-semigroup given in [8].

2. Preliminaries

Let $\mathcal{S}$ be an LA-semigroup, by an LA-subsemigroup of $\mathcal{S}$, we means a non-empty subset $\mathcal{A}$ of $\mathcal{S}$ such that $\mathcal{A}^2 \subseteq \mathcal{A}$.

A non-empty subset $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called a left (right) ideal of $\mathcal{S}$ if $\mathcal{S}\mathcal{A} \subseteq \mathcal{A}$ $(\mathcal{A}\mathcal{S} \subseteq \mathcal{A})$.

A non-empty subset $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called a two-sided ideal or simply an ideal if it is both a left and a right ideal of $\mathcal{S}$.

A non-empty subset $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called a generalized bi-ideal of $\mathcal{S}$ if $(\mathcal{A}\mathcal{S})\mathcal{A} \subseteq \mathcal{A}$.

An LA-subsemigroup $\mathcal{A}$ of $\mathcal{S}$ is called a bi-ideal of $\mathcal{S}$ if $(\mathcal{A}\mathcal{S})\mathcal{A} \subseteq \mathcal{A}$.

A non-empty subset $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called an interior ideal of $\mathcal{S}$ if $(\mathcal{S}\mathcal{A})\mathcal{S} \subseteq \mathcal{A}$.

A non-empty subset $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called a quasi ideal of $\mathcal{S}$ if $\mathcal{S}\mathcal{A} \cap \mathcal{A}\mathcal{S} \subseteq \mathcal{A}$.

An LA-subsemigroup $\mathcal{A}$ of an LA-semigroup $\mathcal{S}$ is called a $(1, 2)$-ideal of $\mathcal{S}$ if $(\mathcal{A}\mathcal{S})\mathcal{A}^2 \subseteq \mathcal{A}$.

The following definitions are available in [6].

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy LA-subsemigroup of $\mathcal{S}$ if $f(xy) \geq f(x) \wedge f(y)$ for all x, $y \in \mathcal{S}$.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy left (right) ideal of $\mathcal{S}$ if $f(xy) \geq f(y)$ $(f(xy) \geq f(x))$ for all x, $y \in \mathcal{S}$.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy two-sided ideal of $\mathcal{S}$ if it is both a fuzzy left and a fuzzy right ideal of $\mathcal{S}$.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy generalized bi-ideal of $\mathcal{S}$ if $f((xa)y) \geq f(x) \wedge f(y)$, for all x, a and $y \in \mathcal{S}$.

A fuzzy LA-subsemigroup f of an LA-semigroup $\mathcal{S}$ is called a fuzzy bi-ideal of $\mathcal{S}$ if $f((xa)y) \geq f(x) \wedge f(y)$, for all x, a and $y \in \mathcal{S}$.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy interior ideal of $\mathcal{S}$ if $f((xa)y) \geq f(a)$, for all x, a and $y \in \mathcal{S}$.

Characteristic function of an LA-semigroup $\mathcal{S}$ is denoted by $C_{\mathcal{S}}(x)$ and defined as $C_{\mathcal{S}}(x) = 1$ for all x in $\mathcal{S}$.

Note that for any two fuzzy subsets f and $\mathcal{S}$ of $\mathcal{S}$, $f \subseteq \mathcal{S}$ means that $f(x) \leq C_{\mathcal{S}}(x)$ for all x in $\mathcal{S}$.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy quasi-ideal of $\mathcal{S}$ if $(f \circ C_{\mathcal{S}}(x)) \cap (C_{\mathcal{S}}(x) \circ f) \subseteq f$.

A fuzzy LA-subsemigroup f of an LA-semigroup $\mathcal{S}$ is called a fuzzy $(1,2)$-ideal of $\mathcal{S}$ if $f((xa)(yz)) \geq f(x) \wedge f(y) \wedge f(z)$ for all x, a, y and $z \in \mathcal{S}$.

Let f and g be any two fuzzy subsets of an LA-semigroup $\mathcal{S}$, then the product $f \circ g$ is defined by,

$$(f \circ g)(a) = \begin{cases} \bigvee\limits_{a=bc} \{f(b) \wedge g(c)\}, & \text{if there exist } b, c \in \mathcal{S}, \text{ such that } a = bc. \\ 0, & \text{otherwise.} \end{cases}$$

The symbols $f \cap g$ and $f \cup g$ will means the following fuzzy subsets of $\mathcal{S}$

$$(f \cap g)(x) = \min\{f(x),\ g(x)\} = f(x) \wedge g(x), \text{ for all } x \text{ in } \mathcal{S}$$

and

$$(f \cup g)(x) = \max\{f(x),\ g(x)\} = f(x) \vee g(x), \text{ for all } x \text{ in } \mathcal{S}.$$

For a fuzzy subset f of an LA-semigroup $\mathcal{S}$ and $\alpha \in (0,1]$, the set $f_\alpha = \{x \in \mathcal{S} : f(x) \geq \alpha\}$ is called a level cut of f.

A fuzzyleft ideal f is called idempotent if $f \circ f = f$.

2.1. Example. Let $\mathcal{S} = \{a, b, c, d, e\}$ be an LA-semigroup with left identity d with the following multiplication table.

.	a	b	c	d	e
a	a	a	a	a	a
b	a	b	b	b	b
c	a	b	d	e	c
d	a	b	c	d	e
e	a	b	e	c	d

Note that $\mathcal{S}$ is non-commutative as $ed \neq de$ and also $\mathcal{S}$ is non-associative because $(cc)d \neq c(cd)$.

Define a fuzzy subset f of $\mathcal{S}$ as follows: $f(a) = 1$ and $f(b) = f(c) = f(d) = f(e) = 0$, then clearly f is a fuzzy two-sided ideal of $\mathcal{S}$.

It is easy to see that every fuzzy left (right, two-sided) ideal of an LA-semigroup $\mathcal{S}$ is a fuzzy LA-subsemigroup of $\mathcal{S}$ but the converse is not true in general. Let us define a fuzzy subset f of $\mathcal{S}$ as follows: $f(a) = 1$, $f(b) = 0$ and $f(c) = f(d) = f(e) = 0.5$, then by routine calculation one can easily check that f is a fuzzy LA-subsemigroup of $\mathcal{S}$ but it is not a fuzzy left (right, two-sided) ideal of $\mathcal{S}$ because $f(bd) \ngeq f(d)$ or $f(db) \ngeq f(d)$.

2.2. Theorem. *For an LA-semigroup $\mathcal{S}$, the following statements are true.*

(i) f_α is a right (left, two-sided) ideal of $\mathcal{S}$ if f is a fuzzy right (left) ideal of $\mathcal{S}$.

(ii) f_α is a bi-(generalized bi-) ideal of $\mathcal{S}$ if f is a fuzzy bi-(generalized bi-) ideal of $\mathcal{S}$.

Proof. (i): Let $\mathcal{S}$ be an LA-semigroup and let f be a fuzzy right ideal of $\mathcal{S}$. If $x, y \in \mathcal{S}$ such that $x \in f_\alpha$, then $f(x) \geq \alpha$ and therefore $f(xy) \geq f(x) \geq \alpha$ implies that $xy \in f_\alpha$. This shows that f_α is a right ideal of $\mathcal{S}$. If f is a fuzzy left ideal of $\mathcal{S}$, then $f(yx) \geq f(x) \geq \alpha$ implies that $yx \in f_\alpha$. This shows that f_α is a left ideal of $\mathcal{S}$.

(ii): Let $\mathcal{S}$ be an LA-semigroup and let f be a fuzzy bi-(generalized bi-) ideal of $\mathcal{S}$. If x, y and $z \in \mathcal{S}$ such that x and $z \in f_\alpha$, then $f(x) \geq \alpha$ and $f(z) \geq \alpha$, therefore $f((xy)z) \geq f(x) \wedge f(z) \geq \alpha$ implies that $(xy)z \in f_\alpha$. Which shows that f_α is a generalized bi ideal of $\mathcal{S}$. Now let $x, y \in f_\alpha$, then $f(x) \geq \alpha$ and $f(y) \geq \alpha$ and therefore $f(xy) \geq f(x) \wedge f(y) \geq \alpha$ implies that $xy \in f_\alpha$. Thus f_α is a bi ideal of $\mathcal{S}$. ∎

Note that the converses of (i) and (ii) are not true in general. Define a fuzzy subset f of an LA-semigroup $\mathcal{S}$ in Example 2.1 as follows: $f(a) = 0.2$, $f(b) = 0.9$, $f(c) = f(d) =$

$f(e) = 0$. Let $\alpha = 0.2$, then it is easy to see that $f_\alpha = \{a, b\}$ and one can easily verify from Example 2.1 that $\{a, b\}$ is a right (left, generalized bi-, bi-) ideal of $\mathcal{S}$ but $f(ba) \ngeq f(b)$ ($f(ab) \ngeq f(b)$, $f((ba)b) \ngeq f(b)$) implies that f is not a fuzzy right (left, generalized bi-, bi) ideal of $\mathcal{S}$.

2.3. Lemma. [3] *Every fuzzy right ideal of an LA-semigroup $\mathcal{S}$ with left identity becomes a fuzzy left ideal of $\mathcal{S}$.*

Note that the converse of the above is not true in general. If we define a fuzzy subset f of an LA-semigroup $\mathcal{S}$ in Example 3.1 as follows: $f(a) = 0.8, f(b) = 0.5, f(c) = 0.4, f(d) = 0.3$ and $f(e) = 0.6$, then it is easy to observe that f is a fuzzy left ideal of $\mathcal{S}$ but it is not a fuzzy right ideal of $\mathcal{S}$, because $f(bd) \ngeq f(b)$.

Assume that $\mathcal{S}$ is an LA-semigroup and let $F(\mathcal{S})$ denote the set of all fuzzy subsets of $\mathcal{S}$, then $(F(\mathcal{S}), \circ)$ is an LA-semigroup and satisfies all the basic laws of an LA-semigroup [3].

2.4. Lemma. [3] *For a fuzzy subset f of an LA-semigroup $\mathcal{S}$, the following conditions are true.*

(i) f is a fuzzy left (right) ideal of $\mathcal{S}$ if and only if $C_\mathcal{S}(x) \circ f \subseteq f$ ($f \circ C_\mathcal{S}(x) \subseteq f$).
(ii) f is a fuzzy LA-subsemigroup of $\mathcal{S}$ if and only if $f \circ f \subseteq f$.

2.5. Lemma. [3] *For any non-empty subsets $\mathcal{A}$ and $\mathcal{B}$ of an LA-semigroup $\mathcal{S}$, the following conditions are true.*

(i) $C_\mathcal{A} \circ C_\mathcal{B} = C_{\mathcal{A}\mathcal{B}}$
(ii) $C_\mathcal{A} \cap C_\mathcal{B} = C_{\mathcal{A}\cap \mathrm{B}}$

2.6. Lemma. [3] *Let $\mathcal{A}$ be a non-empty subset of an LA-semigroup $\mathcal{S}$. Then the following properties holds.*

(i) $\mathcal{A}$ is an LA-subsemigroup of $\mathcal{S}$ if and only if $C_\mathcal{A}$ is a fuzzy LA-subsemigroup of $\mathcal{S}$.
(ii) $\mathcal{A}$ is a left (right, two-sided) ideal of $\mathcal{S}$ if and only if $C_\mathcal{A}$ is a fuzzy left (right, two-sided) ideal of $\mathcal{S}$.

3. Fuzzy ideals in Right Regular LA-semigroups

An element a of an LA-semigroup $\mathcal{S}$ is called a right regular if there exists $x \in \mathcal{S}$ such that $a = a^2x$ and $\mathcal{S}$ is called right regular if every element of $\mathcal{S}$ is right regular.

An LA-semigroup considered in Example 2.1 is right regular because, $a = a^2d$, $b = b^2c$, $c = c^2c$, $d = d^2d$, $e = e^2e$.

Note that in an LA-semigroup $\mathcal{S}$ with left identity, $\mathcal{S} = \mathcal{S}^2$.

3.1. Example. Let us consider an LA-semigroup $\mathcal{S} = \{a, b, c, d, e\}$ with left identity d in the following Cayley's table.

.	a	b	c	d	e
a	a	a	a	a	a
b	a	e	e	c	e
c	a	e	e	b	e
d	a	b	c	d	e
e	a	e	e	e	e

Note that $\mathcal{S}$ is not right regular because for $c \in \mathcal{S}$, there does not exists $x \in \mathcal{S}$ such that $c = c^2x$.

Note that if f is any fuzzy subset of an LA-semigroup $\mathcal{S}$ with left identity then $\mathcal{S}$ is right regular if $f(x) = f(x^2)$ holds for all x in $\mathcal{S}$. But the converse is not true in general.

3.2. Example. Let $\mathcal{S} = \{a, b, c, d, e\}$ be a right regular LA-semigroup with left identity d in the following multiplication table.

.	a	b	c	d	e
a	b	a	a	a	a
b	a	b	b	b	b
c	a	b	c	d	e
d	a	b	e	c	d
e	a	b	d	e	c

Let us consider a right regular LA-semigroup $\mathcal{S}$ in Example 3.2. Define a fuzzy subset f of $\mathcal{S}$ as follows: $f(a) = 0.6$, $f(b) = 0.2$ and $f(c) = f(d) = f(e) = 0.9$, then it is easy to see that $f(a) \neq f(a^2)$ for $a \in \mathcal{S}$.

3.3. Lemma. *If f is a fuzzy interior ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity, then $f(ab) = f(ba)$ holds for all a, b in $\mathcal{S}$.*

Proof. Assume that f is a fuzzy interior ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a \in \mathcal{S}$, then $a = a^2x$ for some x in $\mathcal{S}$. Now by using (1), (4) and (3), we have

$$f(a) = f((aa)x) = f((xa)a) = f((xa)((aa)x)) = f((aa)((xa)x))$$

$$\begin{aligned} &= f((ea^2)((xa)x)) \geq f(a^2) = f(aa) = f(a((aa)x)) \\ &= f((aa)(ax)) = f((xa)(aa)) = f((xa)a^2) \geq f(a). \end{aligned}$$

Which implies that $f(a) = f(a^2)$ for all a in $\mathcal{S}$.
Now by using (3), (4) and (2), we have

$$\begin{aligned} f(ab) &= f((ab)^2) = f((ab)(ab)) = f((ba)(ba)) \\ &= f((e(ba))(ba)) \geq f(ba) = f(b((aa)x)) \\ &= f((aa)(bx)) = f((ab)(ax)) \\ &= f((e(ab))(ax)) \geq f(ab). \end{aligned}$$

∎

The converse is not true in general. For this, let us define a fuzzy subset f of a right regular LA-semigroup $\mathcal{S}$ in Example 2.1 as follows: $f(a) = 0.1$, $f(b) = 0.2, f(c) = 0.6, f(d) = 0.4$ and $f(e) = 0.6$, then it is easy to see that $f(ab) = f(ba)$ holds for all a and b in $\mathcal{S}$ but f is not a fuzzy interior ideal of $\mathcal{S}$ because $f((ab)c) \ngeq f(b)$.

3.4. Lemma. *For any fuzzy subset f of a right regular LA-semigroup $\mathcal{S}$, $C_{\mathcal{S}}(x) \circ f = f$.*

Proof. Since S is right regular, therefore for each a in S there exists x such that $a = a^2x$, now using left invertive law, we get $a = (xa)a$. Then

$$C_{\mathcal{S}}(x) \circ f(a) = \bigvee_{a=(xa)a} \{C_{\mathcal{S}}(x)(xa) \wedge f(a)\} = \bigvee_{a=(xa)a} \{1 \wedge f(a) = f(a).$$

Hence $C_{\mathcal{S}}(x) \circ f = f$. ∎

3.5. Lemma. *In a right regular LA-semigroup $\mathcal{S}$, $f \circ C_{\mathcal{S}}(x) = f$ and $C_{\mathcal{S}}(x) \circ f = f$ holds for every fuzzy two-sided ideal f of $\mathcal{S}$.*

Proof. Let $\mathcal{S}$ be a right regular LA-semigroup. Now for every $a \in \mathcal{S}$ there exists $x \in \mathcal{S}$ such that $a = a^2x$. Then by using (1), we have $a = (aa)x = (xa)a$, therefore

$$\begin{aligned}(f \circ C_{\mathcal{S}}(x))(a) &= \bigvee_{a=(xa)a} \{f(xa) \wedge C_{\mathcal{S}}(x)(a)\} \geq f(xa) \wedge C_{\mathcal{S}}(x)(a) \\ &\geq f(a) \wedge 1 = f(a).\end{aligned}$$

It is easy to observe from Lemma 3.4 that $C_{\mathcal{S}}(x) \circ f = f$ holds for every fuzzy two-sided ideal f of $\mathcal{S}$. ∎

3.6. Corollary. *In a right regular LA-semigroup* $\mathcal{S}$, $C_{\mathcal{S}}(x) \circ C_{\mathcal{S}}(x) = C_{\mathcal{S}}(x)$.

Proof. It is simple. ∎

3.7. Lemma. *A fuzzy subset f of a right regular LA-semigroup $\mathcal{S}$ is a fuzzy left ideal of $\mathcal{S}$ if and only if it is a fuzzy right ideal of $\mathcal{S}$.*

Proof. Assume that f is a fuzzy left ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a, b \in \mathcal{S}$, then $a = a^2x$ for some x in $\mathcal{S}$. Now by using (1), we have

$$f(ab) = f((a^2x)b) = f((bx)a^2) \geq f(a^2) = f(aa) \geq f(a).$$

This shows that$f$ is a fuzzy right ideal of $\mathcal{S}$.

Similarly we can show that every fuzzy right ideal of $\mathcal{S}$ is a fuzzy left ideal of $\mathcal{S}$. ∎

3.8. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$.
(ii) f is a fuzzy two-sided ideal of $\mathcal{S}$.

Proof. $(i) \Longrightarrow (ii)$: Assume that f is a fuzzy $(1, 2)$-ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a \in \mathcal{S}$, then there exists $y \in \mathcal{S}$ such that $a = a^2y$. Now by using (4), (1) and (3), we have

$$\begin{aligned}f(xa) &= f(x((aa)y)) = f((aa)(xy)) = f((((aa)y)a)(xy)) \\ &= f(((ay)(aa))(xy)) = f(((aa)(ya))(xy)) \\ &= f(((xy)(ya))(aa)) = f(((ay)(yx))a^2) \\ &= f(((((aa)y)y)(yx))a^2) = f((((yy)(aa))(yx))a^2) \\ &= f((((aa)y^2)(yx))a^2) = f((((yx)y^2)(aa))a^2) \\ &= f((a(((yx)y^2)a))(aa)) \geq f(a) \wedge f(a) \wedge f(a) = f(a).\end{aligned}$$

This shows that f is a fuzzy left ideal of $\mathcal{S}$ and by using Lemma 3.7, f is a fuzzy two-sided ideal of $\mathcal{S}$.

$(ii) \Longrightarrow (i)$ is obvious. ∎

3.9. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$.
(ii) f is a fuzzy quasi ideal of $\mathcal{S}$.

Proof. $(i) \Longrightarrow (ii)$ is an easy consequence of Theorem 3.8 and Lemma 3.5.

$(ii) \Longrightarrow (i)$: Assume that f is a fuzzy quasi ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a \in \mathcal{S}$, then there exists $x \in \mathcal{S}$ such that $a = a^2x$. Now by using $(1), (3)$ and (4), we have

$$a = (aa)x = (xa)a = (xa)(ea) = (ae)(ax) = a((ae)x),$$

therefore

$$(f \circ C_{\mathcal{S}}(x))(a) = \bigvee_{a=a((ae)x)} \{f(a) \wedge C_{\mathcal{S}}(x)((ae)x)\} \geq f(a) \wedge 1 = f(a) \ .$$

Now by using Lemmas 3.4, 3.6 and (2), we have

$$\begin{aligned} f \circ C_{\mathcal{S}}(x) &= (C_{\mathcal{S}}(x) \circ f) \circ (C_{\mathcal{S}}(x) \circ C_{\mathcal{S}}(x)) = (C_{\mathcal{S}}(x) \circ C_{\mathcal{S}}(x)) \circ (f \circ C_{\mathcal{S}}(x)) \\ &= C_{\mathcal{S}}(x) \circ (f \circ C_{\mathcal{S}}(x)) \supseteq C_{\mathcal{S}}(x) \circ f. \end{aligned}$$

Which shows that $C_{\mathcal{S}}(x) \circ f \subseteq (f \circ C_{\mathcal{S}}(x)) \cap (C_{\mathcal{S}}(x) \circ f)$. As f is a fuzzy quasi ideal of $\mathcal{S}$, thus we get $C_{\mathcal{S}}(x) \circ f \subseteq f$. Now by using Lemmas 2.4 and 3.7, f is a fuzzy two-sided ideal of $\mathcal{S}$. Thus by Theorem 3.8, f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$. ∎

3.10. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy bi-ideal of $\mathcal{S}$.
(ii) f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$.

Proof. $(i) \Longrightarrow (ii)$: Assume that $\mathcal{S}$ is a right regular LA-semigroup with left identity and let $x, a, y, z \in \mathcal{S}$, then there exists $x' \in \mathcal{S}$ such that $x = x^2x'$. Let f be a fuzzy bi-ideal of $\mathcal{S}$, then by using $(3), (1)$ and (4), we have

$$\begin{aligned} f((xa)(yz)) &= f((zy)(ax)) = f(((ax)y)z) \\ &\geq f((ax)y) \wedge f(z) \geq f(ax) \wedge f(y) \wedge f(z) \\ &= f(a((xx)x')) \wedge f(y) \wedge f(z) \\ &= f((xx)(ax')) \wedge f(y) \wedge f(z) \\ &= f(((ax')x)x) \wedge f(y) \wedge f(z) \\ &= f(((ax')((xx)x'))x) \wedge f(y) \wedge f(z) \\ &= f(((ax')((xx)(ex')))x) \wedge f(y) \wedge f(z) \\ &= f(((ax')((x'e)(xx)))x) \wedge f(y) \wedge f(z) \\ &= f(((ax')(x((x'e)x)))x) \wedge f(y) \wedge f(z) \\ &= f(x((ax')((x'e)x))x) \wedge f(y) \wedge f(z) \\ &\geq f(x) \wedge f(x) \wedge f(y) \wedge f(z) \\ &= f(x) \wedge f(y) \wedge f(z). \end{aligned}$$

Which shows that f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$.

$(ii) \Longrightarrow (i)$: Again let $\mathcal{S}$ be a right regular LA-semigroup with left identity, then for any a, b, x and $y \in \mathcal{S}$ there exist a', b', $x'$and $y' \in \mathcal{S}$ such that $a = a^2a'$, $b = b^2b'$, $x = x^2x'$and $y = y^2y'$. Let f be a fuzzy $(1, 2)$-ideal of $\mathcal{S}$, then by using (4) and (1), we have

$$\begin{aligned} f((xa)y) &= f(xa)((yy)y') = (yy)((xa)y') = (y'(xa))(yy) \\ &= (x(y'a))(yy) \geq f(x) \wedge f(y) \wedge f(y) \geq f(x) \wedge f(y). \end{aligned}$$

Which shows that f is a fuzzy bi-ideal of $\mathcal{S}$. ∎

3.11. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy $(1, 2)$-ideal of $\mathcal{S}$.
(ii) f is a fuzzy interior ideal of $\mathcal{S}$.

Proof. $(i) \Longrightarrow (ii)$: Let $\mathcal{S}$ be a right regular LA-semigroup with left identity and let $x, a, y \in \mathcal{S}$. Then for a there exists $u \in \mathcal{S}$ such that $a = a^2u$. Let f be a fuzzy $(1,2)$-ideal of $\mathcal{S}$. Then by using $(4), (1)$, (2) and (3), we have

$$\begin{aligned}(xa)y &= (x(a^2u))y = (a^2(xu))y = (y(xu)(aa) = (y(xu)((a^2u)(a^2u))\\ &= (y(xu)((a^2a^2)(uu)) = (y(xu)((uu)(a^2a^2)) = (y(xu)(a^2(u^2a^2))\\ &= a^2((y(xu)(u^2a^2)) = ((u^2a^2)(y(xu))a^2 = ((a^2u^2)(y(xu))a^2\\ &= (((y(xu)u^2)a^2)a^2 = (((y(xu)u^2)(aa))(aa) = ((a((y(xu)u^2)a))(aa)\\ &= (av)(aa), \text{ where } v = (y(xu)u^2)a.\end{aligned}$$

Therefore $f((xa)y) = f((av)(aa)) \geq f(a) \wedge f(a) \wedge f(a)$. Which shows that f is a fuzzy interior ideal of $\mathcal{S}$.

$(ii) \Longrightarrow (i)$: Again let $\mathcal{S}$ be a right regular LA-semigroup with left identity and let $x, a, y, z \in \mathcal{S}$, then there exist x' and $z' \in \mathcal{S}$ such that $x = x^2x'$ and $z = z^2z'$. Now by using (3), we have

$$f((xa)(yz)) = f((zy)(ax)) \geq f(y).$$

Now by using (1) and (3), we have

$$\begin{aligned}f((xa)(yz)) &= f((((xx)x')a)(yz)) = f(((ax')(xx))(yz))\\ &= f(((xx)(x'a))(yz)) = f((((x'a)x)x)(yz)) \geq f(x).\end{aligned}$$

Now by using (4), we have

$$\begin{aligned}f((xa)(yz)) &= f((xa)(y(((zz)z')))) = f((xa)((zz)(yz')))\\ &= f((zz)((xa)(yz'))) \geq f(z).\end{aligned}$$

Thus we get that $f((xa)(yz)) \geq f(x) \wedge f(y) \wedge f(z)$.

Let a,$b \in \mathcal{S}$ then there exist a',$b' \in \mathcal{S}$ such that $a = a^2a'$ and $b = b^2b'$. Now by using (1), (3) and (4), we have

$$f(ab) = f(((aa)a')b) = f((ba')(aa)) = f((aa)(a'b)) \geq f(a)$$

and

$$f(ab) = f(a((bb)b')) = f((bb)(ab')) \geq f(b).$$

Thus f is a fuzzy $(1,2)$-ideal of $\mathcal{S}$. ∎

3.12. Corollary. *Fuzzy two-sided ideals, fuzzy bi-ideals, fuzzy generalized bi-ideals, fuzzy $(1,2)$-ideals, fuzzy interior ideals and fuzzy quasi-ideals coincide in a right regular LA-semigroup with left identity.*

3.13. Lemma. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy quasi ideal of $\mathcal{S}$.
(ii) $(f \circ C_{\mathcal{S}}(x)) \cap (C_{\mathcal{S}}(x) \circ f) = f$.

Proof. $(i) \Longrightarrow (ii)$ is followed by Lemma 3.5 and Theorem 3.9.
$(ii) \Longrightarrow (i)$ is obvious. ∎

3.14. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy bi-(generalized bi-) ideal of $\mathcal{S}$.
(ii) $(f \circ C_{\mathcal{S}}(x)) \circ f = f$ and $f \circ f = f$.

Proof. $(i) \Longrightarrow (ii)$: Assume that f is a fuzzy bi-ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a \in \mathcal{S}$, then there exists $x \in \mathcal{S}$ such that $a = a^2x$. Now by using (1), (4) and (3), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (x((aa)x))a = ((aa)(xx))a \\ &= ((xx)(aa))a = (((aa)x)x)a = (((xa)a)x)a \\ &= (((x((aa)x))a)x)a = ((((aa)(xx))a)x)a \\ &= ((((xx)(aa))a)x)a = (((a(x^2a))a)x)a. \end{aligned}$$

Therefore

$$\begin{aligned} ((f \circ C_{\mathcal{S}}(x)) \circ f)(a) &= \bigvee_{a=(((a(x^2a))a)x)a} \{(f \circ C_{\mathcal{S}}(x))(((a(x^2a))a)x) \wedge f(a)\} \\ &\geq \bigvee_{((a(x^2a))a)x=((a(x^2a))a)x} \{f((a(x^2a))a) \wedge C_{\mathcal{S}}(x)(x)\} \wedge f(a) \\ &\geq f((a(x^2a))a) \wedge 1 \wedge f(a) \\ &\geq f(a) \wedge f(a) \wedge f(a) = f(a). \end{aligned}$$

Now by using (1), (4) and (3), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (x((aa)x))a = ((aa)(xx))a \\ &= ((xx)(aa))a = (a(x^2a))a. \end{aligned}$$

Therefore

$$\begin{aligned} ((f \circ C_{\mathcal{S}}(x)) \circ f)(a) &= \bigvee_{a=(a(x^2a))a} \{(f \circ C_{\mathcal{S}}(x))((a(x^2a))) \wedge f(a)\} \\ &= \bigvee_{a=(a(x^2a))a} \left(\bigvee_{a(x^2a)=a(x^2a)} \{f(a) \wedge C_{\mathcal{S}}(x)(x^2a)\} \right) \wedge f(a) \\ &= \bigvee_{a=(a(x^2a))a} \left(\bigvee_{a(x^2a)=a(x^2a)} \{f(a) \wedge 1\} \right) \wedge f(a) \\ &= \bigvee_{a=(a(x^2a))a} \left(\bigvee_{a(x^2a)=a(x^2a)} f(a) \right) \wedge f(a) \\ &= \bigvee_{a=(a(x^2a))a} \{f(a) \wedge f(a)\} = f(a). \end{aligned}$$

Thus $(f \circ C_{\mathcal{S}}(x)) \circ f = f$.
Again by using (1), (4) and (3), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (x((aa)x))a = ((aa)(xx))a \\ &= (((xx)a)a)a = (((xx)((aa)x))a)a \\ &= (((xx)((xa)a))a)a = (((xx)((ae)(ax)))a)a \\ &= (((xx)(a((ae)x)))a)a = ((a((xx)((ae)x)))a)a \end{aligned}$$

Therefore

$$\begin{aligned} (f \circ f)(a) &= \bigvee_{a=((a((xx)((ae)x)))a)a} \{f((a((xx)((ae)x)))a) \wedge f(a)\} \\ &\geq f((a((xx)((ae)x)))a) \wedge f(a) \\ &\geq f(a) \wedge f(a) \wedge f(a) = f(a), \end{aligned}$$

therefore by using Lemma 2.4, $f \circ f = f$.
$(ii) \Longrightarrow (i)$: Let f be a fuzzy subset of a right regular LA-semigroup $\mathcal{S}$, then

$$\begin{aligned} f((xy)z) &= ((f \circ C_{\mathcal{S}}(x)) \circ f)((xy)z) \\ &= \bigvee_{(xy)z=(xy)z} \{(f \circ C_{\mathcal{S}}(x))(xy) \wedge f(z)\} \\ &\geq \bigvee_{xy=xy} \{f(x) \wedge C_{\mathcal{S}}(x)(y)\} \wedge f(z) \\ &\geq f(x) \wedge 1 \wedge f(z) = f(x) \wedge f(z). \end{aligned}$$

Since $f \circ f = f$, therefore by Lemma 2.4, f is a fuzzy LA-subsemigroup of $\mathcal{S}$. This shows that f is a fuzzy bi ideal of $\mathcal{S}$. ∎

3.15. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy interior ideal of $\mathcal{S}$.
(ii) $(C_{\mathcal{S}}(x) \circ f) \circ C_{\mathcal{S}}(x) = f$.

Proof. It is simple. ∎

3.16. Theorem. *In a right regular LA-semigroup $\mathcal{S}$ with left identity, the following statements are equivalent.*

(i) f is a fuzzy $(1,2)$-ideal of $\mathcal{S}$.
(ii) $(f \circ C_{\mathcal{S}}(x)) \circ (f \circ f) = f$ and $f \circ f = f$.

Proof. $(i) \Longrightarrow (ii)$: Let f be a fuzzy $(1,2)$-ideal of a right regular LA-semigroup $\mathcal{S}$ with left identity and let $a \in \mathcal{S}$, then there exists $x \in \mathcal{S}$ such that $a = a^2x$. Now by using (1) and (4), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (xa)((aa)x) = (aa)((xa)x) \\ &= (a((aa)x))((xa)x) = ((aa)(ax))((xa)x) \\ &= (((xa)x)(ax))(aa) = (a(((xa)x)x))(aa). \end{aligned}$$

Therefore

$$((f \circ C_{\mathcal{S}}(x)) \circ (f \circ f))(a) = \bigvee_{a=(a(((xa)x)x))(aa)} \{(f \circ C_{\mathcal{S}}(x))(a(((xa)x)x)) \wedge (f \circ f)(aa)\}.$$

Now

$$\begin{aligned} (f \circ C_{\mathcal{S}}(x))(a(((xa)x)x)) &= \bigvee_{a(((xa)x)x)=a(((xa)x)x)} \{f(a) \wedge C_{\mathcal{S}}(x)(((xa)x)x)\} \\ &\geq f(a) \wedge C_{\mathcal{S}}(x)(((xa)x)x) = f(a) \end{aligned}$$

and

$$(f \circ f)(aa) = \bigvee_{aa=aa} \{f(a) \wedge f(a)\} \geq f(a).$$

Thus we get

$$((f \circ C_{\mathcal{S}}(x)) \circ (f \circ f))(a) \geq f(a).$$

Now by using (4), (1) and (3), we have

$$\begin{aligned} a &= (aa)x = (((aa)x)((aa)x))x = ((aa)(((aa)x)x))x \\ &= ((aa)((xx)(aa)))x = ((aa)(x^2(aa)))x \\ &= (x(x^2(aa)))(aa) = (x(a(x^2a)))(aa) \\ &= (a(x(x^2a)))(aa) = (a(x(x^2((aa)x))))(aa) \\ &= (a(x((aa)x^3)))(aa). \end{aligned}$$

Therefore

$$((f \circ C_{\mathbb{S}}(x)) \circ (f \circ f))(a) = \bigvee_{a=(a(x((aa)x^3)))(aa)} \{(f \circ C_{\mathbb{S}}(x))(a(x((aa)x^3))) \wedge (f \circ f)(aa)\}.$$

Now

$$\begin{aligned} (f \circ C_{\mathbb{S}}(x))(a(x((aa)x^3))) &= \bigvee_{a(x((aa)x^3))=a(x((aa)x^3))} \{f(a) \wedge C_{\mathbb{S}}(x)(x((aa)x^3))\} \\ &= \bigvee_{a(x((aa)x^3))=a(x((aa)x^3))} f(a) \end{aligned}$$

and

$$(f \circ f)(aa) = \bigvee_{aa=aa} \{f(a) \wedge f(a)\} = \bigvee_{aa=aa} f(a).$$

Therefore

$$\begin{aligned} (f \circ C_{\mathbb{S}}(x))(a(x((aa)x^3))) \wedge (f \circ f)(aa) &= \bigvee_{a(x((aa)x^3))=a(x((aa)x^3))} f(a) \wedge \bigvee_{aa=aa} f(a) \\ &= \bigvee_{a(x((aa)x^3))=a(x((aa)x^3))} \{f(a) \wedge f(a)\}. \end{aligned}$$

Thus from above, we get

$$\begin{aligned} ((f \circ C_{\mathbb{S}}(x)) \circ (f \circ f))(a) &= \bigvee_{a=(a(x((aa)x^3)))(aa)} \left(\bigvee_{a(x((aa)x^3))=a(x((aa)x^3))} \{f(a) \wedge f(a)\} \right) \\ &= \bigvee_{a=(a(x((aa)x^3)))(aa)} \{f(a) \wedge f(a) \wedge f(a)\} \\ &\leq \bigvee_{a=(a(x((aa)x^3)))(aa)} f((a(x((aa)x^3)))(aa)) = f(a). \end{aligned}$$

Therefore $(f \circ C_{\mathbb{S}}(x)) \circ (f \circ f) = f$.
Now by using (1) and (4), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (x((aa)x))a = ((aa)(xx))a = ((a((aa)x))x^2)a \\ &= (((aa)(ax))x^2)a = ((x^2(ax))(aa))a = ((ax^3)(aa))a. \end{aligned}$$

Thus

$$\begin{aligned} (f \circ f)(a) &= \bigvee_{a=((ax^3)(aa))a} \{f(((ax^3)(aa))) \wedge f(a)\} \\ &\geq f(a) \wedge f(a) \wedge f(a) = f(a). \end{aligned}$$

Now by using Lemma 2.4, $f \circ f = f$.

$(ii) \Longrightarrow (i)$: Let f be a fuzzy subset of a right regular LA-semigroup $\mathcal{S}$. Now since $f \circ f = f$, therefore by Lemma 2.4, f is a fuzzy LA-subsemigroup of $\mathcal{S}$

$$\begin{aligned}
f((xa)(yz)) &= ((f \circ C_{\mathcal{S}}(x)) \circ (f \circ f))((xa)(yz)) \\
&= ((f \circ C_{\mathcal{S}}(x)) \circ f)((xa)(yz)) \\
&= \bigvee_{(xa)(yz)=(xa)(yz)} \{(f \circ C_{\mathcal{S}}(x))(xa) \wedge f(yz)\}. \\
&\geq (f \circ C_{\mathcal{S}}(x))(xa) \wedge f(yz) \\
&= \bigvee_{(xa)=(xa)} \{f(x) \wedge C_{\mathcal{S}}(x)(a)\} \wedge f(yz) \\
&\geq f(x) \wedge 1 \wedge f(y) \wedge f(z) \\
&= f(x) \wedge f(y) \wedge f(z).
\end{aligned}$$

This shows that $f((xa)(yz)) \geq f(x) \wedge f(y) \wedge f(z)$, therefore f is a fuzzy $(1,2)$-ideal of $\mathcal{S}$. ∎

3.17. Theorem. *Let $\mathcal{S}$ be an LA-semigroup with left identity, then the following conditions are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) Every fuzzy left ideal of $\mathcal{S}$ is idempotent.

Proof. $(i) \Longrightarrow (ii)$: Let $\mathcal{S}$ be an LA-semigroup with left identity. Let $a \in \mathcal{S}$, then since $\mathcal{S}$ is right regular so by using (1),

$$a = (aa)x = (xa)a.$$

Let f be a fuzzy left ideal of $\mathcal{S}$, then clearly $f \circ f \subseteq f$ and therefore

$$(f \circ f)(a) = \bigvee_{a=(xa)a} \{f((xa)a) \wedge f(a)\} \geq f(a) \wedge f(a) = f(a).$$

Thus f is idempotent.

$(ii) \Longrightarrow (i)$: Assume that every fuzzy left ideal of an LA-semigroup $\mathcal{S}$ with left identity is idempotent. Since $\mathcal{S}a$ is a left ideal of $\mathcal{S}$, therefore by Lemma 2.6, its characteristic function $C_{\mathcal{S}a}$ is a fuzzy left ideal of $\mathcal{S}$. Since $a \in \mathcal{S}a$, therefore $C_{\mathcal{S}a}(a) = 1$. Now by using the given assumption and Lemma 2.5, we have

$$C_{\mathcal{S}a} \circ C_{\mathcal{S}a} = C_{\mathcal{S}a} \text{ and } C_{\mathcal{S}a} \circ C_{\mathcal{S}a} = C_{(\mathcal{S}a)^2}.$$

Thus we have $(C_{(\mathcal{S}a)^2})(a) = (C_{\mathcal{S}a})(a) = 1$, which implies that $a \in (\mathcal{S}a)^2$. Now by using (3) and (2), we have

$$a \in (\mathcal{S}a)^2 = (\mathcal{S}a)(\mathcal{S}a) = (a\mathcal{S})(a\mathcal{S}) = a^2\mathcal{S}.$$

This shows that $\mathcal{S}$ is right regular. ∎

Note that if an LA-semigroup has a left identity then $C_{\mathcal{S}}(x) \circ C_{\mathcal{S}}(x) = C_{\mathcal{S}}(x)$.

3.18. Theorem. *For an LA-semigroup $\mathcal{S}$ with left identity, the following conditions are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) $f = (C_{\mathcal{S}}(x) \circ f)^2$, where f is any fuzzy left ideal of $\mathcal{S}$.

Proof. $(i) \Longrightarrow (ii)$: Assume that $\mathcal{S}$ is a right regular LA-semigroup and let f be any fuzzy left ideal of $\mathcal{S}$, then clearly $C_{\mathcal{S}}(x) \circ f$ is a fuzzy left ideal of $\mathcal{S}$. Now by using Theorem 3.17, $C_{\mathcal{S}}(x) \circ f$ is idempotent and, therefore, we have

$$(C_{\mathcal{S}}(x) \circ f)^2 = C_{\mathcal{S}}(x) \circ f \subseteq f.$$

Now let $a \in \mathcal{S}$, since $\mathcal{S}$ is right regular, therefore there exists $x \in \mathcal{S}$ such that $a = a^2x$ and by using (1), we have

$$a = (aa)x = (xa)a = (xa)((aa)x) = (xa)((xa)a)$$

Therefore

$$\begin{aligned}(C_{\mathcal{S}}(x) \circ f)^2(a) &= \bigvee_{a=(xa)((xa)a)} \{(C_{\mathcal{S}}(x) \circ f)(xa) \wedge (C_{\mathcal{S}}(x) \circ f)((xa)a)\} \\ &\geq (C_{\mathcal{S}}(x) \circ f)(xa) \wedge (C_{\mathcal{S}}(x) \circ f)((xa)a) \\ &= \bigvee_{xa=xa} \{C_{\mathcal{S}}(x)(x) \wedge f(a)\} \wedge \bigvee_{(xa)a=(xa)a} \{C_{\mathcal{S}}(x)(xa) \wedge f(a)\} \\ &\geq C_{\mathcal{S}}(x)(x) \wedge f(a) \wedge C_{\mathcal{S}}(x)(xa) \wedge f(a) = f(a).\end{aligned}$$

Thus we obtain $f = (C_{\mathcal{S}}(x) \circ f)^2$.

$(ii) \Longrightarrow (i)$: Let $f = (C_{\mathcal{S}}(x) \circ f)^2$ holds for any fuzzy left ideal f of $\mathcal{S}$, then by given assumption, we have

$$f = (C_{\mathcal{S}}(x) \circ f)^2 \subseteq f^2 = f \circ f \subseteq C_{\mathcal{S}}(x) \circ f \subseteq f.$$

Thus by using Theorem 3.17, $\mathcal{S}$ is right regular. ∎

An LA-semigroup $\mathcal{S}$ is called a left (right) duo if every left (right) ideal of $\mathcal{S}$ is a two-sided ideal of $\mathcal{S}$ and is called a duo if it is both a left and a right duo.

Consider an LA-semigroup $\mathcal{S}$ in Example 3.1, the right ideals of $\mathcal{S}$ are $\{a, b, c, e\}$ and $\{a, e\}$ which are also two-sided ideals of $\mathcal{S}$. Thus $\mathcal{S}$ is a right duo. On the other hand, the left ideals of $\mathcal{S}$ are $\{a, b, e\}$, $\{a, c, e\}$, $\{a, b, c, e\}$ and $\{a, e\}$. Note that $\mathcal{S}$ is not a left duo because $\{a, b, e\}$ and $\{a, c, e\}$ are not the right ideals of $\mathcal{S}$.

An LA-semigroups considered in Examples 2.1 and 3.2 are duo because in both examples, the only right (left, two-sided) ideals of $\mathcal{S}$ are $\{a, b\}$.

An LA-semigroup $\mathcal{S}$ is called a fuzzy left (right) duo if every fuzzy left (right) ideal of $\mathcal{S}$ is a fuzzy two-sided ideal of $\mathcal{S}$ and is called a fuzzy duo if it is both a fuzzy left and a fuzzy right duo.

By Lemma 3.7, every right regular LA-semigroup $\mathcal{S}$ with left identity is a fuzzy left (right) duo.

3.19. Theorem. *A right regular LA-semigroup $\mathcal{S}$ with left identity is a left (right) duo if and only if it is a fuzzy left (right) duo.*

Proof. Let a right regular LA-semigroup $\mathcal{S}$ be a left duo and assume that f is any fuzzy left ideal of $\mathcal{S}$. Let $a, b \in \mathcal{S}$, then $a \in (aa)\mathcal{S}$. Now $\mathcal{S}a$ is a left ideal of $\mathcal{S}$, therefore by hypothesis, $\mathcal{S}a$ is a two sided ideal of $\mathcal{S}$. Therefore by using (1), we have

$$ab \in ((aa)\mathcal{S})b = ((\mathcal{S}a)a)b \subseteq ((\mathcal{S}a)\mathcal{S})\mathcal{S} \subseteq (\mathcal{S}a).$$

Thus $ab = ca$ for some $c \in \mathcal{S}$. Now $f(ab) = f(ca) \geq f(a)$, implies that f is a fuzzy right ideal of $\mathcal{S}$ and therefore $\mathcal{S}$ is a fuzzy left duo.

Conversely, assume that $\mathcal{S}$ is a fuzzy left duo and let L be any left ideal of $\mathcal{S}$. Now by Lemma 2.6, the characteristic function C_L of L is a fuzzy left ideal of $\mathcal{S}$. Thus by hypothesis C_L is a fuzzy two-sided ideal of $\mathcal{S}$ and by using Lemma 2.6, L is a two sided ideal of $\mathcal{S}$. Thus $\mathcal{S}$ is a left duo.

Now again let $\mathcal{S}$ be a right regular LA-semigroup such that $\mathcal{S}$ is a right duo and assume that f is any fuzzy right ideal of $\mathcal{S}$. Let $a, b \in \mathcal{S}$, then there exists $x \in \mathcal{S}$ such that $b = b^2x$. Now clearly $b^2 \in b^2\mathcal{S}$ and since $b^2\mathcal{S}$ is a right ideal of $\mathcal{S}$, therefore

$$b = b^2x \in (b^2\mathcal{S})\mathcal{S} \subseteq b^2\mathcal{S}.$$

As $b^2\mathcal{S}$ is a right ideal of $\mathcal{S}$, therefore by hypothesis $b^2\mathcal{S}$ is a two sided ideal of $\mathcal{S}$. Now by using (1), we have

$$ab \in a(b^2\mathcal{S}) \subseteq \mathcal{S}(b^2\mathcal{S}) \subseteq b^2\mathcal{S}.$$

Thus $ab = (bb)c$ for some $c \in \mathcal{S}$. Now $f(ab) = f((bb)c) \geq f(b)$, implies that f is a fuzzy left ideal of $\mathcal{S}$ and therefore $\mathcal{S}$ is a fuzzy right duo.

The Converse is simple. ∎

3.20. Theorem. *Let $\mathcal{S}$ be a right regular LA-semigroup with left identity, then the following statements are equivalent.*

(i) f is a fuzzy left ideal of $\mathcal{S}$.
(ii) f is a fuzzy right ideal of $\mathcal{S}$.
(iii) f is a fuzzy two-sided ideal of $\mathcal{S}$.
(iv) f is a fuzzy bi-ideal of $\mathcal{S}$.
(v) f is a fuzzy generalized bi-ideal of $\mathcal{S}$.
(vi) f is a fuzzy $(1,2)$-ideal of $\mathcal{S}$.
(vii) f is a fuzzy interior ideal of $\mathcal{S}$.
$(viii)$ f is a fuzzy quasi ideal of $\mathcal{S}$.
(ix) $f \circ C_{\mathcal{S}}(x) = f$ and $C_{\mathcal{S}}(x) \circ f = f$.

Proof. $(i) \Longrightarrow (ix)$: Let f be a fuzzy left ideal of a right regular LA-semigroup $\mathcal{S}$. Let $a \in \mathcal{S}$, then there exists $a' \in \mathcal{S}$ such that $a = a^2a'$. Now by using (1) and (3), we have

$$a = (aa)a' = (a'a)a \text{ and } a = (aa)a' = (aa)(ea') = (a'e)(aa).$$

Therefore

$$(f \circ C_{\mathcal{S}}(x))(a) = \bigvee_{a=(a'a)a} \left\{f(a'a) \wedge C_{\mathcal{S}}(x)(a)\right\} \geq f(a'a) \wedge 1 \geq f(a)$$

and

$$(C_{\mathcal{S}}(x) \circ f)(a) = \bigvee_{a=(a'e)(aa)} \left\{C_{\mathcal{S}}(x)(a'e) \wedge f(aa)\right\} \geq 1 \wedge f(aa) \geq f(a).$$

Now by using Lemmas 3.7 and 2.4, we get that $f \circ C_{\mathcal{S}}(x) = f$ and $C_{\mathcal{S}}(x) \circ f = f$.

$(ix) \Longrightarrow (viii)$ is obvious.

$(viii) \Longrightarrow (vii)$: Let f be a fuzzy quasi ideal of a right regular LA-semigroup $\mathcal{S}$. Now for $a \in \mathcal{S}$ there exists $a' \in \mathcal{S}$ such that $a = a^2a'$ and therefore by using (3) and (4), we have

$$(xa)y = (xa)(ey) = (ye)(ax) = a((ye)x)$$

also

$$\begin{aligned} (xa)y &= (x((aa)a'))y = ((aa)(xa'))y = ((a'x)(aa))y \\ &= (a((a'x)a))y = (y((a'x)a))a. \end{aligned}$$

Since f is a fuzzy quasi ideal of $\mathcal{S}$, therefore by Lemma 3.13, we have

$$f((xa)y) = ((f \circ C_{\mathcal{S}}(x)) \cap (C_{\mathcal{S}}(x) \circ f))((xa)y) = (f \circ C_{\mathcal{S}}(x))\,((xa)y) \wedge (C_{\mathcal{S}}(x) \circ f)\,((xa)y).$$

Now

$$(f \circ C_{\mathcal{S}}(x))\,((xa)y) = \bigvee_{(xa)y=a((ye)x)} \{f(a) \wedge C_{\mathcal{S}}(x)((ye)x)\} \geq f(a)$$

and

$$\left(C_{\mathcal{S}}(x) \circ f\right)((xa)y) = \bigvee_{(xa)y=(y((a^{'}x)a))a} \left\{ C_{\mathcal{S}}(x)(y((a^{'}x)a)) \wedge f(a) \right\} \geq f(a).$$

Which implies that $f((xa)y) \geq f(a)$. Thus f is a fuzzy interior ideal of $\mathcal{S}$.
$(vii) \Longrightarrow (vi)$ is followed by Theorem 3.11.
$(vi) \Longrightarrow (v)$ is followed by Theorem 3.10.
$(v) \Longrightarrow (iv)$: It is simple.
$(iv) \Longrightarrow (iii)$ is followed by Theorems 3.10 and 3.8.
$(iii) \Longrightarrow (ii)$ and $(ii) \Longrightarrow (i)$ are easy consequences of Lemma 3.7. ∎

3.21. Theorem. *For an LA-semigroup $\mathcal{S}$ with left identity, the following conditions are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) $f \cap g \subseteq f \circ g$, for every fuzzy right ideal f and g of $\mathcal{S}$, where f and g are fuzzy semiprime.
(iii) $f \cap g \subseteq f \circ g$, for every fuzzy left ideal f and g of $\mathcal{S}$, where f and g are fuzzy semiprime.

Proof. $(i) \Rightarrow (iii)$: Assume that f and g are fuzzy left ideals of $\mathcal{S}$ with left identity. Let a be any element in $\mathcal{S}$, since $\mathcal{S}$ is right regular, so exists x in $\mathcal{S}$, such that $a = a^2x$. Now by using (1), (4), (3) and (2), we have

$$\begin{aligned} a &= (aa)x = (xa)a = (xa)(a^2x) = a^2((xa)x) \\ &= (aa)((xa)x) = (x(xa))(aa) = (xa)((xa)a). \end{aligned}$$

Therefore

$$\begin{aligned} (f \circ g)(a) &= \bigvee_{a=(xa)((xa)a)} \{f(xa) \wedge g((xa)a)\} \geq f(xa) \wedge g((xa)a) \\ &\geq f(a) \wedge g(a) = (f \cap g)(a). \end{aligned}$$

$(iii) \Rightarrow (ii)$ can be followed from Lemma 2.3.
$(ii) \Rightarrow (i)$: Assume that f and g are any fuzzy left ideals of $\mathcal{S}$ with left identity and let R and $R^{'}$ be any right ideals of $\mathcal{S}$, then by Lemma 2.6, C_R and $C_{R^{'}}$ are fuzzy right ideals of $\mathcal{S}$. Let $a \in R \cap R^{'}$,therefore by Lemma 2.5 and given assumption, we have

$$1 = C_{R\cap R^{'}}(a) = (C_R \cap C_{R^{'}})(a) \subseteq (C_R \circ C_{R^{'}})(a) = (C_{RR^{'}})(a),$$

which implies that $R \cap R^{'} \subseteq RR^{'}$. Since f and g are fuzzy semiprime, so R and $R^{'}$ are fuzzy semiprime. As $a^2\mathcal{S}$ is a right ideal of $\mathcal{S}$ and clearly $a^2 \in a^2\mathcal{S}$, therefore $a \in a^2\mathcal{S}$. Now by using (4), we have

$$a \in a^2\mathcal{S} \cap a^2\mathcal{S} \subseteq (a^2\mathcal{S})(a^2\mathcal{S}) = a^2((a^2\mathcal{S})\mathcal{S}) \subseteq a^2\mathcal{S}.$$

This shows that $\mathcal{S}$ is right regular. ∎

A subset A of an LA-semigroup $\mathcal{S}$ is called semiprime if $a^2 \in A$ implies $a \in A$.

The subset $\{a, b\}$ of an LA-semigroup $\mathcal{S}$ in Example 2.1 is semiprime.

A fuzzy subset f of an LA-semigroup $\mathcal{S}$ is called a fuzzy semiprime if $f(a) \geq f(a^2)$ for all a in $\mathcal{S}$.

Let us define a fuzzy subset f of an LA-semigroup $\mathcal{S}$ in Example 3.1 as follows: $f(a) = 0.2, f(b) = 0.5, f(c) = 0.6, f(d) = 0.1$ and $f(e) = 0.4$, then f is a fuzzy semiprime.

3.22. Lemma. *For a right regular LA-semigroup $\mathcal{S}$, the following holds.*

(i) Every fuzzy right ideal of $\mathcal{S}$ is a fuzzy semiprime.
(ii) Every fuzzy left ideal of $\mathcal{S}$ is a fuzzy semiprime if $\mathcal{S}$ has a left identity.

Proof. (i) : It is simple.
(ii) : Let f be a fuzzy left ideal of a right regular LA-semigroup $\mathcal{S}$ and let $a \in \mathcal{S}$, then there exists $x \in \mathcal{S}$ such that $a = a^2x$. Now by using (3), we have

$$f(a) = f((aa)(ex)) = f((xe)a^2) \geq f(a^2).$$

This shows that f is a fuzzy semiprime. ∎

Right, left and two-sided ideals of an LA-semigroup $\mathcal{S}$ are semiprime if and only if their characteristic functions are fuzzy semiprime.

3.23. Lemma. *Let $\mathcal{S}$ be an LA-semigroup with left identity, then the following statements are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) Every fuzzy right (left, two-sided) ideal of $\mathcal{S}$ is fuzzy semiprime.

Proof. (i) $\Longrightarrow$ (ii) : It follows from Lemma 3.22.
(ii) $\Longrightarrow$ (i) : Assume that $\mathcal{S}$ is an LA-semigroup with left identity and let every fuzzy right (left, two-sided) ideal of $\mathcal{S}$ be fuzzy semiprime. Since $a^2\mathcal{S}$ is a right and also a left ideal of $\mathcal{S}$, therefore, $a^2\mathcal{S}$ is semiprime. Now clearly $a^2 \in a^2\mathcal{S}$, therefore $a \in a^2\mathcal{S}$, which shows that $\mathcal{S}$ is right regular. ∎

3.24. Theorem. *The following statements are equivalent for an LA-semigroup $\mathcal{S}$ with left identity.*

(i) $\mathcal{S}$ is right regular.
(ii) Every fuzzy right ideal of $\mathcal{S}$ is fuzzy semiprime.
(iii) Every fuzzy left ideal of $\mathcal{S}$ is fuzzy semiprime.

Proof. (i) $\Longrightarrow$ (iii) and (ii) $\Longrightarrow$ (i) are followed by Lemma 3.23.
(iii) $\Longrightarrow$ (ii) : Assume that $\mathcal{S}$ is an LA-semigroup and let f be a fuzzy right ideal of $\mathcal{S}$, then by using Lemma 2.3, f is a fuzzy left ideal of $\mathcal{S}$ and therefore by given assumption f is a fuzzy semiprime. ∎

3.25. Theorem. *For an LA-semigroup $\mathcal{S}$ with left identity, the following conditions are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) Every fuzzy two-sided ideal of $\mathcal{S}$ is fuzzy semiprime.
(iii) Every fuzzy bi-ideal of $\mathcal{S}$ is fuzzy semiprime.
(iv) $f(a) = f(a^2)$, for all fuzzy two sided ideal f of $\mathcal{S}$, for all $a \in \mathcal{S}$.
(v) $f(a) = f(a^2)$, for all fuzzy bi-ideal f of $\mathcal{S}$, for all $a \in \mathcal{S}$.

Proof. (i) $\Rightarrow$ (v) : Assume that f is any fuzzy bi-ideal of $\mathcal{S}$. Let a be any element of $\mathcal{S}$. Since $\mathcal{S}$ is right regular, so there exists x in $\mathcal{S}$, such that $a = a^2x$. Now by using (1), (4) and (3), we have

$$\begin{aligned} f(a) &= f((aa)x) = f((xa)a) = f((x(a^2x)a) = f((a^2x^2)a) \\ &= f((ax^2)a^2) = f(((a^2x)x^2)a^2) = f(((x^2x)a^2)a^2) \\ &= f(((x^2x)(aa))a^2) = f(((aa)(x^2x))a^2) \\ &= f((a^2x^3)a^2) \geq f(a^2) \wedge f(a^2) = f(aa) \\ &= f((a^2x)a) = f(((aa)(ex))a) = f(((xe)(aa))a) \\ &= f((a((xe)a))a) \geq f(a) \wedge f(a) = f(a). \end{aligned}$$

This shows that $f(a) = f(a^2)$ for all a in $\mathcal{S}$. Clearly $(v) \Rightarrow (iv)$.

$(iv) \Rightarrow (i)$: Since $a^2\mathcal{S}$ is a two sided ideal of $\mathcal{S}$ with left identity, therefore it is clear to see that $a^2 \in a^2\mathcal{S}$. Now by Lemma 2.6, $C_{a^2\mathcal{S}}$ is a fuzzy two sided ideal of $\mathcal{S}$ and by given assumption , we have $C_{a^2\mathcal{S}}(a) = C_{a^2\mathcal{S}}\left(a^2\right) = 1$. Therefore $a \in a^2\mathcal{S}$, which shows that $\mathcal{S}$ is a right regular.

It is easy to observe that $(ii) \iff (iv)$ and $(iii) \iff (v)$. ∎

3.26. Theorem. *For an LA-semigroup $\mathcal{S}$ with left identity, the following conditions are equivalent.*

(i) $\mathcal{S}$ is right regular.
(ii) Every right ideal of $\mathcal{S}$ is semiprime.
(iii) Every fuzzy right ideal of $\mathcal{S}$ is fuzzy semiprime.
(iv) $f(a) = f(a^2)$, for every fuzzy right ideal f of $\mathcal{S}$ and for all a in $\mathcal{S}$.
(v) $f(a) = f(a^2)$, for every fuzzy left ideal f of $\mathcal{S}$ and for all a in $\mathcal{S}$.

Proof. $(i) \Rightarrow (v)$: Assume that f is a fuzzy left ideal of $\mathcal{S}$. Let a be any element in $\mathcal{S}$, since $\mathcal{S}$ is right regular, so exists x in $\mathcal{S}$, such that $a = a^2x$. Now by using (3), we have

$$f(a^2x) = f((aa)(ex)) = f((xe)(aa)) \geq f(aa) \geq f(a),$$

therefore $f(a) = f(a^2)$ for all a in $\mathcal{S}$.

From Lemma 2.3, it is clear that $(v) \Rightarrow (iv)$ and $(iv) \Rightarrow (iii)$ are obvious.

Now $(iii) \Rightarrow (ii)$ and $(ii) \Rightarrow (i)$ are easy. ∎

References

[1] P. Holgate, Groupoids satisfying a simple invertive law, The Math. Stud., $1-4$ (61) (1992), $101-106$.

[2] M. A. Kazim and M. Naseeruddin, On almost semigroups, The Alig. Bull. Math., 2 (1972), $1-7$.

[3] M. Khan and M. N. A Khan, On fuzzy Abel-Grassmann's groupoids, Advances in Fuzzy Mathematics, 3 (2010), $349-360$.

[4] Madad Khan and N. Ahmad, Characterizations of left almost semigroups by their ideals, Journal of Advanced Research in Pure Mathematics, 2 (2010), $61-73$.

[5] N. Kuroki, Fuzzy bi-ideals in semigroups, Comment. Math. Univ. St. Pauli, 27 (1979), $17-21$.

[6] J. N. Mordeson, D. S. Malik and N. Kuroki, *Fuzzy semigroups*. Springer-Verlag, Berlin, Germany (2003).

[7] Q. Mushtaq and S. M. Yusuf, On LA-semigroups, The Alig. Bull. Math., 8 (1978), $65-70$.

[8] Q. Mushtaq and S. M. Yousuf, On LA-semigroup defined by a commutative inverse semigroup, Math. Bech., 40(1988), $59-62$.

[9] M. Naseeruddin, Some studies in almost semigroups and flocks, Ph.D., thesis. Aligarh Muslim University, Aligarh, India, (1970).

[10] P. V. Protić and N. Stevanović, AG-test and some general properties of Abel-Grassmann's groupoids. PU. M. A., 4 (6) (1995), $371-383$.

[11] N. Stevanović and P. V. Protić, Composition of Abel-Grassmann's 3-bands, Novi Sad, J. Math., 2, 34 (2004), $175-182$.

[12] A. Rosenfeld, Fuzzy groups, J. Math. Anal. Appl., 35 (1971), $512-517$.

[13] L. A. Zadeh, *Fuzzy sets*. Inform. Control, 8 (1965), $338-353$.

[14] L. A. Zadeh, *Probability theory should be based on fuzzy logics-a contentious view*, Soft Methadology and Random Informations Systems, Adv. Soft Comput., vols. $3-7$, Springer, Berlin (2004).

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 587 – 595

Generalizations of prime submodules

S. Moradi * and A. Azizi †

Abstract

Let R be a commutative ring with identity and M a unitary R-module, and $n > 1$ an integer number. As a generalization of the concept of prime submodules, a proper submodule N of M will be called n-almost prime, if for $r \in R$ and $x \in M$ with $rx \in N \setminus (N : M)^{n-1}N$, either $x \in N$ or $r \in (N : M)$. We study n-almost prime submodules, in this paper.

2000 AMS Classification: 13C99, 13C13, 13E05, 13F05, 13F15.

Keywords: Essential multiplicatively closed subsets; n-almost prime submodules; Prime submodules.

Received 06/01/2011 : Accepted 15/01/2013 Doi : 10.15672/HJMS.20154411041

1. Introduction

Throughout this paper all rings are commutative with identity and all modules are unitary. Also we consider R to be a commutative ring with identity, M an R-module, $n > 1$ a positive integer and $\mathbb{N}$ the set of positive integers.

Let N be a submodule of an $R-$module M. The set $\{r \in R | rM \subseteq N\}$ is denoted by $(N : M)$ and particularly we denote $\{r \in R | rN = 0\}$ by $ann(N)$.

Let N a proper submodule of M. It is said that N is a *prime submodule* of M, if for $r \in R$ and $x \in M$ with $rx \in N$, either $x \in N$ or $rM \subseteq N$. In this case, if $P = (N : M)$, then P is a prime ideal. The concept of prime submodules has been studied in many papers in recent years (see, for example, [3, 8]).

*Department of Mathematics, College of Sciences, Shiraz University, Shiraz, 71457-44776, IRAN
Email: smoradi@shirazu.ac.ir
† Email:aazizi@shirazu.ac.ir

2. n-Almost Prime Submodules

According to [1] an ideal I of R is called an *n-almost prime* ideal if for $a, b \in R$ with $ab \in I \setminus I^n$, either $a \in I$ or $b \in I$. The case $n = 2$ is called an *almost prime* ideal and it is due to [5]. We will generalize this definition to modules as follows:

Definition. Let $n > 1$ be an integer number. A proper submodule N of M will be called n-almost prime, if for $r \in R$ and $x \in M$ with $rx \in N \setminus (N : M)^{n-1}N$, either $x \in N$ or $r \in (N : M)$. A 2-almost prime submodule will be called an almost prime submodule.

Evidently every prime submodule is an n-almost prime submodule, for any integer $n > 1$.

The following remark is an evident consequence of the definition of being almost prime submodules.

Remark.

(i) The zero submodule is an almost prime submodule.
(ii) Let N be a proper submodule of M such that $(N : M)^{n-1}N = N$. Then N is n-almost prime.
(iii) Let N be a proper submodule of a torsion-free divisible module M. Then N is prime if and only if N is n-almost prime.
(iv) Every n-almost prime submodule of an R-module M is $m-$almost prime, where $3 \leq n$ and $1 < m \leq n$.

2.1. Lemma. Let M be an R-module, and I an ideal of R.

(i) If $n \in \mathbb{N}$, then $(IM : M)^n M = I^n M$.
(ii) If K is a submodule of M such that $(K : M)$ is a maximal ideal, then K is a prime submodule.
(iii) If $1 < n \in \mathbb{N}$ such that $M \neq IM = I^n M$, then IM is an n-almost prime submodule.
(iv) Let F be a free R-module. Then I is an n-almost prime ideal of R if and only if IF is an n-almost prime submodule of F.
(v) Consider the R-module $F = \oplus_{i \in \mathbb{N}} R$ and let $N = I \oplus (\oplus_{1<i\in\mathbb{N}} R)$. Then the following are equivalent:
    (a) N is a prime submodule of F;
    (b) N is an n-almost prime submodule of F;
    (c) I is a prime ideal of R.

Proof. The proofs of (i),(ii) and (iii) are clear.

(iv) Consider $F = \oplus_{i\in\alpha} R$. It is easy to see that $(IF : F) = I$, for any ideal I of R. Then I is a proper ideal of R if and only if IF is a proper submodule of F. Also $(IF : F)^{n-1} IF = I^n F$.

Suppose I is a proper ideal of R, which is not n-almost prime. Then there exist $a, b \in R \setminus I$ such that $ab \in I \setminus I^n$. So $a(b, 0, 0, 0, \cdots) \in IF \setminus I^n F$, but $a \notin I = (IF : F)$, also $(b, 0, 0, 0, \cdots) \notin IF$, that is IF is not an n-almost prime submodule.

For the converse, suppose I is an n-almost prime ideal of R. We consider the following two cases:

Case 1. $F = R \oplus R$, that is $rank\ F = 2$.

Let $r(a, b) \in IF \setminus I^n F$, where $r \in R \setminus (IF : F) = I$ and $a, b \in R$. Then $ra, rb \in I$, and ra or rb is not in I^n. Without loss of generality, we may assume $ra \notin I^n$. Then $ra \in I \setminus I^n$ and as $r \notin I$, $a \in I$. Similarly if $rb \notin I^n$, then $b \in I$ and so $(a, b) \in IF$.

Now let $rb \in I^n$. Then $r(a + b) \in I$, and $ra \notin I^n$, and so $r(a + b) \in I \setminus I^n$, and $r \notin I$, hence $a + b \in I$. Also $a \in I$, therefore $b \in I$, that is $(a, b) \in IF$.

Case 2. F is a free module of arbitrary rank.

If $a \in F$, then $a \in \oplus_{i=1}^{n} Ra_i$, where $a_1, a_2, \cdots, a_n \in F$ for some integer n. Now by using case 1, we get the results.

(v) The proofs of (a) $\Longrightarrow$ (b) and (c) $\Longrightarrow$ (a) are straightforward.

(b) $\Longrightarrow$ (c) It is easy to see that I is an n-almost prime ideal of R. Now if I is not a prime ideal, then there exists $a, b \in R \setminus I$ such that $ab \in I$. Since I is an n-almost prime ideal, $ab \in I^n$. Therefore $a(b, 1, 1, 1, \cdots) \in N \setminus (N : M)^{n-1}N$, however $a \notin I = (N : M)$ and $(b, 1, 1, 1, \cdots) \notin N$, which is a contradiction. ■

Examples.

(1) If I is an ideal of R generated by idempotents, then by Lemma 2.1(iii), IM is an almost prime submodule, or $IM = M$, for any R-module M. For a specific example, let R' be an arbitrary ring, and consider $R = \prod_{n=1}^{\infty} R'$ and $I = \oplus_{n=1}^{\infty} R'$, particularly I is an almost prime ideal.

(2) Let $R = K[[X^3, X^4, X^5]]$, where K is a field, and $I = \langle X^3, X^4 \rangle$. By [1, Example 11], I is an almost prime ideal, which is not a 3-almost prime ideal.

Let F be a free R-module. By Lemma 2.1(iv), the submodule IF is an almost prime submodule, which is not a 3-almost prime submodule.

(3) Let R be an Artinian ring. Then for any ideal I of R, there exists an $n \in \mathbb{N}$ such that $I^n = I^{n+1}$. So the ideal $J = I^n$ is an almost prime ideal, and by Lemma 2.1(iv), for any free R-module F, the submodule JF is an almost prime submodule.

Let M, M' be two R-modules. For a projective resolution

$\cdots \xrightarrow{f_3} P_2 \xrightarrow{f_2} P_1 \xrightarrow{f_1} P_0 \xrightarrow{\varepsilon} M \longrightarrow 0,$ of M, consider the complexes

$\cdots \xrightarrow{f_3} P_2 \xrightarrow{f_2} P_1 \xrightarrow{f_1} P_0 \xrightarrow{f_0=0} 0,$ and $\cdots \xrightarrow{f_3 \otimes 1} P_2 \otimes M' \xrightarrow{f_2 \otimes 1} P_1 \otimes M' \xrightarrow{f_1 \otimes 1} P_0 \otimes M' \xrightarrow{f_0 \otimes 1} 0.$

Now recall that $Tor_n(M, M')$ is defined to be $Tor_n(M, M') = \frac{Ker(f_n \otimes 1)}{Im(f_{n+1} \otimes 1)}$.

2.2. Proposition. Let M be an R-module, and suppose that I is an ideal of R with $IM \neq M$. If $Tor_1(\frac{R}{I}, \frac{M}{IM}) = 0$, then IM is an n-almost prime submodule for each $1 < n \in \mathbb{N}$.

Proof. Put $K = IM$. By the short exact sequence $0 \to K \to M \to \frac{M}{K} \to 0$, and according to [7, Theorem 6.26], there is an exact sequence

$$0 = Tor_1(\frac{R}{I}, \frac{M}{K}) \xrightarrow{f} \frac{R}{I} \otimes_R K \xrightarrow{g} \frac{R}{I} \otimes_R M.$$

The natural homomorphism $h : K \longrightarrow \frac{M}{IM}$ induces a homomorphism $\bar{h} : \frac{K}{IK} \longrightarrow \frac{M}{IM}$. Also note that there is an isomorphism $\theta_L : \frac{R}{I} \otimes_R L \longrightarrow \frac{L}{IL}$, for each R-module L.

In the following diagram the rows are exact and it is easy to see that the rectangle is commutative:

$$\begin{array}{ccccccc} & & 0 = Tor_1(\frac{R}{I}, \frac{M}{K}) & \xrightarrow{f} & \frac{R}{I} \otimes_R K & \xrightarrow{g} & \frac{R}{I} \otimes_R M \\ & & & & \downarrow \theta_K & & \downarrow \theta_M \\ 0 & \longrightarrow & Ker\bar{h} & \longrightarrow & \frac{K}{IK} & \xrightarrow{\bar{h}} & \frac{M}{IM} \end{array}$$

It follows that $Ker\bar{h} \cong Kerg = Imf = 0$. On the other hand, $Ker\bar{h} = \frac{IM}{IK}$, hence $K = IM = IK$. Therefore by Lemma 2.1, $(K : M)K = (IM : M)IM = I(IM : M)M = I(IM) = IK = K$, and evidently $(K : M)K = K$ implies that K is an n-almost prime submodule for each $1 < n \in \mathbb{N}$. ■

2.3. Corollary. Let M be an R-module and I an ideal of R with $IM \neq M$. Then IM is an n-almost prime submodule of M, for each $1 < n \in \mathbb{N}$, if one of the following holds:

(i) $\frac{M}{IM}$ is a flat $\frac{R}{Ann\ M}$-module.

(ii) $\frac{R}{I}$ is a flat R-module.

Proof. Put $K = IM$. Note that $\frac{(K:_R M)}{Ann\ M} = (K :_{\frac{R}{Ann\ M}} M)$, thus K is an n-almost prime R-submodule of M, if and only if it is an n-almost prime $\frac{R}{Ann\ M}$-submodule of M. Therefore we can replace $\frac{R}{Ann\ M}$ with R for simplification.

We know that if $\frac{R}{I}$ or $\frac{M}{K}$ is a flat R-module, then $Tor_1(\frac{R}{I}, \frac{M}{K}) = 0$ (see for example [7, Theorem 7.2]). Now the proof follows from Proposition 2.2. ■

The following example shows that the converse of Corollary 2.3 is not necessarily true.

Example. Let $M = R = \mathbb{Z}$, and $I = 2\mathbb{Z}$. Then evidently $2\mathbb{Z}$ is a prime ideal [resp. submodule] of R [resp. the R-module M], with $Ann\ M = 0$. However $\frac{R}{I}$ is not a flat R-module, since it is not torsion-free.

Recall that a ring R is called a Von Neumann regular ring, if for any $a \in R$, $Ra = Ra^2$. By [7, Corollary 4.10], every semi-simple ring is a Von Neumann regular ring.

2.4. Corollary. Let M be an R-module, where R is a Von Neumann regular ring and suppose I is an ideal of R. If $IM \neq M$, then IM is an n-almost prime submodule for each $1 < n \in \mathbb{N}$.

Proof. According to [7, Theorem 4.9], every module over a Von Neumann regular ring is flat. So the proof is given by Corollary 2.3. ■

2.5. Lemma. Let N be an n-almost prime submodule of M.

(i) If there exist $x \in M \setminus N$ and $r \in R \setminus (N : M)$ with $rx \in N$, then $rN \cup (N : M)x \subseteq (N : M)^{n-1}N$.

(ii) If $0 \neq x + N \in \frac{M}{N}$, where $x \in M$, then $(ann(x + N))N \subseteq (N : M)N$.

(iii) $(N : M)N = (\bigcup_{x \in M \setminus N} ann(x + N))N$.

Proof. (i) As N is n-almost prime, $rx \in (N : M)^{n-1}N$. Let y be an arbitrary element of N. Then $y + x \notin N$ and $r(y + x) = ry + rx \in N$ and since N is n-almost prime, $r(y + x) \in (N : M)^{n-1}N$. Therefore $ry \in (N : M)^{n-1}N$, and so $rN \subseteq (N : M)^{n-1}N$.

Now let s be an arbitrary element of $(N : M)$. Clearly $r+s \notin (N : M)$ and $(r+s)x \in N$ and as N is n-almost prime, $(r + s)x \in (N : M)^{n-1}N$. Then since $rx \in (N : M)^{n-1}N$, $sx \in (N : M)^{n-1}N$. Hence $(N : M)x \subseteq (N : M)^{n-1}N$.

(ii) Let $r \in ann(x + N)$. Then $rx \in N$. If $r \in (N : M)$, then clearly $rN \subseteq (N : M)N$. If $r \notin (N : M)$, then in this case by part (i), $rN \subseteq (N : M)^{n-1}N \subseteq (N : M)N$.

(iii) Evidently $(N : M) \subseteq \bigcup_{x \in M \setminus N} ann(x + N)$. Then by part (ii) we have,

$(N : M)N \subseteq (\bigcup_{x \in M \setminus N} ann(x + N))N \subseteq \bigcup_{x \in M \setminus N}(ann(x + N)N) \subseteq (N : M)N$. ■

2.6. Proposition. Let I be an ideal of a ring R and N a submodule of an R-module M.

(i) If $IM \neq IN$, $IN \neq N$, then $K = IN$ is n-almost prime if and only if $K = (K : M)^{n-1}K$.

(ii) If for some positive integer $k > 1$, $I^{k-1}M \neq I^kM = K$, then K is n-almost prime if and only if $K = (K : M)^{n-1}K$. Consequently in this case K is almost prime if and only if K is n-almost prime, for any (or some) positive integer $n \geq 3$.

(iii) Let R be an integral domain and M a Noetherian module with $ann(N) = 0$. Then for every proper ideal I of R with $IM \neq IN$, IN is not n-almost prime.

Proof. (i) If $K = (K : M)^{n-1}K$, then clearly K is n-almost prime. Now assume K is n-almost prime. Evidently K is almost prime. If $K \neq (K : M)K$, then consider $a \in I$ and $x \in N$, where $ax \notin (K : M)K$. Then since $ax \in IN = K \setminus (K : M)K$, either $a \in (K : M)$ *or* $x \in K$. Let $a \in (K : M)$. As $K = IN \subset IM$, $I \nsubseteq (K : M)$ and so we can choose an element $r \in I \setminus (K : M)$. As $rx \in IN = K$, Lemma 2.5(i) implies that $(K : M)x \subseteq (K : M)K$, and so $ax \in (K : M)K$.

Now suppose that $x \in K$. By our assumption $N \nsubseteq K$, hence there exists $z \in N \setminus K$. Note that $az \in IN = K$. Again by Lemma 2.5(i), $aK \subseteq (K : M)K$. Then in this case $ax \in (K : M)K$.

Therefore $K = (K : M)K$, and consequently $K = (K : M)^{n-1}K$.

(ii) We have $IM \neq K$, otherwise $I^{k-1}M \subseteq IM = K = I^kM \subseteq I^{k-1}M$, which is impossible. Now apply part (i) for $N = I^{k-1}M$.

(iii) Clearly $N \neq IN$, otherwise by Nakayama's lemma, there exists $s \in I$ such that $(s+1)N = 0$ and since $ann(N) = 0$, $1 = -s \in I$, which is a contradiction with the fact that $I \neq R$.

Note that $(IN : M)N \subseteq IN$. If IN is n-almost prime, then IN is almost prime and so by part (i), $IN = (IN : M)IN = I(IN : M)N \subseteq I^2N \subseteq IN$, that is $IN = I^2N$. Again by Nakayama's lemma, for some $t \in I$, $(t+1)IN = 0$. As $ann(N) = 0$, $(t+1)I = 0$. So $1 = -t \in I$ or $I = 0$, which is a contradiction with the fact that $I \neq R$ and $IM \neq IN$. Consequently IN is not n-almost prime. ■

Recall that a ring R is said to be ZPI-ring, if every non-zero proper ideal of R can be written as a product of prime ideals of R (see [6, Chapters VI and IX]). According to [6, Theorem 9.10], every ZPI-ring is a Noetherian ring.

2.7. Theorem. Let M be an R-module and I an ideal of R with $IM \neq M$.

(i) If R is a ZPI-ring and IM is an n-almost prime submodule, then $IM = I^nM$, or $IM = PM$, where P is a prime ideal of R.

(ii) If R is a Dedekind domain, then IM is an n-almost prime submodule if and only if $IM = I^nM$ or IM is a prime submodule of M.

(iii) If (R, m) is a local ZPI-ring and IM is finitely generated, then IM is n-almost prime if and only if $IM = 0$ *or* $IM = mM$.

Proof. (i) Let $I = P_1^{k_1}...P_m^{k_m}$, where $P_i's$ are distinct prime ideals of R and $k_i's$ are positive integers.

Assume that $IM \neq PM$ for each prime ideal P of R. Then $IM = P_1^{k_1}...P_m^{k_m}M$ and without loss of generality we may suppose that $IM \neq P_1^{k_1-1}P_2^{k_2}...P_m^{k_m}M$ and $(k_1 - 1) + k_2 + k_3 + \cdots + k_m > 0$.

Put $N = P_1^{k_1-1}P_2^{k_2}...P_m^{k_m}M$ and $K = IM$. Then $K = P_1N$ and $P_1M \neq K$ and $K \neq N$, then by Proposition 2.6(i), $K = (K : M)^{n-1}K$, that is $IM = (IM : M)^{n-1}(IM)$, and by Lemma 2.1(i), $(IM : M)^{n-1}(IM) = I(IM : M)^{n-1}M = I^nM$. Thus $IM = I^nM$.

(ii) Let R be a Dedekind domain and suppose IM is an n-almost prime submodule. By part (i), $IM = I^nM$, or $IM = PM$, where P is a prime ideal of R.

If $IM = PM$, where P is a prime ideal of R, then $P = 0$ or P is a maximal ideal of R. Evidently $P = 0$ implies that $I^nM = 0 = IM$. Now suppose P is a maximal ideal of R. As $P \subseteq (PM : M)$, we have $P = (PM : M)$ or $PM = M$. By our hypothesis $PM = IM \neq M$, then $(IM : M) = (PM : M) = P$ and so IM is a prime submodule of M, by Lemma 2.1(ii).

Now for the converse, suppose that $IM = I^nM$. Then by Lemma 2.1(iii), IM is n-almost prime.

(iii) If $IM = mM$, then by Lemma 2.1(ii), mM is a prime submodule. Also clearly 0 is an n-almost prime submodule.

Now assume that IM is an n-almost prime submodule of M. By [6, Theorem 9.10], R is a Noetherian ring. If $m = m^2$, by Nakayama's lemma, $m = 0$, then R is a field and so $IM = 0$.

Now let $m^2 \neq m$. Choose $x \in m \setminus m^2$. Then $m^2 \subset m^2 + Rx \subseteq m$. By [6, Theorem 9.10], there are no ideals of R strictly between m^2 and m. So $m^2 + Rx = m$ and by Nakayama's lemma, $m = Rx$.

Now let P be a non-zero prime ideal of R, and $0 \neq y \in P$. By the Krull Intersection Theorem, we have $\cap_{n=1}^{+\infty} m^n = 0$. Thus there is a positive integer k such that $y \in m^k$ and $y \notin m^{k+1}$. Since $y \in m^k = Rx^k$, there exists an element $u \in R$ such that $y = ux^k$, and since $y \notin m^{k+1}$, $u \notin m$. Then u is a unit element of R. Hence $x^k = u^{-1}y \in P$. We know that P is a prime ideal of R, so $x \in P$, that is $m = P$. Whence m is the only nonzero prime ideal of R. Now by part (i), $IM = I^nM$ or $IM = mM$

If $IM = mM$, then Lemma 2.1(ii) implies that IM is a prime submodule.

In case $IM = I^nM$, Nakayama's lemma implies that $IM = 0$. ■

The following result is an obvious consequence of the above theorem.

2.8. Corollary. Let R be a ZPI-ring and I a proper ideal of R.

(i) I is an n-almost prime ideal if and only if $I = I^n$ or I is a prime ideal.
(ii) If (R, m) is a local ring, then I is an n-almost prime ideal if and only if $I = 0$ *or* $I = m$.

2.9. Proposition. Let M be an R-module, and I an ideal which is a product of a finite number of maximal ideals of R. Then IM is an n-almost prime submodule if and only if IM is a prime submodule of M, or $IM = I^nM$.

Proof. For each maximal ideal P of R, we have $P \subseteq (PM : M)$, then by Lemma 2.1(ii), PM is a prime submodule or $PM = M$. Thus if IM is an n-almost prime submodule, which is not a prime submodule, then there exist maximal ideals P_i, $1 \leq i \leq m$ and positive numbers k_i, $1 \leq i \leq m$ such that $IM = P_1^{k_1}P_2^{k_2}\cdots P_m^{k_m}M$ and $IM \neq P_1^{k_1-1}P_2^{k_2}\cdots P_m^{k_m}M$. Therefore if we put $N = P_1^{k_1-1}P_2^{k_2}\cdots P_m^{k_m}M$ and $K = P_1N$, since K is not prime, we get $K \neq P_1M$, also $K = P_1N \neq N$, hence by Proposition 2.6(i), $K = (K : M)^{n-1}K$.

Consequently by Lemma 2.1(i), $K = IM = (IM : M)^{n-1}(IM) = I(IM : M)^{n-1}M = I^nM$.

For the converse suppose $IM = I^nM$. Then according to Lemma 2.1(iii), IM is n-almost prime. ■

Recall that a *multiplicatively closed subset* of a ring R is a subset S such that $0 \notin S$ and $1 \in S$ and $xy \in S$ for each $x, y \in S$.

The following result studies when the localization of an n-almost prime submodule is n-almost prime.

2.10. Proposition. Let N be an n-almost prime submodule of an R-module M, and S a multiplicatively closed subset of R.

(i) If $S \cap (N : M) = \emptyset$ and for some $x \in M \setminus N$, $S \cap ((N : M)^{n-1}N : x) = \emptyset$, then $S^{-1}N \neq S^{-1}M$.
(ii) If $S^{-1}N \neq S^{-1}M$, then $S^{-1}N$ is an n-almost prime submodule of $S^{-1}M$.

Proof. (i) Let $x \in M \setminus N$. If $S^{-1}N = S^{-1}M$, then there exists an element $s \in S$ such that $sx \in N$. Since $S \cap ((N : M)^{n-1}N : x) = \emptyset$, $sx \notin (N : M)^{n-1}N$. As N is an n-almost prime submodule and $x \notin N$, $s \in (N : M) \cap S$, which is a contradiction. Hence $S^{-1}N \neq S^{-1}M$.

(ii) Let for $\frac{r}{s} \in S^{-1}R$, $\frac{y}{t} \in S^{-1}M$, $\frac{r}{s}\frac{y}{t} \in S^{-1}N \setminus (S^{-1}N : S^{-1}M)^{n-1}S^{-1}N$. Then there exists an element $u \in S$ such that $ury \in N$. If $ury \in (N : M)^{n-1}N$, then $\frac{ry}{st} = \frac{ury}{ust} \in S^{-1}((N : M)^{n-1}N) \subseteq (S^{-1}N : S^{-1}M)^{n-1}S^{-1}N$, a contraction. Hence $ury \in N \setminus (N : M)^{n-1}N$. As N is almost prime, either $ur \in (N : M)$ or $y \in N$, so either $\frac{r}{s} = \frac{ur}{us} \in S^{-1}(N : M) \subseteq (S^{-1}N : S^{-1}M)$ or $\frac{y}{t} \in S^{-1}N$. ■

3. Essential multiplicatively closed subsets

Recall that an ideal I of a ring R is said to be *essential* if $I \cap J \neq 0$, for each non-zero ideal J of R (that is $J \nsubseteq 0$). In this section we introduce a similar notion for multiplicatively closed subsets of R, and we find some connections between this notion and n-almost primes.

Definition. Let S be a multiplicatively closed subset of R and P a prime ideal of R with $S \cap P = \emptyset$. Then S will be called P-essential, if $S \cap J \neq \emptyset$, for each ideal J with $J \nsubseteq P$.

Evidently $R \setminus P$ is a P-essential multiplicatively closed subset, for each prime ideal P of R.

Recall that a multiplicatively closed subset S of R is said to be *saturated* if

$$xy \in S \Longleftrightarrow x, y \in S.$$

The following lemma is a well known result (see [2, p. 44, Exercise 7 (ii)]).

3.1. Lemma. Let S be a multiplicatively closed subset of R. Then

$$\bar{S} = R \setminus \cup\{P | P \text{ is a prime ideal with } P \cap S = \emptyset\}$$

is a saturated multiplicatively closed subset of R containing S and there is no saturated multiplicatively closed subset of R strictly between S and $\bar{S}$.

It is obvious that for each prime ideal P of R, the ring R_P is a local ring and the ideal P_P is a maximal ideal of R_P. The following result shows that $S^{-1}R$ being a local ring is indeed related to P-essentiality of S.

Let S be a multiplicatively closed subset of R. For any ideal J of $S^{-1}R$, we consider $J^c = \{r \in R \mid r/1 \in J\}$.

3.2. Proposition. Let S be a multiplicatively closed subset of R and P a prime ideal of R with $S \cap P = \emptyset$. Then the following are equivalent:

(i) S is P-essential;
(ii) $S^{-1}R = R_P$;
(iii) $\bar{S} = R \setminus P$;
(iv) $S^{-1}P$ is the only maximal ideal of $S^{-1}R$.

Proof. *(i)* $\Rightarrow$ *(ii)* Clearly $S^{-1}R \subseteq R_P$, since $S \subseteq R \setminus P$. Now suppose that $\frac{y}{t} \in R_P$. Hence as $t \in R \setminus P$, for some $r \in R$, we have $rt \in S \subseteq R \setminus P$, and so $r \in R \setminus P$. Then $\frac{y}{t} = \frac{ry}{rt} \in S^{-1}R$ and hence $S^{-1}R = R_P$.

(ii) $\Rightarrow$ *(iii)* Since $P \cap S = \emptyset$, by Lemma 3.1, $\bar{S} \subseteq R \setminus P$. Now let $r \in R \setminus P$. We have $1/r \in R_P = S^{-1}R$, then there exists $s \in S$, $x \in R$ with $1/r = x/s$. Thus for some $s' \in S$ we have $s'rx = ss' \in S \subseteq \bar{S}$, and so $r \in \bar{S}$, because $\bar{S}$ is saturated.

(iii) $\Rightarrow$ *(iv)* Let m be a maximal ideal of $S^{-1}R$. Then m^c is a prime ideal of R with $m^c \cap S = \emptyset$. Note that $R \setminus (P \cup m^c)$ is a saturated multiplicatively closed subset of R and since $S \subseteq R \setminus (P \cup m^c) \subseteq (R \setminus P) = \bar{S}$, Lemma 3.1 implies that $R \setminus (P \cup m^c) = (R \setminus P) = \bar{S}$. Hence $m^c \subseteq (P \cup m^c) = P$, and thus $m = S^{-1}(m^c) \subseteq S^{-1}P$, and so $m = S^{-1}P$, because of maximality of m.

(iv) $\Rightarrow$ *(i)* Let J be an ideal of R such that $J \nsubseteq P$. If $J \cap S = \emptyset$, since $S^{-1}P$ is the only maximal ideal of $S^{-1}R$, we have $S^{-1}J \subseteq S^{-1}P$. So $J \subseteq P$, which is impossible. Consequently S is P-essential. ■

3.3. Theorem. Let N be an n-almost prime submodule of an R-module M with $I = (N : M)$. Then $S = [(R \setminus I) \cup (I^{n-1}N : M)] \setminus P$ is P-essential, for each prime ideal P of R.

Proof. First to prove that S is multiplicatively closed, let $r, s \in S$. If $rs \notin S$, then $rs \in I$. Also $rs \notin P$, because if $rs \in P$, then $r \in P$ or $s \in P$, although $S \cap P = \emptyset$. Thus $rs \in I \setminus P$.

If $r \in (I^{n-1}N : M)$ or $s \in (I^{n-1}N : M)$, then $rs \in (I^{n-1}N : M) \setminus P$, and so $rs \in S$.

Now on the contrary suppose $r, s, rs \notin (I^{n-1}N : M)$. Hence there exists $m \in M$ such that $rsm \notin I^{n-1}N$, and we know that $rs \in I = (N : M)$, therefore $rsm \in N \setminus I^{n-1}N$.

As $r, s \in S \subseteq [(R \setminus I) \cup (I^{n-1}N : M)]$ and $r, s \notin (I^{n-1}N : M)$, we have $r, s \notin I = (N : M)$. Note that $rsm \in N \setminus I^{n-1}N$ and $r, s \notin (N : M)$ and N is n-almost prime, thus $m \in N$.

Now consider $m' \in M \setminus N$. If $rsm' \notin I^{n-1}N$, the above argument shows that $m' \in N$, which is impossible.

Then we may assume $rsm' \in I^{n-1}N$. Thus for $x = m + m'$, we have $rsx \in N \setminus I^{n-1}N$. Now since $m \in N$ and $m' \notin N$, we have $x \notin N$, consequently $r \in (N : M) = I$ or $s \in (N : M) = I$, which is a contradiction.

Next we will prove that S is P-essential. Let J be an ideal of R such that $J \nsubseteq P$. If $I \subseteq P$, then $S = R \setminus P$, and obviously S is P-essential. So suppose that $I \nsubseteq P$.

If $J \cap S = \emptyset$, it is easy to see $J \cap [(R \setminus I) \cup (I^{n-1}N : M)] \subseteq P$ and so $J \subseteq I \cup P$. Therefore $J \subseteq I$

Note that $I = (N : M)$, so $I^n M = I^{n-1}(N : M)M \subseteq I^{n-1}N$, that is $I^n \subseteq (I^{n-1}N : M)$. Hence $J^n \subseteq J \cap I^n \subseteq J \cap (I^{n-1}N : M) \subseteq P$, which is impossible. Consequently $J \cap S \neq \emptyset$ and so S is P-essential. ■

3.4. Corollary. Let I be an ideal of R such that $N(R) \subseteq I^n$ and consider $S_P = [(R \setminus I) \cup I^n] \setminus P$. Then the following are equivalent:

(i) I is n-almost prime;
(ii) S_P is multiplicatively closed for any prime ideal P;
(iii) S_P is multiplicatively closed for any minimal prime ideal P.

Proof. $(i) \Rightarrow (ii)$ The proof is given by Theorem 3.3.

$(ii) \Rightarrow (iii)$ The proof is evident.

$(iii) \Rightarrow (i)$ Let $ab \in I \setminus I^n$. So $ab \notin N(R)$ and there exists a minimal prime ideal P of R such that $ab \notin P$. Then $ab \notin S_P$. Hence $a \notin S_P$ or $b \notin S_P$. Therefore $a \in I$ or $b \in I$. ■

3.5. Corollary. Let R be an integral domain and M an R-module.

(i) If N is an n-almost prime submodule of M with $(N : M) = I$, then $S = [(R \setminus I) \cup (IN : M)] \setminus \{0\}$ is a multiplicatively closed subset of R and $S^{-1}R$ is a field.
(ii) An ideal I of R is n-almost prime if and only if $S = [(R \setminus I) \cup I^n] \setminus \{0\}$ is a multiplicatively closed subset of R. When this is the case, $S^{-1}R$ is a field.

Proof. (i) The proof is given by Theorem 3.3 and Proposition 3.2.

(ii) The proof of the first part is given by Corollary 3.4. By part (i), $S^{-1}R$ is a field, if I is n-almost prime. ■

The following remark studies the converse of the above corollary.

Remark. Let S be a saturated multiplicatively closed subset of R such that $S^{-1}R$ is a field. Then there exist prime ideals I and P of R such that $S = [(R \setminus P) \cup I^2] \setminus P$.

Proof. Since S is a multiplicatively closed subset of R, there exists a prime ideal P of R with $P \cap S = \emptyset$. Then $S^{-1}P$ is a proper ideal of $S^{-1}R$ and $S^{-1}R$ is a field, so $S^{-1}P$ is the only maximal ideal of $S^{-1}R$. Hence by Proposition 3.2, $\bar{S} = R \setminus P$. Note that S is a saturated multiplicatively closed subset of R, then by Lemma 3.1, $S = \bar{S} = R \setminus P$. Thus it is enough to consider $I = P$. ■

Acknowledgment. The authors would like to thank the referee for his comments on this paper.

References

[1] Anderson, D.D. and Batainen, Malik *Generalizations of primes ideals*, Comm. Algebra, **36**, 86–696, 2008.

[2] Atiah, M. F. and Mcdonald, I. G. *Introduction to commutative algebra* (Addison-Wesley, 1969).

[3] Azizi, A. *Prime submodules and flat modules*, Acta Math. Sinica, English Series, **23** (1), 147–152, 2007.

[4] Azizi, A. *Strongly irreducible ideals*, J. Aust. Math. Soc., **84**, 145–154, 2008.

[5] Bhatwadekar, S. M. and Sharma, P. K. *Uniqe factorization and birth of almost prime*, Comm. Algebra, **33**, 43–49, 2005.

[6] Larsen, M. D. and McCarthy, P. J. *Multiplicative theory of ideals* (Academic press, Inc., 1971).

[7] Rotman, J. J. *An introduction to homological algebra, second edition* (Springer, 2009).

[8] Tiras, Y. and Harmanci, A. and Smith, P. F. *A characterization of prime submodules*, J. Algebra, **212** (2), 743–752, 1999.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 597 – 606

Common fixed point theorems in cone Banach type spaces

Fridoun Moradlou * and Peyman Salimi †

Abstract

In this paper, we give some generalized theorems on points of coincidence and common fixed points for two weakly compatible mappings on a cone Banach type space.

2000 AMS Classification: Primary 47H10, 54H25, 55M20.

Keywords: cone normed type spaces, common fixed point, weakly compatible mappings

Received 15/03/2012 : Accepted 18/06/2013 Doi : 10.15672/HJMS.20154410017

1. Introduction

In 1980, Rzepecki [15] provide a generalization of metric spaces. He defined a metric d_E on a set X by $d_E : X \times X \to S$, where E is a Banach space and S is a normal cone in E with partial order $\preceq$, and he generalized the fixed point theorems of Maia type. In 1987, Lin [9] introduced the notion of K-metric spaces and considered some results of Khan and Imdad [7] in K-metric spaces. In 2007, Huang and Zhang [8] introduced cone metric spaces and defined some properties of convergence of sequences and completeness in cone metric spaces, also they proved a fixed point theorem of cone metric spaces. Beginning around the year 2007, the fixed point theorems in cone metric spaces have been extensively proved by a number of authors and there are many interesting results concerning these theorems (see [1]–[3], [5], [11]–[14]).

In this paper, we propose the notion of cone Banach type spaces and prove the generalization of some known results on points of coincidence and the generalization of some common fixed point theorems for two weakly compatible mappings in cone Banach type spaces.

*Department of Mathematics, Sahand University of Technology, Tabriz, Iran.
Email: moradlou@sut.ac.ir, fridoun.moradlou@gmail.com

†Department of Mathematics, Sahand University of Technology, Tabriz, Iran.
Email:peyman_salimi_63@yahoo.com

2. preliminaries

2.1. Definition. [11] Let E be a real Banach space with norm $\|.\|$ and P be a subset of E. P is called a cone if and only if the following conditions are satisfied:

(P1) P is closed, nonempty and $P \neq \{0\}$;
(P1) $a, b \geq 0$ and $x, y \in P \Rightarrow ax + by \in P$;
(P3) $x \in P$ and $-x \in P \Rightarrow x = 0$.

Let $P \subset E$ be a cone, we define a partial ordering $\preceq$ on E with respect to P by $x \preceq y$ if and only if $y - x \in P$. we write $x \prec y$ whenever $x \preceq y$ and $x \neq y$, while $x \ll y$ will stand for $y - x \in intP$(interior of P). The cone $P \subset E$ is called normal if there is a positive real number k such that for all $x, y \in E$,

$$0 \preceq x \preceq y \Rightarrow \|x\| \leq k\|y\|.$$

The least positive number satisfying the above inequality is called the normal constant of P. It is clear that $k \geq 1$. Rezapour and Hamlbarani [14] proved that existence of an ordered Banach space E with cone P which is not normal but with $intP \neq \emptyset$.

Throughout this paper, we assume that E is a real Banach space and P is a cone suth that $intP \neq \emptyset$.

2.2. Definition. [11]. Let X be a nonempty set. A function $d : X \times X \to E$ is said to be a cone b-metric function on X with the constant $K \geq 1$ if the following conditions are satisfied:

(1) $0 \preceq d(x, y)$, for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$;
(2) $d(x, y) = d(y, x)$ for all $x, y \in X$;
(3) $d(x, z) \preceq K\Big(d(x, y) + d(y, z)\Big)$ for all $x, y, z \in X$;

then the pair (X, d) is called the cone b-metric space (or cone metric type space (in brief $CMTS$)).

2.3. Definition. [5] Let X be a vector space over $\mathbb{R}$. Suppose the mapping $\|.\|_P : X \to E$ satisfies:

(i) $\|x\|_P \succ 0$ for all $x \in X$;
(ii) $\|x\|_P = 0$ if and only if $x = 0$;
(iii) $\|x + y\|_P \preceq \|x\|_P + \|y\|_P$ for all $x, y \in X$;
(iv) $\|kx\|_P = |k|\|x\|_P$ for all $x \in X$ and all $k \in \mathbb{R}$;

then $\|.\|_P$ is called cone norm on X and the pair $(X, \|.\|_P)$ is called a cone normed space (in brief CNS). Note that each CNS is cone metric space (in brief CMS). Indeed, $d(x, y) = \|x - y\|_P$.

Similar to the definition of $CMTS$, we give the following definition:

2.4. Definition. Let X be a vector space over $\mathbb{R}$. Suppose the mapping $\|.\|_P : X \to E$ satisfies:

(i) $\|x\|_P \succeq 0$ for all $x \in X$;
(ii) $\|x\|_P = 0$ if and only if $x = 0$;
(iii) $\|x + y\|_P \preceq K\Big(\|x\|_P + \|y\|_P\Big)$ for all $x, y \in X$ and for constant $K \geq 1$ (triangle - type inequality);
(iv) $\|rx\|_P = |r|\|x\|_P$ for all $x \in X$ and all $r \in \mathbb{R}$;

then the pair $(X, \|.\|_P)$ is called a *cone normed type space* (in brief $CNTS$).

Note that each $CNTS$ is $CMTS$. Indeed, $d(x, y) = \|x - y\|_P$.

2.5. Example. *Let* $C_b(X) = \{f : X \to \mathbb{C} : \sup_{x\in X} |f(x)| < \infty\}$. *Define* $\|.\|_P : C_b(X) \to \mathbb{R}$ *by*

$$\|f\|_P = \sqrt[3]{\sup_{x\in X} |f(x)|^3}.$$

Then $\|.\|_P$ *satisfies the following properties:*

(i) $\|f\|_P > 0$ *for all* $f \in C_b(X)$;

(ii) $\|f\|_P = 0$ *if and only if* $f = 0$;

(iii) $\|f+g\|_P \leq \sqrt[3]{4}\Big(\|f\|_P + \|g\|_P\Big)$ *for all* $f, g \in C_b(X)$;

(iv) $\|rf\|_P = |r|\|f\|_P$ *for all* $f \in C_b(X)$ *and all* $r \in \mathbb{R}$.

2.6. Definition. Let $(X, ||.||_P, K)$ be a $CNTS$, let $\{x_n\}$ be a sequence in X and $x \in X$. Then

(i) $\{x_n\}$ converges to x whenever for every $c \in E$ with $0 \ll c$ there is a natural number N, such that $\|x_n - x\|_P \ll c$ for all $n > N$. It is denoted by $\lim_{n\to\infty} x_n = x$ or $x_n \to x$ as $n \to \infty$;

(ii) $\{x_n\}$ is a Cauchy sequence whenever for every $c \in E$ with $0 \ll c$ there is a natural number N, such that $\|x_n - x_m\|_P \ll c$ for all $n, m > N$;

(iii) $(X, \|.\|_P, K)$ is a complete cone normed type space if every Cauchy sequence is convergent. Complete cone normed type spaces will be called *cone Banach type spaces.*

2.7. Lemma. *Let* $(X, \|.\|_P, K)$ *be a* $CNTS$, P *be a normal cone with normal constant* M, *and* $\{x_n\}$ *be a sequence in* X. *Then,*

(i) *the sequence* $\{x_n\}$ *converges to* x *if and only if* $\|x_n - x\|_P \to 0$, *as* $n \to \infty$;

(ii) *the sequence* $\{x_n\}$ *is Cauchy if and only if* $\|x_n - x_m\|_P \to 0$ *as* $n, m \to \infty$;

(iii) *if the sequence* $\{x_n\}$ *converges to* x *and the sequence* $\{y_n\}$ *converges to* y, *then* $\|x_n - y_n\|_P \to \|x - y\|_P$.

Proof. The proof is similar to proof of Lemmas 1-5 of [8], by taking $d(x,y) = \|x-y\|_P$. □

From now on, we assume that P is a normal cone with $intP \neq \emptyset$.

2.8. Lemma. *Let* $\{y_n\}$ *be a sequence in a cone Banach type space* $(X, \|.\|_P, K)$ *such that*

$$d(y_n, y_{n+1}) \leq \lambda d(y_{n-1}, y_n),$$

for some $0 < \lambda < 1/K$ *and all* $n \in \mathbb{N}$, *where* $d(x,y) = \|x - y\|_P$. *Then* $\{y_n\}$ *is a Cauchy sequence in* $(X, \|.\|_P, K)$.

2.9. Definition. Let S and T be two self-mappings on a cone metric type space (X, d). A point $z \in X$ is called a coincidence point of S and T if $Sz = Tz$, and it is called a common fixed point of S and T if $Sz = z = Tz$. Moreover, a pair of self-mappings (S, T) is called weakly compatible on X if they commute at their coincidence points, i.e.,

$$z \in X, \quad Sz = Tz \Rightarrow STz = TSz.$$

2.10. Theorem. *Let* C *be a subset of a cone Banach type space* $(X, \|.\|_P, K)$ *and* $d : X \times X \to E$ *be such that* $d(x,y) = \|x - y\|_P$. *Suppose that* $F, T : C \longrightarrow C$ *are two mappings such that* $TC \subset FC$ *and* FC *is closed and convex. If there exists some constant* $1 - \frac{1}{K} < \frac{r}{2} < 1$ *such that*

$$d(Fy, Ty) + rd(Fx, Fy) \preceq d(Fx, Tx), \tag{2.1}$$

for all $x, y \in C$, *then* F *and* T *have at least one point of coincidence. Moreover, if* F *and* T *are weakly compatible, then* F *and* T *have a unique common fixed point.*

Proof. Let $x_0 \in C$ be arbitrary. we define a sequence $\{Fx_n\}$ in the following relation:

$$Fx_{n+1} := \frac{Fx_n + Tx_n}{2}, \quad n = 0, 1, 2, \cdots. \tag{2.2}$$

We see that

$$Fx_n - Tx_n = 2\left(Fx_n - \Big(\frac{Fx_n + Tx_n}{2}\Big)\right) = 2(Fx_n - Fx_{n+1}), \tag{2.3}$$

which implies

$$d(Fx_n, Tx_n) = ||Fx_n - Tx_n||_P = 2||Fx_n - Fx_{n+1}||_P = 2d(Fx_n, Fx_{n+1}), \tag{2.4}$$

for $n = 0, 1, 2, \cdots$. Now, letting $x = x_{n-1}$ and $y = x_n$ in (2.1), using (2.4), we can conclude that

$$2d(Fx_n, Fx_{n+1}) + rd(Fx_{n-1}, Fx_n) \preceq 2d(Fx_{n-1}, Fx_n). \tag{2.5}$$

So

$$d(Fx_n, Fx_{n+1}) \preceq (1 - \frac{r}{2})d(Fx_{n-1}, Fx_n), \tag{2.6}$$

where $1 - \frac{r}{2} < \frac{1}{K}$. Hence by Lemma 2.8, $\{Fx_n\}$ is a Cauchy sequence in FC. Then there exists $z \in C$ such that $Fx_n \to Fz$. Also by (2.2) we can obtain $Tx_n \to Fz$. So by (2.1) we have

$$d(Fz, Tz) \preceq d(Fz, Tz) + rd(Fx_n, Fz) \preceq d(Fx_n, Tx_n). \tag{2.7}$$

Therefore by taking the limit as $n \to \infty$ in (2.7), we obtain $d(Fz, Tz) = 0$, that is, z is a point of coincidence of F and T. Therefore F and T have at least one point of coincidence.

Put $w = Fz = Tz$. If F and T are weakly compatible mappings, then $FTz = TFz$, so $Fw = Tw$.

Now, we show that w is a fixed point of F. Putting $x = w$ and $y = z$ in (2.1), we get

$$d(Fz, Tz) + rd(Fw, Fz) \preceq d(Fw, Tw). \tag{2.8}$$

Hence $d(Fw, Fz) = 0$. That is, $Fw = w$. Therefore $Fw = Tw = w$. So we conclude that $w = Fw = Tw$ is a common fixed point of F and T.

To prove the uniqueness of w, suppose that w_1 is another common fixed point F and T. Replacing x and y by w and w_1 in (2.1), respectively, we get

$$d(Fw_1, Tw_1) + rd(Fw, Fw_1) \preceq d(Fw, Tw). \tag{2.9}$$

Thus,

$$d(w_1, w) \preceq 0.$$

So $w = w_1$. Then w is the unique common fixed point of F and T. □

2.11. Corollary. *Let C be a closed and convex subset of a cone Banach type space $(X, \|.\|_P, K)$ and $d : X \times X \to E$ be such that $d(x, y) = \|x-y\|_P$. Suppose that $T : C \to C$ is a mapping for which there exists some constant $1 - \frac{1}{K} < \frac{r}{2} < 1$ such that*

$$d(y, Ty) + rd(x, y) \preceq d(x, Tx),$$

for all $x, y \in C$. Then T has a unique fixed point.

2.12. Theorem. *Let C be a subset of a cone Banach type space $(X, \|.\|_P, K)$ such that $1 < K \leq 2$. Let $d : X \times X \to E$ be such that $d(x, y) = \|x - y\|_P$. Suppose that $F, T : C \longrightarrow C$ are two mappings such that $TC \subset FC$ and FC is closed and convex. If there exists some constant $1 - \frac{1}{K} < \frac{r}{2} < 1$ such that*

$$d(Tx, Ty) + (1 - \frac{1}{K})d(Fy, Ty) + rd(Fx, Fy) \preceq \frac{1}{2}d(Fx, Tx), \tag{2.10}$$

for all $x, y \in C$, then F and T have at least one point of coincidence. Moreover, if F and T are weakly compatible, then F and T have a unique common fixed point.

Proof. Similar to proof of Theorem 2.10, we construct the sequence $\{Fx_n\}$, therefore

$$Fx_n - Tx_{n-1} = \frac{Fx_{n-1} + Tx_{n-1}}{2} - Tx_{n-1} = \frac{Fx_{n-1} - Tx_{n-1}}{2},$$

which implies that

$$d(Fx_n, Tx_{n-1}) = \frac{1}{2} d(Fx_{n-1}, Tx_{n-1}). \tag{2.11}$$

Using the triangle-type inequality, we get

$$d(Fx_n, Tx_n) - Kd(Fx_n, Tx_{n-1}) \preceq Kd(Tx_{n-1}, Tx_n) \tag{2.12}$$

It follows from (2.3) and (2.11) that

$$\frac{2}{K} d(Fx_n, Fx_{n+1}) - d(Fx_n, Fx_{n-1}) \preceq d(Tx_{n-1}, Tx_n). \tag{2.13}$$

Replacing x and y by x_{n-1} and x_n in (2.10) and using (2.3) and (2.13), we can obtain

$$\begin{aligned} \frac{2}{K} d(Fx_n, Fx_{n+1}) - d(Fx_{n-1}, Fx_n) \quad &+ \quad 2(1 - \frac{1}{K}) d(Fx_n, Fx_{n+1}) \\ &+ \quad rd(Fx_{n-1}, Fx_n) \preceq d(Fx_{n-1}, Fx_n). \end{aligned}$$

Thus,

$$d(Fx_n, Fx_{n+1}) \preceq (1 - \frac{r}{2}) d(Fx_{n-1}, Fx_n),$$

where $1 - \frac{r}{2} < \frac{1}{K}$. Hence by Lemma 2.8, $\{Fx_n\}$ is a Cauchy sequence in FC. Then there exists $z \in C$ such that $Fx_n \longrightarrow Fz$. Substituting $x = x_n$ and $y = z$ in (2.10), we get

$$\begin{aligned} (1 - \frac{1}{K}) d(Fz, Tz) &\preceq d(Tx_n, Tz) + (1 - \frac{1}{K}) d(Fz, Tz) + rd(Fx_n, Fz) \\ &\preceq \frac{1}{2} d(Fx_n, Tx_n). \end{aligned} \tag{2.14}$$

Therefore by taking the limit as $n \to \infty$ in (2.14), we obtain $d(Fz, Tz) = 0$. Then we conclude that z is a point of coincidence of F and T.

Let $w = Fz = Tz$. If F and T are weakly compatible mappings, then $FTz = TFz$, so $Fw = Tw$.

Now, we show that w is a fixed point of F. Putting $x = w$ and $y = z$ in (2.10), we have

$$d(Tw, Tz) + (1 - \frac{1}{K}) d(Fz, Tz) + rd(Fw, Fz) \preceq \frac{1}{2} d(Fw, Tw).$$

Then

$$(r + 1) d(Fw, w) \preceq 0.$$

Therefore $Fw = Tw = w$. So we conclude that $w = Fw = Tw$ is a common fixed point of F and T.

To prove the uniqueness of w, suppose that w_1 is another common fixed point of F and T. Replacing x and y by w and w_1 in (2.10), respectively, we get

$$d(Tw, Tw_1) + (1 - \frac{1}{K}) d(Fw_1, Tw_1) + rd(Fw, Fw_1) \preceq \frac{1}{2} d(Fw, Tw). \tag{2.15}$$

Thus,

$$(1 + r) d(w_1, w) \preceq 0.$$

So $w = w_1$. Then w is the unique common fixed point of F and T. □

2.13. Corollary. *Let C be a closed and convex subset of a cone Banach type space $(X, ||.||_P, K)$ such that $1 < K \leq 2$ and $d : X \times X \to E$ be such that $d(x, y) = ||x - y||_P$. Suppose that $T : C \longrightarrow C$ is a mapping which satisfies the condition*

$$d(Tx, Ty) + (1 - \frac{1}{K})d(y, Ty) + rd(x, y) \preceq \frac{1}{2}d(x, Tx),$$

for all $x, y \in C$, where $1 - \frac{1}{K} < \frac{r}{2} < 1$, then T has a unique fixed point.

2.14. Theorem. *Let C be a subset of a cone Banach type space $(X, \|.\|_P, K)$ and $d : X \times X \to E$ be such that $d(x, y) = \|x - y\|_P$. Suppose that $F, T : C \to C$ are two mappings such that $TC \subset FC$ and FC is closed and convex. If there exist a, b, s satisfying*

$$0 < s + |a|K^{\frac{1}{2} - \frac{1}{2}sgn(a)} - 2b < 2(aK^{-sgn(a)} + b), \tag{2.16}$$

and

$$ad(Tx, Ty) + b\Big(d(Fx, Tx) + d(Fy, Ty)\Big) \preceq sd(Fx, Fy), \tag{2.17}$$

for all $x, y \in C$, then F and T have at least one point of coincidence. Moreover if $a > s$ and F and T are weakly compatible, then F and T have a unique common fixed point.

Proof. Similar to proof of Theorem 2.10, we construct the sequence $\{Fx_n\}$. We claim that the inequality (2.17) for $x = x_{n-1}$ and $y = x_n$ implies that

$$\begin{aligned} &2aK^{-\text{sgn}(a)}d(Fx_n,Fx_{n+1}) - |a|K^{\frac{1}{2}-\frac{1}{2}\text{sgn}(a)}d(Fx_{n-1}, Fx_n) \\ &\qquad + 2b\Big(d(Fx_{n-1}, Fx_n) + d(Fx_n, Fx_{n+1})\Big) \preceq sd(Fx_{n-1}, Fx_n), \end{aligned} \tag{2.18}$$

for all a, b, s that satisfy (2.16). To see this, replacing x and y by x_{n-1} and x_n in (2.17), respectively, we obtain

$$ad(Tx_{n-1}, Tx_n) + b\Big(d(Fx_{n-1}, Tx_{n-1}) + d(Fx_n, Tx_n)\Big) \preceq sd(Fx_{n-1}, Fx_n). \tag{2.19}$$

Let $a \geq 0$, using (2.3), (2.13) and (2.19), we have

$$\begin{aligned} &\frac{2a}{K}d(Fx_n, Fx_{n+1}) - ad(Fx_n, Fx_{n-1}) \\ &\qquad + 2b\Big(d(Fx_{n-1}, Fx_n) + d(Fx_n, Fx_{n+1})\Big) \preceq sd(Fx_{n-1}, Fx_n), \end{aligned}$$

which is equivalent to (2.18), since $\text{sgn}(a) = 0$ or 1.

Now suppose that $a < 0$, consider the inequality

$$d(Tx_{n-1}, Tx_n) \preceq K\Big(d(Tx_{n-1}, Fx_n) + d(Fx_n, Tx_n)\Big),$$

which is equivalent to

$$ad(Tx_{n-1}, Tx_n) \succeq Ka\Big(d(Tx_{n-1}, Fx_n) + d(Fx_n, Tx_n)\Big). \tag{2.20}$$

It follows from (2.19) and (2.20) that

$$\begin{aligned} &aK(d(Tx_{n-1}, Fx_n) + d(Fx_n, Tx_n)) \\ &\qquad + b\Big(d(Fx_{n-1}, Tx_{n-1}) + d(Fx_n, Tx_n)\Big) \preceq sd(Fx_{n-1}, Fx_n). \end{aligned} \tag{2.21}$$

Using (2.4), (2.11) and (2.21), we get

$$\begin{aligned} &aKd(Fx_{n-1}, Fx_n) + 2aKd(Fx_n, Fx_{n+1}) \\ &\qquad + 2b\Big(d(Fx_{n-1}, Fx_n) + d(Fx_n, Fx_{n+1})\Big) \preceq sd(Fx_{n-1}, Fx_n) \end{aligned}$$

which is equivalent to (2.18), since $\text{sgn}(a) = -1$. Hence, we established our claim.

It follows from (2.18) that

$$d(Fx_n, Fx_{n+1}) \preceq \frac{s + |a|K^{\frac{1}{2} - \frac{1}{2}\mathrm{sgn}(a)} - 2b}{2(aK^{-\mathrm{sgn}(a)} + b)} d(Fx_{n-1}, Fx_n),$$

where $\frac{s + |a|K^{\frac{1}{2} - \frac{1}{2}\mathrm{sgn}(a)} - 2b}{2(aK^{-\mathrm{sgn}(a)} + b)} < 1$. Hence by Lemma 2.8, $\{Fx_n\}$ is a Cauchy sequence in FC. Then there exists $z \in C$ such that $Fx_n \to Fz$, so $Tx_n \to Fz$. Now, using (2.17), we have

$$ad(Tx_n, Tz) + b\Big(d(Fx_n, Tx_n) + d(Fz, Tz)\Big) \preceq sd(Fx_n, Fz). \tag{2.22}$$

Thus by taking the limit as $n \to \infty$ in (2.22), we obtain

$$(a + b)d(Fz, Tz) \preceq 0.$$

Since $aK^{-\mathrm{sgn}(a)} \leq a$, we get $a + b > 0$. Hence, $d(Fz, Tz) = 0$. So z is a point of coincidence of F and T.

If F and T are weakly compatible, then $FTz = TFz$. Therefore $Fw = Tw$, where $w = Fz = Tz$.

Now, we show that w is a unique common fixed point of T and F. Substituting $x = w$ and $y = z$ in (2.17), we obtain

$$ad(Tw, Tz) + b\Big(d(Fw, Tw) + d(Fz, Tz)\Big) \preceq sd(Fw, Fz),$$

which yields that

$$(a - s)d(Tw, w) \preceq 0.$$

Since $a > s$, we have $Tw = w$. Therefore $Fw = Tw = w$. This means w is a common fixed point of F and T.

To prove the uniqueness of w, suppose that w_1 is another common fixed point of F and T. Replacing x and y by w_1 and w in (2.17), we get

$$ad(Tw_1, Tw) + b\Big(d(Fw_1, Tw_1) + d(Fw, Tw)\Big) \preceq sd(Fw_1, Fw).$$

Thus,

$$(a - s)d(w_1, w) \preceq 0.$$

So $w = w_1$. Therefore w is the unique common fixed point of F and T. □

2.15. Corollary. *Let C be a closed and convex subset of a cone Banach type space $(X, \|.\|_P, K)$ and $d : X \times X \to E$ be such that $d(x, y) = \|x-y\|_P$. Suppose that $T : C \to C$ is a mapping for which there exist a, b, s such that*

$$0 < s + |a|K^{\frac{1}{2} - \frac{1}{2}sgn(a)} - 2b < 2(aK^{-sgn(a)} + b),$$

and

$$ad(Tx, Ty) + b\Big(d(x, Tx) + d(y, Ty)\Big) \preceq sd(x, y),$$

for all $x, y \in C$, then T has at least one fixed point. Moreover, if $a > s$, then T has a unique fixed point.

2.16. Theorem. *Let C be a subset of a cone Banach type space $(X, \|.\|_P, K)$ and $d : X \times X \to E$ be such that $d(x, y) = \|x - y\|_P$. Suppose that $F, T : C \to C$ are two mappings such that $TC \subset FC$ and FC is closed and convex. If there exist a, b satisfying*

$$1 < b < 1 + \frac{(2a - 1)K - 1}{2K^2} \ \ \& \ \ a > \frac{K + 1}{2K}, \tag{2.23}$$

and

$$(2.24)\qquad ad(Fy,Ty)+d(Fy,Tx)\preceq bd(Fx,Tx)+\frac{1}{K}d(Fx,Fy),$$

for all $x,y\in C$, then F and T have at least one point of coincidence. Moreover, if $K>1$ and F and T are weakly compatible, then F and T have a unique common fixed point.

Proof. Let $x_0\in C$ be arbitrary, we define a sequence $\{Fx_n\}$ in the following relation:

$$(2.25)\qquad Fx_{n+1}:=\frac{(2K-1)Fx_n+Tx_n}{2K},\quad n=0,1,2,\cdots,$$

we see that

$$(2.26)\qquad Fx_n-Tx_n=2K\left(Fx_n-\Big(\frac{(2K-1)Fx_n+Tx_n}{2K}\Big)\right)=2K(Fx_n-Fx_{n+1}),$$

which implies

$$(2.27)\qquad d(Fx_n,Tx_n)=2Kd(Fx_n,Fx_{n+1}).$$

Similarly

$$Fx_n-Tx_{n-1}=\frac{(2K-1)Fx_{n-1}+Tx_{n-1}}{2K}-Tx_{n-1}=(\frac{2K-1}{2K})(Fx_{n-1}-Tx_{n-1}),$$

then

$$(2.28)\qquad d(Fx_n,Tx_{n-1})=(\frac{2K-1}{2K})d(Fx_{n-1},Tx_{n-1}).$$

Replacing x and y by x_{n-1} and x_n in (2.24), respectively, we get

$$(2.29)\qquad ad(Fx_n,Tx_n)+d(Fx_n,Tx_{n-1})\preceq bd(Fx_{n-1},Tx_{n-1})+\frac{1}{K}d(Fx_{n-1},Fx_n).$$

It follows from (2.27), (2.28) and (2.29) that

$$2aKd(Fx_n,Fx_{n+1})+(2K-1)d(Fx_n,Fx_{n-1})\preceq 2bKd(Fx_{n-1},Fx_n)+\frac{1}{K}d(Fx_{n-1},Fx_n).$$

Therefore

$$d(Fx_n,Fx_{n+1})\preceq\frac{(2bK+\frac{1}{K}-2K+1)}{2aK}d(Fx_{n-1},Fx_n),$$

where $\frac{(2bK+\frac{1}{K}-2K+1)}{2aK}<\frac{1}{K}$. Hence by Lemma 2.8, $\{Fx_n\}$ is a Cauchy sequence in FC. Then there exists $z\in C$ such that $Fx_n\to Fz$, so $Tx_n\to Fz$. Replacing x and y by x_n and z in (2.24), respectively, we get

$$(2.30)\qquad ad(Fz,Tz)+d(Fz,Tx_n)\preceq bd(Fx_n,Tx_n)+\frac{1}{K}d(Fx_n,Fz).$$

Then by taking the limit as $n\to\infty$ in (2.30), we obtain $d(Fz,Tz)=0$. So we conclude that z is a point of coincidence of F and T.

If F and T are weakly compatible, then $FTz=TFz$. Therefore $Fw=Tw$, where $w=Fz=Tz$.

Now, we show that w is a unique common fixed point of F and T. Substituting $x=w$ and $y=z$ in (2.24), we obtain

$$ad(Fz,Tz)+d(Fz,Tw)\preceq bd(Fw,Tw)+\frac{1}{K}d(Fw,Fz),$$

which implies that

$$(1-\frac{1}{K})d(w,Tw)\preceq 0.$$

Hence $w=Tw$, therefore w is a common fixed point of F and T.

To prove the uniqueness of w, suppose that w_1 is another common fixed point of F and T. Replacing x and y by w and w_1 in (2.24), respectively, we have

$$ad(Fw_1, Tw_1) + d(Fw_1, Tw) \preceq bd(Fw, Tw) + \frac{1}{K}d(Fw, Fw_1).$$

Thus,

$$(1 - \frac{1}{K})d(w, w_1) \preceq 0.$$

So $w = w_1$. Therefore w is the unique common fixed point of F and T. □

2.17. Corollary. *Let C be a closed and convex subset of a cone Banach type space $(X, \|.\|_P, K)$ and $d : X\times X \to E$ be such that $d(x, y) = \|x-y\|_P$. Suppose that $T : C \to C$ is a mapping for which there exist a, b satisfying*

$$1 < b < 1 + \frac{(2a-1)K - 1}{2K^2} \ \ \& \ \ a > \frac{K+1}{2K},$$

and

$$ad(y, Ty) + d(y, Tx) \preceq bd(x, Tx) + \frac{1}{K}d(x, y),$$

for all $x, y \in C$, then T has a fixed point. Moreover, if $K > 1$, then T has a unique fixed point.

References

[1] T. Abdeljawad, E. Karapinar and K. Tas, *Common fixed point theorems in cone banach spaces*, Hacet. J. Math. Stat. **40** (2) (2011), 211–217.

[2] S. Alizadeh, F. Moradlou and Peyman Salimi, *Some fixed point results for (α, β)-(ψ, ϕ)-contractive mappings*, Filomat **28** (2014), 635–647.

[3] A. S. Cvetkovic, M. P. Stanic, S. Dimitrijevic and S. Simic, *Common Fixed Point Theorems for Four Mappings on Cone Metric Type Space*, Fixed point Theory Appl. 15 pages (2011) Article ID 589725.

[4] K. Deimling, *Nonlinear Functional Analysis,* Springer-Verlag, (1985).

[5] E. Karapinar, *Fixed point theorems in cone Banach spaces*, Fixed Point Theory and Appl. 9 pages (2009) Article ID 60928.

[6] G. Jungck. and B. E. Rhoades, *Fixed points for set valued functions without continuity*, Indian J. Pure and Appl. Math. **29** (3) (1998), 227–238.

[7] M. S. Khan and M. A. Imdad, *A common fixed point theorem for a class of mappings*, Indian J. Pure Appl. Math. **14** (1983), 1220–1227.

[8] L. G. Huang and X. Zhang, *Cone metric spaces and fixed point theorems of contractive mappings*, J. Math. Anal. Appl. **332** (2007), 1468–1476.

[9] S. D. Lin, *A common fixed point theorem in abstract spaces*, Indian J. Pure Appl. Math. **18** (8) (1987) 685–690.

[10] M. G. Maia, *Un'Osservazione sulle contrazioni metriche*, Ren. Sem. Mat. Univ. Padova **40** (1968), 139–143.

[11] N. Hussain and M. H. Shah, *KKM maps in cone b-metric spaces*, Comput. Math. Appl. **73** (2010) 3123–3129.

[12] F. Moradlou, P. Salimi and P. Vetro, *Some new extensions of Edelstein-Suzuki-type fixed point theorem to G-metric and G-cone metric spaces,* Acta Math. Sci. Ser. B Engl. Ed. 2013, **33** B (4):1049–1058.

[13] F. Moradlou, P. Salimi and P. Vetro, *Fixed point results for r-(η, ξ, ψ)-contractive mappings of type (I), (II) and (III),* Filomat **27** (2013), 403–410.

[14] Sh. Rezapour and R. Hamlbarani, *Some notes on the paper "one metric spaces and fixed point theorems of contractive mappings"*, J. Math. Anal. Appl. **347** (2008), 719–724.

[15] B. Rzepecki, *On fixed point theorems of Maia type*, Publications De Linstitut Mathématique **28** (1980), 179-186.

[16] D. Turkoglu and M. Abuloha, *Cone metric spaces and fixed point theorems in diametrically contractive mappings*, Acta. Math. Sin. **26** (3) (2010), 489–496.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 607 – 614

Non-selfadjoint matrix Sturm-Liouville operators with eigenvalue-dependent boundary conditions

Murat OLGUN *

Abstract

In this paper we investigate discrete spectrum of the non-selfadjoint matrix Sturm-Liouville operator L generated in $L^2(\mathbb{R}_+, S)$ by the differential expression

$$\ell(y) = -y'' + Q(x)y \quad , \qquad x \in \mathbb{R}_+ : [0, \infty),$$

and the boundary condition $y\prime(0) - \left(\beta_0 + \beta_1\lambda + \beta_2\lambda^2\right) y(0) = 0$ where Q is a non-selfadjoint matrix valued function. Also using the uniqueness theorem of analytic functions we prove that L has a finite number of eigenvalues and spectral singularities with finite multiplicities.

2000 AMS Classification: 34B24, 34L40 , 47A10.

Keywords: Eigenvalues, Spectral Singularities, Spectral Analysis, Sturm-Liouville Operator, Non-Selfadjoint Matrix Operator.

Received 07/01/2014 : Accepted 04/03/2014 Doi : 10.15672/HJMS.2015449424

1. Introduction

The study of the spectral analysis of non self-adjoint Sturm-Liouville operators was begun by Naimark [23] in 1954. He studied the spectral analysis of non-selfadjoint differential operators with continuous and discrete spectrum. Also he investigated the existence of spectral singularities in the continuous spectrum of the non-selfadjoint differential operator. Spectral singularities are poles of the resolvent's kernel which are in the continuous spectrum and are not eigen-values [26]. General notion of the sets of spectral singularities for closed linear operators on a Banach space was given by Nagy in [22]. Let L_0 denote the operator generated in $L^2(\mathbb{R}_+)$ by the differential expression

(1.1) $$\ell_0(y) = -y'' + v(x)y \quad , \qquad x \in \mathbb{R}_+$$

and the boundary condition

$$y\prime(0) - hy(0) = 0$$

where v is a complex valued function and $h \in \mathbb{C}$.

*Ankara University, Faculty of Sciences Department of Mathematics Ankara, TURKEY
Email: olgun@ankara.edu.tr

In [23] it is shown that if

$$\int_0^\infty \exp(\varepsilon x)\,|v(x)|\,dx < \infty,$$

for some $\varepsilon > 0$, then L_0 has a finite number of eigenvalues and spectral singularities with a finite multiplicities. Pavlov [25] established the dependence of the structure of the spectral singularities of L_0 on the behavior of the potential function at infinity. The spectral analysis of the non-selfadjoint operator, generated in $L^2(\mathbb{R}_+)$ by (1.1) and the integral boundary condition

$$\int_0^\infty B(x)\,y(x)\,dx + \alpha y'(0) - \beta y(0) = 0$$

where $B \in L^2(\mathbb{R}_+)$ is a complex-valued function, and $\alpha, \beta \in \mathbb{C}$, was investigated in detail by Krall [15],[16].

Some problems of spectral theory of differential and some other types of operators with spectral singularities were also studied in [1],[3]-[7],[17],[18]. The spectral analysis of the non self-adjoint operator, generated in $L^2(\mathbb{R}_+)$ by (1.1) and the boundary condition

$$\frac{y\prime(0)}{y(0)} = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2$$

where $\alpha_i \in \mathbb{C}$, $i = 0,\ 1,\ 2$ with $\alpha_2 \neq 0$ was investigated by Bairamov et al. [8].

The all above mentioned papers related with differential and difference operators are of scalar coefficients.Spectral analysis of the selfadjoint differential and difference operators with matrix coefficients are studied in [2],[9]-[11],[14].

Let S be a n-dimensional $(n < \infty)$ Euclidian space. We denote by $L^2(\mathbb{R}_+, S)$ the Hilbert space of vector-valued functions with values in S and the norm

$$\|f\|^2_{L_2(\mathbb{R}_+,S)} = \int_0^\infty \|f(x)\|^2_S\,dx.$$

Let L denote the operator generated in $L^2(\mathbb{R}_+, S)$ by the matrix differential expression

$$\ell(y) = -y'' + Q(x)\,y, \qquad x \in \mathbb{R}_+$$

and the boundary condition $y(0) = 0$, where Q is a non-selfadjoint matrix-valued function (i.e. $Q \neq Q^*$). In [24], [12] discrete spectrum of the non-selfadjoint matrix Sturm-Liouville operator was investigated. Let us consider the BVP in $L_2(\mathbb{R}_+, S)$

$$(1.2) \qquad -y'' + Q(x)y = \lambda^2 y,\ x\epsilon\mathbb{R}_+,$$

$$(1.3) \qquad y\prime(0) - \left(\beta_0 + \beta_1\lambda + \beta_2\lambda^2\right)y(0) = 0$$

where Q is a non self-adjoint matrix-valued function and β_0, β_1, β_2 are non self-adjoint matrices with $\det\beta_2 \neq 0$.

In this paper using the uniqueness theorem of analytic functions we investigate the eigenvalues and the spectral singularities of L. In particular we prove that L has a finite number of eigenvalues and spectral singularities with finite multiplicities, if the condition

$$\lim_{x\to\infty} Q(x) = 0\ ,\ \int_0^\infty e^{\epsilon x}\left\|Q'(x)\right\|dx < \infty,\ \epsilon > 0,$$

holds, where $\|.\|$ denote norm in S. We also show that the analogue of the Pavlov condition for L is the form

$$\lim_{x\to\infty} Q(x) = 0 \ , \ \int_0^\infty e^{\epsilon\sqrt{x}} \left\|Q'(x)\right\| dx < \infty, \ \epsilon > 0.$$

2. Jost Solution

Let us consider the matrix Sturm-Liouville equation

$$-y'' + Q\left(x\right)y = \lambda^2 y, \ x \in \mathbb{R}_+ \tag{2.1}$$

where Q is a non-selfadjoint matrix-valued function and

$$\int_0^\infty x\left\|Q\left(x\right)\right\| dx < \infty \tag{2.2}$$

holds. The bounded matrix solution of (2.1) satisfying the condition

$$\lim_{x\to\infty} y\left(x,\lambda\right)e^{-i\lambda x} = I, \ \lambda \in \overline{\mathbb{C}}_+ := \{\lambda : \lambda \in \mathbb{C}, \ \operatorname{Im}\lambda \geq 0\}$$

will be denoted by $F\left(x,\lambda\right)$. The solution $F\left(x,\lambda\right)$ is called Jost solution of (2.1). It has been shown that, under the condition (2.2), the Jost solution has the representation

$$F\left(x,\lambda\right) = e^{i\lambda x}I + \int_x^\infty K\left(x,t\right)e^{i\lambda t}dt \tag{2.3}$$

where I denotes the identity matrix in S and the matrix function $K\left(x,t\right)$ satisfies

(2.4)

$$K(x,t) = \frac{1}{2}\int_{\frac{x+t}{2}}^{\infty} Q(s)ds + \frac{1}{2}\int_x^{\frac{x+t}{2}}\int_{t+x-s}^{t+s-x} Q(s)K(s,v)dvds \quad + \frac{1}{2}\int_{\frac{x+t}{2}}^{\infty}\int_s^{t+s-x} Q(s)K(s,v)dvds.$$

$K\left(x,t\right)$ is continuously differentiable with respect to their arguments and

$$\left\|K\left(x,t\right)\right\| \leq c\alpha\left(\frac{x+t}{2}\right) \tag{2.5}$$

$$\left\|K_x\left(x,t\right)\right\| \leq \frac{1}{4}\left\|Q\left(\frac{x+t}{2}\right)\right\| + c\alpha\left(\frac{x+t}{2}\right) \tag{2.6}$$

$$\left\|K_t\left(x,t\right)\right\| \leq \frac{1}{4}\left\|Q\left(\frac{x+t}{2}\right)\right\| + c\alpha\left(\frac{x+t}{2}\right) \tag{2.7}$$

where $\alpha\left(x\right) = \int_x^\infty \left\|Q\left(s\right)\right\| ds$ and $c > 0$ is a constant. Therefore, $F\left(x,\lambda\right)$ is analytic with respect to λ in $\mathbb{C}_+ := \{\lambda : \lambda \in \mathbb{C}_+, \operatorname{Im}\lambda > 0\}$ and continuous on the real axis $\left([2], [17], [19]\right)$.

We will denote the matrix solution of (2.1) satisfying the initial conditions

$$G\left(0,\lambda\right) = I \quad , \quad G'\left(0,\lambda\right) = \beta_0 + \beta_1\lambda + \beta_2\lambda^2$$

by $G\left(x,\lambda\right)$. Let us define the following functions:

$$A_{\pm}(\lambda) = F_x\left(0,\pm\lambda\right) - (\beta_0 + \beta_1\lambda + \beta_2\lambda^2)F(0,\pm\lambda) \qquad \lambda\epsilon\bar{\mathbb{C}}_{\pm}, \tag{2.8}$$

where $\bar{\mathbb{C}}_{\pm} = \{\lambda : \lambda \in \mathbb{C}, \ \pm \operatorname{Im} \lambda \geq 0\}$. It is obvious that the functions $A_{+}(\lambda)$ and $A_{-}(\lambda)$ are analytic in $\mathbb{C}_{+}$ and $\mathbb{C}_{-}$, respectively and continuous on the real axis.It is clear that the resolvent of L defined by the following

$$\mathbf{R}_{\lambda}(L)\varphi = \int_{0}^{\infty} R(x,\xi;\lambda)\varphi(\xi)\,d\xi, \qquad \varphi \in L^{2}(\mathbb{R}_{+},S) \tag{2.9}$$

where

$$R(x,\xi;\lambda) = \begin{cases} R_{+}(x,\xi;\lambda) & , \quad \lambda \in \mathbb{C}_{+} \\ R_{-}(x,\xi;\lambda) & , \quad \lambda \in \mathbb{C}_{-} \end{cases}$$

$$R_{\pm}(x,\xi;\lambda) = \begin{cases} -F(x,\pm\lambda)\,A_{\pm}^{-1}(\lambda)\,G^{t}(\xi,\lambda), & 0 \leq \xi \leq x \\ -G(x,\lambda)\left[A_{\pm}^{t}(\lambda)\right]^{-1} F(\xi,\pm\lambda), & x \leq \xi < \infty, \end{cases} \tag{2.10}$$

and $G^{t}(\xi,\lambda)$ and $A_{\pm}^{t}(\lambda)$ denotes the transpose of the matrix function $G(\xi,\lambda)$ and $A_{\pm}(\lambda)$ respectively.

In the following we will denote the class of non self-adjoint matrix-valued valued absolutely continuous functions in $\mathbb{R}_{+}$ by $AC(\mathbb{R}_{+})$.

2.1. Lemma. *If*

$$Q \epsilon AC(\mathbb{R}_{+}) \ , \quad \lim_{x\to\infty} Q(x) = 0 \quad , \quad \int_{0}^{\infty} x^{3}\left\|Q'(x)\right\| < \infty \tag{2.11}$$

then $K_{tt}(x,t)$ exist and

$$\begin{aligned} K_{tt}(x,t) = &-\frac{1}{8}Q'(\frac{t}{2}) + \frac{1}{2}\int_{0}^{\infty} Q(s)K_{t}(s,t+s)ds \\ &- \frac{1}{4}Q(\frac{t}{2})K(\frac{t}{2},\frac{t}{2}) \\ &- \frac{1}{2}\int_{0}^{\frac{t}{2}} Q(s)\left[K_{t}(s,t-s) + K_{t}\left(t-x+s\right)\right]ds. \end{aligned} \tag{2.12}$$

Proof. The proof of lemma direct consequently of (2.4) . ∎

From (2.5)-(2.7) and (2.12) we obtain that

$$\|K_{tt}(0,t)\| \leq c\left\{\left\|Q'(\frac{t}{2})\right\| + t\left\|Q(\frac{t}{2})\right\| + t\alpha(\frac{t}{2}) + \alpha_{1}(\frac{t}{2})\right\} \tag{2.13}$$

holds, where $\alpha_{1}(t) = \int_{t}^{\infty} \alpha(s)\,ds$ and $c > 0$ is a constant.

2.2. Lemma. *Under the condition* (2.11)*, A_{+} and A_{-} have the representations*

$$A_{+}(\lambda) = -\beta_{2}\lambda^{2} + A\lambda + B + \int_{0}^{\infty} F^{+}(t)e^{i\lambda t}dt, \quad \lambda \in \bar{\mathbb{C}}_{+}, \tag{2.14}$$

$$A_{-}(\lambda) = -\beta_{2}\lambda^{2} + C\lambda + D + \int_{0}^{\infty} F^{-}(t)e^{-i\lambda t}dt, \quad \lambda \in \bar{\mathbb{C}}_{-}, \tag{2.15}$$

where A, B, C, D are non self-adjoint matrices in S, and $F^{\pm} \in L_{1}(\mathbb{R}_{+})$.

Proof. Using (2.3), (2.4) and (2.8) we get (2.14), where

$$A = i - \beta_1 - i\beta_2 K(0,0),$$
$$(2.16) \qquad B = -K(0,0) - \beta_0 - i\beta_1 K(0,0) + \beta_2 K_t(0,0),$$
$$F^+(t) = K_x(0,t) - \beta_0 K(0,t) - i\beta_1 K_t(0,t) + \beta_2 K_{tt}(0,0).$$

From (2.5) − (2.7) and (2.13), $F^+ \in L_1(\mathbb{R}_+)$. By similar way we obtain (2.15) and $F^- \in L_1(\mathbb{R}_+)$. ∎

2.3. Theorem. *$A_+(\lambda)$ and $A_-(\lambda)$ have the asymptotic behavior:*

$$(2.17) \quad A_\pm(\lambda) = -\beta_2\lambda^2 + A\lambda + B + o(1) \qquad \lambda \in \bar{\mathbb{C}}_\pm,\ |\lambda| \to \infty.$$

Proof. The proof is obvious from (2.5) − (2.7) and (2.13)). ∎

We will denote the continuous spectrum of L by σ_c. From Theorem 2 ([22], page 303) we get that

$$(2.18) \quad \sigma_c = \mathbb{R}.$$

3. Eigenvalues and Spectral Singularities of L

Let us suppose that

$$(3.1) \quad f_\pm(\lambda) := \det A_\pm(\lambda).$$

We denote the set of eigenvalues and spectral singularities of L by $\sigma_d(L)$ and $\sigma_{ss}(L)$, respectively. By the definition of eigenvalues and spectral singularities of differential operators we can write

$$(3.2) \quad \sigma_d(L) = \{\lambda{:}\lambda \in \mathbb{C}_+,\ f_+(\lambda) = 0\} \cup \{\lambda{:}\lambda \in \mathbb{C},\ f_-(\lambda) = 0\}$$

$$(3.3) \quad \sigma_{ss}(L) = \{\lambda : \lambda \in \mathbb{R}\backslash\{0\},\ f_+(\lambda) = 0\} \cup \{\lambda : \lambda \in \mathbb{R}\backslash\{0\},\ f_-(\lambda) = 0\}$$

[22], [23], [26]. It is clear that $\sigma_{ss}(L) \subset \mathbb{R}$.

3.1. Definition. The multiplicity of a zero of f_+ in $\overline{\mathbb{C}}_+$ (or f_- in $\overline{\mathbb{C}}_-$) is defined as the multiplicity of the corresponding eigenvalue and spectral singularity of L.

In order to investigate the quantitative properties of the eigenvalues and the spectral singularities of L, we need to discuss the quantitative properties of the zeros of f_+ and f_- in $\bar{\mathbb{C}}_+$ and $\bar{\mathbb{C}}_-$, respectively. Assume that

$$M_1^\pm = \{\lambda : \lambda \in \mathbb{C}_\pm,\ f_\pm(\lambda) = 0\}$$

and

$$M_2^\pm = \{\lambda : \lambda \in \mathbb{R},\ f_\pm(\lambda) = 0\}.$$

From (3.3) and (3.4), we get

$$(3.4) \quad \sigma_d(L) = M_1^+ \cup M_1^-,$$

and

$$(3.5) \quad \sigma_{ss}(L) = M_2^+ \cup M_2^- - \{0\}.$$

3.2. Theorem. *Under the condition* (2.11)

i) *The set $\sigma_d(L)$ is bounded and has at most countable number of elements and its limit points can lie only in a bounded subinterval of the real axis.*
ii) *The set $\sigma_{ss}(L)$ is bounded and $\mu(\sigma_{ss}(L)) = 0$, where $\mu(\sigma_{ss}(L))$ denotes the linear Lebesque measure of $\sigma_{ss}(L)$.*

Proof. Using (2.5) and (3.1) we get that the function $f_{\pm}$ is analytic in $\mathbb{C}_+$ continuous on the real axis and

$$f_{\pm}(\lambda) = -\lambda^2 \det \beta_2 + O(\lambda), \ \lambda \in \overline{\mathbb{C}}_{\pm}, |\lambda| \to \infty, \tag{3.6}$$

Equation (3.6) shows the boundedness of the sets $\sigma_d(L)$ and $\sigma_{ss}(L)$. From the analyticity of the function $f_{\pm}$ in $\mathbb{C}_{\pm}$ we obtain that $\sigma_d(L)$ has at most countable number of elements and its limit points can lie only in a bounded subinterval of the real axis . By the boundary value uniqueness theorem of analytic functions, we find that $\mu(\sigma_{ss}(L)) = 0$, [13]. ∎

We will denote the sets of limit points of M_1^+ and M_2^+ by M_3^+ and M_4^+ respectively and the set of all zeros of A_+ with infinite multiplicity in $\bar{\mathbb{C}}_+$ by M_5^+. Analogously define the sets M_3^-, M_4^- and M_5^-.

It is explicit from the boundary uniqueness theorem of analytic functions that [13]

$$\begin{aligned} &M_1^{\pm} \cap M_5^{\pm} = \varnothing, \quad M_3^{\pm} \subset M_2^{\pm}, \quad M_4^{\pm} \subset M_2^{\pm}, \\ &M_5^{\pm} \subset M_2^{\pm}, \quad M_3^{\pm} \subset M_5^{\pm}, \quad M_4^{\pm} \subset M_5^{\pm} \end{aligned} \tag{3.7}$$

and $\mu(M_3^{\pm}) = \mu(M_4^{\pm}) = \mu(M_5^{\pm}) = 0$.

3.3. Theorem. *If*

$$Q \epsilon AC(\mathbb{R}_+) \ , \ \lim_{x\to\infty} Q(x) = 0 \ , \ \int_0^{\infty} e^{\epsilon x} \left\| Q'(x) \right\| dx < \infty, \quad \epsilon > 0 \tag{3.8}$$

the operator L has a finite number of eigenvalues and spectral singularities and each of them is of finite multiplicity.

Proof. By (2.5) , (2.13) , (2.14) and (3.8) we observe that, the function A_+ has an analytic continuation to the half plane $\operatorname{Im} \lambda > -\frac{\varepsilon}{4}$. So, the limit points of zeros of A_+ in $\overline{\mathbb{C}}_+$ can not lie in $\mathbb{R}$. From analyticity of A_+ for $\operatorname{Im} \lambda > -\frac{\varepsilon}{4}$, we obtain that all zeros of A_+ in $\overline{\mathbb{C}}_+$ have a finite multiplicity. We obtain similar results for A_-. Consequently by (3.4) and (3.5) the sets $\sigma_d(L)$ and $\sigma_{ss}(L)$ have a finite number of elements with a finite multiplicity. ∎

Now let us suppose that hold, the conditions which is weaker than (3.8).

3.4. Theorem. *If*

$$Q \epsilon AC(\mathbb{R}_+) \ , \ \lim_{x\to\infty} Q(x) = 0 \ , \sup_{x\in\mathbb{R}_+} \left[\exp\left(\varepsilon\sqrt{x}\right) \left\| Q'(x) \right\| \right] < \infty, \ \varepsilon > 0 \tag{3.9}$$

holds , then $M_5^+ = M_5^- = \phi$.

Proof. From (3.1) and (3.9) we have f_+ is analytic in $\mathbb{C}_+$ and all of its derivatives are continuous on the $\overline{\mathbb{C}}_+$.For sufficiently large $P > 0$ we have

$$\left| \frac{d^m}{d\lambda^m} f_+(\lambda) \right| \le T_m, \quad m = 0, 1, 2, ..., \lambda \in \bar{\mathbb{C}}_+, |\lambda| < P \tag{3.10}$$

where

$$T_m := 2^m c \int_0^{\infty} t^m e^{-(\varepsilon/2)\sqrt{t}} dt, \ m = 0, 1, 2, ..., \tag{3.11}$$

where $c > 0$ is a constant. Since the function f_+ is not equal to zero identically, using Pavlov's Theorem [25] we get that M_5^+ satisfies

$$\int_0^{a} \ln G(s)\, d\mu\left(M_5^+, s\right) > -\infty \tag{3.12}$$

where $G(s) = \inf\limits_{m} \frac{T_m s^m}{m!}$, $\mu\left(M_5^+, s\right)$ is the linear Lebesque measure of s-neighborhood of M_5^+ and $a > 0$ is a constant .

We obtain the following estimates for T_m

(3.13) $\quad T_m \leq B b^m m! m^m$

where B and b are constants depending on c and ε. Substituting (3.13) in the definition of $G(s)$, we arrive at

$$G(s) = \inf_{m} \frac{T_m s^m}{m!} \leq B \exp\left(-e^{-1} b^{-1} s^{-1}\right).$$

Now by (3.12), we get

(3.14) $\quad \int\limits_0^a s^{-1} d\mu\left(M_5^+, s\right) < \infty.$

Consequently (3.14) holds for an arbitrary s if and only if $\mu\left(M_5^+, s\right) = 0$ or $M_5^+ = \phi$. In a similar way we can show $M_5^- = \phi$ ∎

3.5. Theorem. *Under the condition* (3.9) *the operator L has a finite number of eigenvalues and spectral singularities and each of them is of a finite multiplicity.*

Proof. We have to show that the functions f_+ and f_- have a finite number of zeros with a finite multiplicities in $\overline{\mathbb{C}}_+$ and $\overline{\mathbb{C}}_-$, respectively. We prove only for f_+.

It follows from (3.7) and Theorem 3.4 that $M_3^+ = M_4^+ = \phi$. So the bounded set M_1^+ and M_1^+ have no limit points, i.e. the function f_+ has only finite number of zeros in $\overline{\mathbb{C}}_+$. Since $M_5^+ = \phi$, these zeros are of finite multiplicity. ∎

References

[1] Adivar, M. and Bairamov, E. Difference equations of second order with spectral singularities, *J. Math. Anal. Appl.* **277** (2003), 714-721.

[2] Agranovich, Z. S. and Marchenko, V. A. The Inverse Problem of Scattering Theory, Gordon and Breach, 1965.

[3] Bairamov, E., Cakar, O. and Celebi, A. O. Quadratic pencil of Schrödinger operators with spectral singularities: Discrete spectrum and principal functions, *J. Math. Anal. Appl.* **216** (1997), 303-320.

[4] Bairamov, E., Cakar, O. and Krall, A. M. An eigenfunction expansion for a quadratic pencil of a Schrödinger operator with spectral singularities, *J. Differential Equations* **151** (1999), 268-289.

[5] Bairamov, E., Cakar, O. and Krall, A. M. Non-selfadjoint difference operators and Jacobi matrices with spectral singularities, *Math. Nachr.* **229** (2001) 5-14.

[6] Bairamov, E. and Celebi, A. O. Spectrum and spectral expansion for the non-selfadjoint discrete Dirac operators, *Quart. J. Math. Oxford* Ser. (2) **50** (1999), 371-384.

[7] Bairamov, E. and Coskun, C. The Structure of the Spectrum a System of Difference Equations, *Appl. Math. Lett.* **18** (2005), 384-394.

[8] Bairamov, E. and Yokus, N. Spectral Singularities of Sturm-Liouville Problems with Eigenvalue-Dependent Boundary Conditions, Abstract and Applied Analysis, **2009**, (2009) 1-8.

[9] Carlson, R. An inverse problem for the matrix Schrödinger equation, *J. Math. Anal. Appl.* **267** (2002), 564-575.

[10] Clark, S. and Gesztesy, F. Weyl-Titchmarsh M-function asymptotics, local uniqueness results, trace formulas and Borg-type theorems for Dirac operators, *Trans Amer. Math. Soc.* **354** (2002) 3475-3534.

[11] Clark, S., Gesztesy, F. and Renger, W. Trace formulas and Borg-type theorems for matrix-valued Jacobi and Dirac finite difference operators, *J. Differential Equations* **219** (2005), 144-182.

[12] Coskun, C. and Olgun, M. Principal functions of non-selfadjoint matrix Sturm-Liouville equations, Journal of Computational and Applied Mathematics, **235** (2011).

[13] Dolzhenko, E. P. Boundary value uniqueness theorems for analytic functions, Math. Notes **25** (1979), 437-442.

[14] Gesztesy, F., Kiselev, A. and Makarov, K. A. Uniqueness results for matrix-valued Schrödinger, Jacobi and Dirac-type operators, *Math. Nachr.* **239** (2002), 103-145.

[15] Krall, A. M. The adjoint of a differential operator with integral boundary conditions, *Proc. Amer. Math. Soc.***16** (1965), 738-742.

[16] Krall, A. M. A nonhomogeneous eigenfunction expansion, *Trans. Amer. Math. Soc*, **117** (1965), 352-361.

[17] Krall, A. M., Bairamov, E. and Cakar, O. Spectrum and spectral singularities of a quadratic pencil of a Schrödinger operator with a general boundary condition, *J. Differentional Equations*, **151** (1999), 252-267.

[18] Krall, A. M., Bairamov, E. and Cakar, O. Spectral analysis of a non-selfadjoint discrete Schrödinger operators with spectral singularities, *Math. Nachr.* **231** (2001), 89-104.

[19] Levitan, B. M. Inverse Sturm-Liouville Problems, VSP, Zeist, (1987).

[20] Lyance, V. E. A differential operator with spectral singularities, I, II, *AMS Translations* 2 (60) 1967, 185-225, 227-283.

[21] Marchenko, V. A. Sturm-Liouville Operators and Applications, Birkhauser Verlag, Basel, (1986).

[22] Nagy, B. Operators with spectral singularities. J. Operator Theory 15, no **2**, (1986), 307-325.

[23] Naimark, M. A. Investigation of the spectrum and the expansion in eigenfunctions of a non-selfadjoint operators of second order on a semi-axis, *AMS Translations* 2(16) 1960, 103-193.

[24] Olgun, M. and Coskun, C. Non-selfadjoint matrix Sturm-Liouville operators with spectral singularities, Applied Mathematics and Computations, **216** (8) (2010).

[25] Pavlov, B. S. On a non-selfadjoint Schrödinger operator. II. (Russian) 1967 Problems of Mathematical Physics, No. 2, Spectral Theory, Diffraction Problems (Russian) pp. 133–157 Izdat. Leningrad. Univ., Leningrad.

[26] Schwartz, J. T. Some non-selfadjoint operators, Comm. Pure and Appl. Math. 13 (1960), 609-639.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 615 – 621

The quenching behavior of a nonlinear parabolic equation with a singular boundary condition

Nuri OZALP * and Burhan SELCUK †

Abstract

In this paper, we study the quenching behavior of solution of a nonlinear parabolic equation with a singular boundary condition. We prove finite-time quenching for the solution. Further, we show that quenching occurs on the boundary under certain conditions. Furthermore, we show that the time derivative blows up at quenching point. Also, we get a lower solution and an upper bound for quenching time. Finally, we get a quenching rate and lower bounds for quenching time.

2000 AMS Classification: 35K55, 35K60, 35B35, 35Q60

Keywords: Nonlinear parabolic equation, singular boundary condition, quenching, quenching point, quenching time, maximum principles.

Received 07/07/2013 : Accepted 20/03/2014 Doi : 10.15672/HJMS.2015449429

1. Introduction

In this paper, we study the quenching behavior of solutions of the following nonlinear parabolic equation with a singular boundary condition:

$$(1.1)\qquad \begin{cases} u_t = u_{xx} + (1-u)^{-p},\ 0 < x < 1,\ 0 < t < T, \\ u_x(0,t) = 0,\ u_x(1,t) = (1-u(1,t))^{-q},\ 0 < t < T, \\ u(x,0) = u_0(x),\ 0 \le x \le 1, \end{cases}$$

where p, q are positive constants and $T \le \infty$. The initial function $u_0 : [0,1] \to (0,1)$ satisfies the compatibility conditions

$$u_0'(0) = 0,\ u_0'(1) = (1 - u_0(1))^{-q}.$$

Throughout this paper, we also assume that the initial function u_0 satisfies the inequalities

*Department of Mathematics, Ankara University, Besevler, 06100, Turkey
Email: nozalp@science.ankara.edu.tr
†Department of Computer Engineering, Karabuk University, Balıklarkayası Mevkii, 78050, Turkey
Email:bselcuk@karabuk.edu.tr

$$(1.2) \qquad u_{xx}(x,0) + (1-u(x,0))^{-p} \geq 0,$$

$$(1.3) \qquad u_x(x,0) \geq 0$$

Our main purpose is to examine the quenching behavior of the solutions of problem (1.1) having two singular heat sources. A solution $u(x,t)$ of the problem (1.1) is said to quench if there exists a finite time T such that

$$\lim_{t\to T^-} \max\{u(x,t) : 0 \leq x \leq 1\} \to 1.$$

From now on, we denote the quenching time of the problem (1) with T.

The concept of quenching was first introduced by Kawarada. In [12], Kawarada has considered an initial-boundary value problem for the parabolic equation $u_t = u_{xx} + 1/(1-u)$. Then, the quenching problems have been studied extensively by several researchers (cf. the surveys by Chan [1, 2] and Kirk and.Roberts [14] and [3], [4], [6], [8], [9],[10], [13], [15], [16], [17], [19]). In the literature, the quenching problems have been less studied with two nonlinear heat sources. We give as examples two of these papers. Chan and Yuen [5] considered the problem

$$\begin{aligned}
&u_t = u_{xx}, \text{ in } \Omega,\\
&u_x(0,t) = (1-u(0,t))^{-p}, \ u_x(a,t) = (1-u(a,t))^{-q}, \ 0 < t < T,\\
&u(x,0) = u_0(x), \ 0 \leq u_0(x) < 1, \text{ in } \bar{D},
\end{aligned}$$

where $a, p, q > 0, T \leq \infty, D = (0,a), \Omega = D \times (0,T)$. They showed that $x = a$ is the unique quenching point in finite time if u_0 is a lower solution, and u_t blows up at quenching. Further, they obtained criteria for nonquenching and quenching by using the positive steady states. Zhi and Mu [20] considered the problem

$$\begin{aligned}
&u_t = u_{xx} + (1-u)^{-p}, \ 0 < x < 1, \ 0 < t < T\\
&u_x(0,t) = u^{-q}(0,t), \ u_x(1,t) = 0, \ 0 < t < T,\\
&u(x,0) = u_0(x), \ 0 < u_0(x) < 1, \ 0 \leq x \leq 1,
\end{aligned}$$

where $p, q > 0$ and $T \leq \infty$. They showed that $x = 0$ is the unique quenching point in finite time if u_0 satisfies $u_0''(x) + (1-u_0(x))^{-p} \leq 0$ and $u_0'(x) \geq 0$. Further, they obtained the quenching rate estimates which is $(T-t)^{1/2(q+1)}$ if T denotes the quenching time. Further, the quenching problems have been less studied with combined power-type nonlinearities ([7], [18]) in the literature. In [18], Xu et.al. the studied the following quenching behavior for the solutions of parabolic equation with combined power-type nonlinearities:

$$\begin{aligned}
&u_t - \Delta u = \sum_{k=2}^{q} (b-u(x,t))^{-k}, \text{ in } \Omega \times (0,T),\\
&u(x,t) = 0, \text{ on } \partial\Omega \times (0,T),\\
&u(x,0) = u_0(x), \text{ in } \Omega,
\end{aligned}$$

where Ω is a bounded domain in R^N with smooth boundary $\partial\Omega, q > 2, b = const > 0$. The initial data $u_0(x) \in C^1(\overline{\Omega})$ is nonnegative in Ω and $sup_{x\in\Omega} u_0(x) < b$. They showed that the solution of the above problem quenches in a finite time, and estimated its quenching time. Finally, they given numerical examples.

Here, we would like to study how the reaction term $(1-u)^{-p}$ and the boundary absorption term $(1-u)^{-q}$ affect the quenching behaviour of the solution of the problem (1.1). In Section 2, we first show that quenching occurs in finite time under the condition (1.2). Then, we show that the only quenching point is $x = 1$ under the condition (1.2) and (1.3). Further we show that u_t blows up at quenching time. In Section 3, we get a lower solution and an upper bound for quenching time. In Section 4, we get a quenching rate and lower bounds for quenching time.

2. Quenching on the boundary and blow-up of u_t

2.1. Remark. We assume that the condition (1.2) and (1.3) is proper. Namely, we can easily construct such a initial function satisfying (1.2),(1.3) and compatibility conditions. Let $0 < A < 1, \alpha = \frac{1}{A(1-A)}$ and $u(x,0) = Ax^{\alpha}$. For example, for $q = 1$ and $A = 0.5$, $u(x,0) = \frac{1}{2}x^4$ satisfies compatibility conditions, (1.2) and (1.3).

2.2. Remark. If u_0 satisfies (1.3), then we get $u_x > 0$ in $(0,1]\times(0,T)$ by the maximum principle. Thus we get $u(1,t) = \max\limits_{0\le x\le 1} u(x,t)$.

2.3. Lemma. If u_0 satisfies (1.2), then $u_t(x,t) \geq 0$ in $[0,1]\times[0,T)$.
Proof. Let us prove it by utilizing Lemma 3.1 of [11]. Let $v = u_t(x,t)$. Then, $v(x,t)$ satisfies

$$\begin{aligned}
&v_t = v_{xx} + p\,(1-u)^{-p-1}\,v,\ 0<x<1,\ 0<t<T,\\
&v_x\,(0,t) = 0,\ v_x\,(1,t) = q\,(1-u(1,t))^{-q-1}\,v(1,t),\ 0<t<T,\\
&v\,(x,0) = u_{xx}\,(x,0) + (1-u\,(x,0))^{-p} \geq 0,\ 0\leq x\leq 1.
\end{aligned}$$

For any fixed $\tau \in (0,T)$, let

$$\begin{aligned}
L &= \max_{0\le x\le 1,\ 0\le t\le \tau}\left(\frac{1}{2}q\,(1-u(x,t))^{-q-1}\right),\\
M &= 2L + 4L^2 + \max_{0\le x\le 1,\ 0\le t\le \tau}\left(p\,(1-u(x,t))^{-p-1}\right).
\end{aligned}$$

Set $w(x,t) = e^{-Mt-Lx^2}v(x,t)$. Then w satisfies

$$\begin{aligned}
&w_t = w_{xx} + 4Lxw_x + cw,\ 0<x<1,\ 0<t\leq\tau,\\
&w_x(0,t) = 0,\ w_x(1,t) = d(t)w(1,t),\ 0<t\leq\tau,\\
&w(x,0)\geq 0,\ 0\leq x\leq 1,
\end{aligned}$$

where

$$c = c(x,t) = 4L^2(x^2-1) + p\,(1-u(x,t))^{-p-1} - \max_{0\le x\le 1,\ 0\le t\le \tau}\left(p\,(1-u(x,t))^{-p-1}\right) \leq 0$$

and

$$d(t) = -\max_{0\le x\le 1,\ 0\le t\le \tau}\left(q\,(1-u(x,t))^{-q-1}\right) + q\,(1-u(1,t))^{-q-1} \leq 0.$$

By the maximum principle and Hopf lemma, we obtain that $w \geq 0$ in $[0,1]\times[0,\tau]$. Thus, $u_t(x,t)\geq 0$ in $[0,1]\times[0,T)$. □

2.4. Theorem. If u_0 satisfies (1.2), then there exist a finite time T, such that the solution u of the problem (1.1) quenches at time T.
Proof. Assume that u_0 satisfies (1.2). Then there exist

$$w = (1-u\,(1,0))^{-q} + \int_0^1 (1-u\,(x,0))^{-p}\,dx > 0.$$

Introduce a mass function; $m\,(t) = \int_0^1 (1-u\,(x,t))\,dx, 0<t<T$. Then

$$m'\,(t) = -\,(1-u\,(1,t))^{-q} - \int_0^1 (1-u\,(x,t))^{-p}\,dx \leq -w,$$

by Lemma 2.3. Thus, $m\,(t) \leq m(0) - wt$; which means that $m\,(T_0) = 0$ for some $T_0, (0 < T \leq T_0)$. Then u quenches in finite time. □

2.5. Theorem. If u_0 satisfies (1.2) and (1.3), then $x=1$ is the only quenching point.
Proof. Define

$$J(x,t) = u_x - \varepsilon\,(x-(1-\eta)) \text{ in } [1-\eta,1]\times[\tau,T),$$

where $\eta\in(0,1), \tau\in(0,T)$ and ε is a positive constant to be specified later. Then, $J(x,t)$ satisfies

$$J_t - J_{xx} = p(1-u)^{-p-1}u_x > 0 \text{ in } (1-\eta,1)\times(\tau,T),$$

since $u_x(x,t)>0$ in $(0,1]\times(0,T)$. Thus, $J(x,t)$ cannot attain a negative interior minimum by the maximum principle. Further, if ε is small enough, $J(x,\tau)>0$ since $u_x(x,t)>0$ in $(0,1]\times(0,T)$. Furthermore, if ε is small enough,

$$\begin{aligned} J(1-\eta,t) &= u_x(1-\eta,t) > 0,\\ J(1,t) &= (1-u(1,t))^{-q} - \varepsilon\eta > 1-\varepsilon\eta > 0 \end{aligned}$$

for $t\in(\tau,T)$. By the maximum principle, we obtain that $J(x,t)>0$, i.e. $u_x > \varepsilon\,(x-(1-\eta))$ for $(x,t)\in[1-\eta,1]\times[\tau,T)$. Integrating this with respect to x from $(1-\eta)$ to 1, we have

$$u(1-\eta,t) < u(1,t) - \frac{\varepsilon\eta^2}{2} < 1 - \frac{\varepsilon\eta^2}{2}.$$

So u does not quench in $[0,1)$. The theorem is proved. □

2.6. Theorem. u_t blows up at the quenching point $x=1$.
Proof. We will prove that u_t blows up at quenching, as in [5]. Suppose that u_t is bounded on $[0,1]\times[0,T)$. Then, there exists a positive constant M such that $u_t<M$. We have $u_{xx}+(1-u)^{-p} < M \Rightarrow u_{xx} < M$. Integrating this twice with respect to x from x to 1, and then from 0 to 1, we have

$$\frac{1}{(1-u(1,t))^q} < \frac{M}{2} + u(1,t) - u(0,t).$$

As $t\to T^-$, the left-hand side tends to infinity, while the right-hand side is finite. This contradiction shows that u_t blows up somewhere. □

3. A lower solution and an upper bound for the quenching time

3.1. Definition. μ is called a lower solution of problem (1.1) if μ satisfies the following conditions:

$$\begin{aligned} &\mu_t - \mu_{xx} \le (1-\mu)^{-p},\ 0<x<1,\ 0<t<T,\\ &\mu_x\,(0,t)=0,\ \mu_x\,(1,t) \le (1-\mu(1,t))^{-q},\ 0<t<T,\\ &\mu\,(x,0) \le u_0\,(x),\ 0\le x\le 1. \end{aligned}$$

It is an upper solution when the inequalities are reversed.

3.2. Lemma. Let u be a solution and μ be a lower solution of problem (1.1) in $[0,1]\times[0,T)$. Then $u\ge\mu$ in $[0,1]\times[0,T)$.
Proof. Let $v(x,t)=u(x,t)-\mu(x,t)$. Then $v(x,t)$ satisfies

$$\begin{aligned} &v_t \ge v_{xx} + p\,(1-\eta)^{-p-1}\,v,\ 0<x<1,\ 0<t<T,\\ &v_x\,(0,t)=0,\ v_x\,(1,t) \ge q\,(1-\xi(1,t))^{-q-1}\,v(1,t),\ 0<t<T,\\ &v\,(x,0)\ge 0,\ 0\le x\le 1, \end{aligned}$$

where $\eta(x,t)$ lies between $u(x,t)$ and $\mu(x,t)$ and $\xi(1,t)$ lies between $u(1,t)$ and $\mu(1,t)$.

For any fixed $\tau \in (0, T)$, let

$$\begin{aligned} L &= \max_{0\le x\le 1,\ 0\le t\le\tau} \left(\frac{1}{2} q\,(1-\xi(x,t))^{-q-1}\right), \\ M &= 2L + 4L^2 + \max_{0\le x\le 1,\ 0\le t\le\tau} \left(p\,(1-\eta(x,t))^{-p-1}\right). \end{aligned}$$

Set $w(x,t) = e^{-Mt-Lx^2} v(x,t)$. Then w satisfies

$$\begin{aligned} &w_t \ge w_{xx} + 4Lxw_x + cw,\ 0 < x < 1,\ 0 < t \le \tau, \\ &w_x(0,t) = 0,\ w_x(1,t) \ge d(t)w(1,t),\ 0 < t \le \tau, \\ &w(x,0) \ge 0,\ 0 \le x \le 1, \end{aligned}$$

where $c = c(x,t) \le 0$ and $d = d(t) \le 0$. By the maximum principle, we obtain that $w \ge 0$ in $[0,1] \times [0,\tau]$. Thus, $u \ge \mu$ in $[0,1] \times [0,T)$. □

3.3. Theorem. $x = 1$ is a quenching point.
Proof. Let $\min_{x\in[0,1]} u_0(x) = c \ge 0$. Define

$$\mu(x,t) = 1 - \left(\frac{(q+1)\left(1-x^2+\tau-t\right)}{2}\right)^{1/(q+1)} \quad \text{in } [0,1]\times[0,\tau],$$

where $\tau = 2(1-c)^{q+1}/(q+1)$. We have

$$\begin{aligned} \mu_t - \mu_{xx} &= \frac{-1}{2}\left(\frac{(q+1)\left(1-x^2+\tau-t\right)}{2}\right)^{-q/(q+1)} \\ &\quad -x^2 q \left(\frac{(q+1)\left(1-x^2+\tau-t\right)}{2}\right)^{(-2q-1)/(q+1)} \\ &\le 0 \end{aligned}$$

for $x \in (0,1), t \in (0,\tau]$. Further,

$$\begin{aligned} \mu_x(0,t) &= 0, \\ \mu_x(1,t) &= (1-\mu(1,t))^{-q} \end{aligned}$$

for $t \in (0,\tau]$. Furthermore,

$$\mu(x,0) = 1 - \left(\frac{(q+1)\left(1-x^2+\tau\right)}{2}\right)^{1/(q+1)} \le 1 - \left(\frac{(q+1)\,\tau}{2}\right)^{1/(q+1)} = c,$$

for $x \in [0,1]$. Thus, $\mu(x,t)$ is a lower solution of the problem (1.1). In addition, at $t = \tau$ and $x = 1$, we get

$$\mu(1,\tau) = 1.$$

Hence, we have

$$u(1,\tau) \ge \mu(1,\tau) = 1$$

by Lemma 3.2. Thus, $x = 1$ is a quenching point. □

3.4. Remark. We can calculate an upper bound for the quenching time. From Theorem 3.3, maximum upper bound is $T = 2/(q+1)$ (for $c = 0$). Also, as in Remark 2.1, $u_0(x) = \frac{1}{2}x^4$ (for $q = 1$), then we have $T = 1$.

4. A quenching rate and lower bounds for the quenching time

In this section, we get a quenching rate and lower bounds for quenching time. Throughout this section, we assume that

(4.1) $\quad u_x(x,0) \geq x(1-u(x,0))^{-q}, 0<x<1,$

(4.2) $\quad u_t(1,t) = u_{xx}(1,t) + (1-u(1,t))^{-p}, 0<t<T.$

4.1. Theorem. If u_0 satisfies $(1.2),(1.3),(4.1)$ and (4.2), then there exists positive constants C_1 and C_2 such that

$$\begin{aligned} \text{if } p &> 2q+1, \text{ then } u(1,t) \geq 1 - C_1(T-t)^{1/(p+1)},\\ \text{if } q &\leq p \leq 2q+1, \text{ then } u(1,t) \geq 1 - C_2(T-t)^{1/(2q+2)}, \end{aligned}$$

for t sufficiently close to T.

Proof. Define

$$J(x,t) = u_x - x(1-u)^{-q} \text{ in } [0,1]\times[0,T).$$

Then, $J(x,t)$ satisfies

$$\begin{aligned} J_t - J_{xx} - p(1-u)^{-p-1}J &= 2q(1-u)^{-q-1}u_x + (p-q)x(1-u)^{-p-q-1}\\ &\quad + xq(q+1)(1-u)^{-q-2}u_x^2, \end{aligned}$$

since $u_x>0$ and $p\geq q$, $J(x,t)$ cannot attain a negative interior minimum. On the other hand, $J(x,0)\geq 0$ by (4.1) and

$$J(0,t)=0, J(1,t)=0,$$

for $t\in(0,T)$. By the maximum principle, we obtain that $J(x,t)\geq 0$ for $(x,t)\in[0,1]\times[0,T)$. Therefore

$$J_x(1,t) = \lim_{h\to 0^+}\frac{J(1,t)-J(1-h,t)}{h} = \lim_{h\to 0^+}\frac{-J(1-h,t)}{h} \leq 0.$$

From (4.2), we get

$$\begin{aligned} J_x(1,t) &= u_{xx}(1,t) - (1-u(1,t))^{-q} - q(1-u(1,t))^{-2q-1}\\ &= u_t(1,t) - (1-u(1,t))^{-p} - (1-u(1,t))^{-q} - q(1-u(1,t))^{-2q-1} \leq 0 \end{aligned}$$

and

$$\begin{aligned} \text{if } p &> 2q+1, \text{ then } u_t(1,t) \leq (q+2)(1-u(1,t))^{-p},\\ \text{if } q &\leq p \leq 2q+1, \text{ then } u_t(1,t) \leq (q+2)(1-u(1,t))^{-2q-1}. \end{aligned}$$

Integrating for t from t to T we get

$$\begin{aligned} \text{if } p &> 2q+1, \text{ then } u(1,t) \geq 1 - C_1(T-t)^{1/(p+1)},\\ \text{if } q &\leq p \leq 2q+1, \text{ then } u(1,t) \geq 1 - C_2(T-t)^{1/(2q+2)}, \end{aligned}$$

where $C_1 = [(q+2)(p+1)]^{1/(p+1)}$ and $C_2 = [(q+2)(2q+2)]^{1/(2q+2)}$. □

4.2. Remark. We can calculate a lower bound for the quenching time. From Theorem 4.1, lower bounds are

$$\begin{aligned} \text{if } p &> 2q+1, \text{ then } T = (1-u_0(1))^{p+1}/(q+2)(p+1),\\ \text{if } q &\leq p \leq 2q+1, \text{ then } T = (1-u_0(1))^{2q+2}/(q+2)(2q+2), \end{aligned}$$

for quenching time T. If we choose, as Remark 1, $u_0(x) = \frac{1}{2}x^4$ (for $q=1$), then we have

$$\begin{aligned} T &\approx 0.0021 \text{ for } p=4, q=1,\\ T &\approx 0.0052 \text{ for } 1\leq p<3, q=1. \end{aligned}$$

References

[1] C.Y. Chan, Recent advances in quenching phenomena, Proc. Dynam. Sys. Appl. 2, 1996, pp. 107-113.

[2] C.Y. Chan, New results in quenching, Proc. of the First World Congress of Nonlinear Analysts, Walter de Gruyter, New York, 1996, pp. 427-434.

[3] C.Y. Chan, X. O. Jiang, Quenching for a degenerate parabolic problem due to a concentrated nonlinear source, Quart. Appl. Math., 62 (2004) 553-568.

[4] C.Y. Chan, N. Ozalp, Singular reactions-diffusion mixed boundary value quenching problems, Dynamical Systems and Applications, World Sci. Ser. Appl. Anal., 4, World Sci. Publ., River Edge, NJ, (1995) 127-137.

[5] C.Y. Chan, S.I. Yuen, Parabolic problems with nonlinear absorptions and releases at the boundaries, Appl. Math. Comput.,121 (2001) 203-209.

[6] C.W. Chang, Y.H. Hsu, H.T. Liu, Quenching behavior of parabolic problems with localized reaction terms, Mathematics and Statistics, 2(1), 2014, 48-53.

[7] K. Deng , M. Xu, Quenching for a nonlinear diffusion equation with a singular boundary condition, Z. Angew. Math. Phys. 50 (1999) 574-584.

[8] K. Deng, C-L. Zhao, Blow-up versus quenching, Comm. Appl. Anal. 7 (2003), 87-100.

[9] N. E. Dyakevich, Existence, uniqueness, and quenching properties of solutions for degenerate semilinear parabolic problems with second boundary conditions, J. Math. Anal. Appl. 338 (2008), 892-901.

[10] M. Fila, H.A. Levine, Quenching on the boundary, Nonlinear Anal. 21 (1993) 795–802.

[11] J.-S. Guo, S.C. Fu, J.C. Tsai, Blow up behavior for a semilinear heat equation with a nonlinear boundary condition, Tohoku Math. J. 55 (2003) 565-581.

[12] H. Kawarada, On solutions of initial-boundary problem for $u_t = u_{xx} + 1/(1-u)$, Publ. Res. Inst. Math. Sci. 10 (1975) 729-736.

[13] L. Ke, S.Ning, Quenching for degenerate parabolic equations, Nonlinear Anal. 34 (1998) 1123-1135.

[14] C.M. Kirk, C.A. Roberts, A quenching problem for the heat equation, Journal of Integral Equations and Applications,Vol. 14, 1 (2002) 53-72.

[15] C.M. Kirk, C.A. Roberts, A review of quenching results in the context of nonlinear volterra equations, Dynamics of Discrete and Impulsive Systems. Series A: Mathematical Analysis, 10 (2003) 343-356.

[16] W. E. Olmstead, C. A. Roberts, Critical speed for quenching, Advances in quenching, Dynamics of Discrete and Impulsive Systems. Series A: Mathematical Analysis, 8 (2001), no. 1, 77-88.

[17] T. Salin, On quenching with logarithmic singularity, Nonlinear Anal. 52 (2003) 261-289.

[18] R. Xu, C. Jin, T. Yu, Y. Liu, On quenching for some parabolic problems with combined power-type nonlinearities, Nonlinear Analysis Real World Applications, Vol. 13, 1 (2012) 333-339.

[19] Y. Yang, J. Yin, C. Jin, A quenching phenomenon for one-dimensional p-Laplacian with singular boundary flux, Appl. Math.Letters. 23 (2010) 955-959.

[20] Y. Zhi, C. Mu, The quenching behavior of a nonlinear parabolic equation with a nonlinear boundary outflux, Appl. Math.Comput. 184 (2007) 624-630.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 623 – 631

On the univalence of an integral operator

Dorina Răducanu *

Abstract

In this paper the method of Loewner chains is used to derive a fairly general and flexible univalence criterion for an integral operator. Two examples involving Bessel and hypergeometric functions are given. Our results include a number of known or new univalence criteria.

2000 AMS Classification: Primary 30C45, 30A10. Secondary 33C05, 33C10.

Keywords: Analytic functions, integral operator, Loewner chain, univalence criterion.

Received 09/03/2012 : Accepted 10/06/2013 Doi : 10.15672/HJMS.2015449650

1. Introduction

Let $\mathcal{U}_r = \{z \in \mathbb{C} : |z| < r,\ 0 < r \leq 1\}$ be the open disk of radius r centered at the origin and let $\mathcal{U} = \mathcal{U}_1$ be the open unit disk.

Denote by $\mathcal{A}$ the class of analytic functions in $\mathcal{U}$ which satisfy the usual normalization $f(0) = f'(0) - 1 = 0$.

Let $\mathcal{S}$ be the subclass of $\mathcal{A}$ consisting of univalent functions.

There are known numerous criteria which ensure that a function $f \in \mathcal{A}$ is in the class $\mathcal{S}$. In Theorem 1.1 some of these criteria are listed.

1.1. Theorem. *Let $f \in \mathcal{A}$. Then, each of the following three conditions implies that $f \in \mathcal{S}$:*

(1.1) $$(1 - |z|^2)\left|\frac{zf''(z)}{f'(z)}\right| \leq 1, \quad z \in \mathcal{U};$$

(1.2) $$\left|c|z|^2 + (1 - |z|^2)\frac{zf''(z)}{f'(z)}\right| \leq 1, \quad z \in \mathcal{U}$$

for some $c \in \mathbb{C}, |c| \leq 1, c \neq -1$;

(1.3) $$\left||z|^2\left[(c+1)f'(z)e^{-\int_0^z a(\tau)d\tau} - 1\right] + z(1 - |z|^2)a(z)\right| \leq 1, \quad z \in \mathcal{U}$$

---

*Faculty of Mathematics and Computer Science, Transilvania University of Braşov, 50091, Iuliu Maniu 50, Braşov, Romania
Email: draducanu@unitbv.ro

for some $c \in \mathbb{C}, |c| \leq 1, c \neq -1$ *and for* $a(z)$ *analytic function in* $\mathcal{U}$.

The univalence criterion given in (1.2) (see [1]) is an extension of Becker's univalence criterion (see [3], [4]) given in (1.1). The univalence criterion (1.3) was obtained by D. Tan (see [19]).

An extension of Becker's criterion, due to N. N. Pascu ensures the univalence of an integral operator.

1.2. Theorem. *([12]) Let* $f \in \mathcal{A}$ *and let* $\alpha \in \mathbb{C}$ *with* $\Re\alpha > 0$. *If*

$$\frac{1-|z|^{2\Re\alpha}}{\Re\alpha}\left|\frac{zf''(z)}{f'(z)}\right| \leq 1, \quad z \in \mathcal{U} \tag{1.4}$$

then, the integral operator

$$F_\alpha(z) = \left[\alpha\int_0^z \tau^{\alpha-1}f'(\tau)d\tau\right]^{1/\alpha} \tag{1.5}$$

is analytic and univalent in $\mathcal{U}$.

During the time many authors (see [5], [6], [7], [8], [9], [11], [18], etc.) have obtained numerous and various conditions which guarantee the univalence of a function in the class $\mathcal{A}$ or the univalence of an integral operator.

In this paper we are mainly interested on the integral operator

$$F_{\alpha,\beta}(z) = \left[\alpha\int_0^z \tau^{\alpha-1}(f'(\tau))^\beta d\tau\right]^{1/\alpha} \tag{1.6}$$

where the function f belongs to the class $\mathcal{A}$ and the parameters α and β are complex numbers such that the integral exists. Here and in the sequel every many-valued function is taken with the principal branch.

For the integral operator $F_{\alpha,\beta}(z)$ we establish a fairly general and flexible univalence criterion which contains a number of known or new results.

2. Univalence criterion

Before proving our main result we need a brief summary of the theory of Loewner chains.

A function $L(z,t) : \mathcal{U} \times [0,\infty) \to \mathbb{C}$ is said to be a *Loewner chain* or a *subordination chain* if:

(i) $L(z,t)$ is analytic and univalent in $\mathcal{U}$ for all $t \geq 0$.
(ii) $L(z,t) \prec L(z,s)$ for all $0 \leq t \leq s < \infty$, where the symbol " $\prec$ " stands for subordination.

The following result due to Pommerenke is often used to obtain univalence criteria.

2.1. Theorem. *([15], [16]) Let* $L(z,t) = a_1(t)z + \dots$ *be an analytic function in* $\mathcal{U}_r$ $(0 < r \leq 1)$ *for all* $t \geq 0$. *Suppose that:*

(i) $L(z,t)$ *is a locally absolutely continuous function of* $t \in [0,\infty)$, *locally uniform with respect to* $z \in \mathcal{U}_r$.
(ii) $a_1(t)$ *is a complex valued continuous function on* $[0,\infty)$ *such that* $a_1(t) \neq 0$, $\lim_{t\to\infty}|a_1(t)| = \infty$ *and*

$$\left\{\frac{L(z,t)}{a_1(t)}\right\}_{t\geq 0}$$

is a normal family of functions in $\mathcal{U}_r$.

(iii) *There exists an analytic function* $p : \mathcal{U} \times [0,\infty) \to \mathbb{C}$ *satisfying* $\Re p(z,t) > 0$ *for all* $(z,t) \in \mathcal{U} \times [0,\infty)$ *and*

$$z\frac{\partial L(z,t)}{\partial z} = p(z,t)\frac{\partial L(z,t)}{\partial t},\ z \in \mathcal{U}_r\ , a.e\ \ t \geq 0. \tag{2.1}$$

Then, for each $t \geq 0$*, the function* $L(z,t)$ *has an analytic and univalent extension to the whole disk* $\mathcal{U}$*, i.e* $L(z,t)$ *is a Loewner chain.*

Our main result contains sufficient conditions for the univalence of the integral operator $F_{\alpha,\beta}(z)$ defined by (1.6).

2.2. Theorem. *Let* $a(z)$ *be an analytic function in* $\mathcal{U}$ *and let* $f \in \mathcal{A}$*. Consider three complex numbers* α, β *and* c *such that* $\Re\alpha > 0, \beta \neq 0$ *and* $|c| \leq 1, c \neq -1$*. Suppose that:*

$$\left|(c+1)(f'(z))^{\beta} e^{-\int\limits_0^z a(\tau)d\tau} - 1\right| \leq 1,\ z \in \mathcal{U} \tag{2.2}$$

and

$$\left||z|^{2\alpha}\left[(c+1)(f'(z))^{\beta} e^{-\int\limits_0^z a(\tau)d\tau} - 1\right] + z\frac{1-|z|^{2\alpha}}{\alpha}a(z)\right| \leq 1,\quad z \in \mathcal{U}\setminus\{0\}\,. \tag{2.3}$$

Then, the integral operator

$$F_{\alpha,\beta}(z) = \left[\alpha\int\limits_0^z \tau^{\alpha-1}(f'(\tau))^{\beta}d\tau\right]^{1/\alpha}$$

is univalent in $\mathcal{U}$*, i.e. is in the class* $\mathcal{S}$*.*

Proof. Define the function

$$f_1(z,t) = \alpha\int\limits_0^{e^{-t}z} \tau^{\alpha-1}(f'(\tau))^{\beta}d\tau \quad z \in \mathcal{U},\ t \geq 0.$$

Since $f \in \mathcal{A}$, $e^{-t}z \in \mathcal{U}$ for all $t \geq 0$ and $z \in \mathcal{U}$, it follows that

$$f_1(z,t) = \left(e^{-t}z\right)^{\alpha} + \sum_{n=2}^{\infty} b_n\left(e^{-t}z\right)^{n+\alpha-1}$$

where $b_n \in \mathbb{C}, n \geq 2$. Consider the function $f_2(z,t)$ such that

$$f_1(z,t) = z^{\alpha}f_2(z,t) \quad z \in \mathcal{U},\ t \geq 0.$$

It is easy to check that $f_2(z,t)$ is analytic in $\mathcal{U}$ for all $t \geq 0$ and

$$f_2(z,t) = e^{-\alpha t} + \sum_{n=2}^{\infty} b_n e^{-t(n+\alpha-1)}z^{n-1}.$$

Since the function $a(z)$ is analytic in $\mathcal{U}$ it follows that the function $f_3(z,t)$ defined by

$$f_3(z,t) = \left(e^{\alpha t} - e^{-\alpha t}\right)e^{\int\limits_0^{e^{-t}z} a(\tau)d\tau}$$

is analytic in $\mathcal{U}$ for all $t \geq 0$.

Then, the function $f_4(z,t)$ given by

$$f_4(z,t) = f_2(z,t) + \frac{1}{c+1}f_3(z,t) \quad z \in \mathcal{U},\ t \geq 0$$

is also analytic in $\mathcal{U}$.

We have

$$f_4(0,t) = f_2(0,t) + \frac{1}{c+1} f_3(0,t) = \frac{e^{\alpha t}}{c+1}(1 + ce^{-2\alpha t}).$$

The conditions $\Re\alpha > 0$ and $|c| \le 1, c \ne -1$ yield $f_4(0,t) \ne 0$ for all $t \ge 0$. Thus, there exists an open disk $\mathcal{U}_{r_1}$ $(0 < r_1 \le 1)$ in which $f_4(z,t) \ne 0$ for all $t \ge 0$. Therefore, we can choose an analytic branch of $[f_4(z,t)]^{1/\alpha}$, which will be denoted by $f_5(z,t)$.

Making use of the previous results, we obtain that the function

$$L(z,t) = zf_5(z,t)$$

or

$$L(z,t) = \left[\alpha \int_0^{e^{-t}z} \tau^{\alpha-1}(f'(\tau))^\beta d\tau + \frac{1}{c+1}\left(e^{\alpha t} - e^{-\alpha t}\right) z^\alpha e^{\int_0^{e^{-t}z} a(\tau)d\tau}\right]^{1/\alpha}$$

is analytic in $\mathcal{U}_{r_1}$ for all $t \ge 0$.

We have $L(z,t) = a_1(t)z + \dots$ for $z \in \mathcal{U}_{r_1}$ and $t \ge 0$, where

$$a_1(t) = e^t \left(\frac{1 + ce^{-2\alpha t}}{c+1}\right)^{1/\alpha}, \quad t \ge 0.$$

From $\Re\alpha > 0$ and $|c| \le 1, c \ne -1$ we obtain $a_1(t) \ne 0$ and $\lim_{t\to\infty} |a_1(t)| = \infty$.

Let $r_2 \in (0, r_1]$ and let $K = \overline{\{z \in \mathbb{C} : |z| \le r_2\}}$. Since the function $L(z,t)$ is analytic in $\mathcal{U}_{r_1}$, there exists $M > 0$ such that $|L(z,t)| \le Me^t$ for $z \in K$ and $t \ge 0$. Also, for $t \ge 0$, it is easy to see that there exists $N > 0$ such that $|a_1(t)| > Ne^t$. It follows that

$$\left|\frac{L(z,t)}{a_1(t)}\right| \le \frac{M}{N}\ , \text{for } z \in K \text{ and } t \ge 0.$$

Thus, $\{L(z,t)/a_1(t)\}_{t\ge 0}$ is a normal family of functions in $\mathcal{U}_{r_1}$.

Elementary calculations show that $\frac{\partial L}{\partial z}(z,t)$ is analytic in $\mathcal{U}_{r_1}$. It follows that $\left|\frac{\partial L}{\partial z}(z,t)\right|$ is bounded on $[0,T]$ for any fixed $T > 0$ and $z \in \mathcal{U}_{r_3}$ $(0 < r_3 \le r_1)$. Therefore, the function $L(z,t)$ is locally absolutely continuous on $[0,\infty)$ locally uniform with respect to $z \in \mathcal{U}_{r_1}$.

Consider the function $p(z,t)$ defined by

$$p(z,t) = z\frac{\partial L}{\partial z}(z,t) / \frac{\partial L}{\partial t}(z,t).$$

In order to prove that the function $p(z,t)$ has an analytic extension in $\mathcal{U}$ and $\Re p(z,t) > 0$ for all $t \ge 0$, we will show that the function $w(z,t)$ given by

$$w(z,t) = \frac{p(z,t) - 1}{p(z,t) + 1} \quad z \in \mathcal{U}_{r_1},\ t \ge 0$$

has an analytic extension in $\mathcal{U}$ and $|w(z,t)| < 1$, for all $z \in \mathcal{U}$ and $t \ge 0$.

Lengthy but elementary calculations give

$$w(z,t) = e^{-2t\alpha}\left[(c+1)(f'(e^{-t}z))^\beta e^{-\int_0^{e^{-t}z} a(\tau)d\tau} - 1\right] + \frac{1}{\alpha}(1 - e^{-2t\alpha})e^{-t}za(e^{-t}z).$$

It is easy to check that $w(z,t)$ is an analytic function in $\mathcal{U}$. We have $w(0,t) = ce^{-2t\alpha}$ and thus

$$(2.4) \qquad |w(0,t)| = |c|e^{-2t\Re\alpha} < 1, \text{ for all } t > 0.$$

For $t = 0$ we obtain

$$w(z,0) = (c+1)(f'(z))^{\beta} e^{-\int\limits_0^z a(\tau)d\tau} - 1, \quad z \in \mathcal{U}.$$

Inequality (2.2) from the hyphothesis, yields

(2.5) $\quad |w(z,0)| < 1 \quad z \in \mathcal{U}.$

Let $t > 0$ and let $z \neq 0$. Since $|e^{-t}z| \leq e^{-t} < 1$ for all $z \in \bar{\mathcal{U}} = \{z \in \mathbb{C} : |z| \leq 1\}$, it follows that $w(z,t)$ is analytic in $\bar{\mathcal{U}}$. Making use of the maximum modulus principle we obtain that, for each fixed $t > 0$, there exists $\theta \in \mathbb{R}$ such that :

$$|w(z,t)| < \max_{|z|=1} |w(z,t)| = |w(e^{i\theta}, t)|.$$

Denote $u = e^{-t}e^{i\theta}$. Then, $|u| = e^{-t}$ and thus

$$|w(e^{i\theta}, t)| = \left| |u|^{2\alpha} \left[(c+1)(f'(u))^{\beta} e^{-\int\limits_0^u a(\tau)d\tau} - 1 \right] + \frac{1 - |u|^{2\alpha}}{\alpha} u a(u) \right|.$$

Inequality (2.3), from the hyphothesis, shows that

(2.6) $\quad |w(e^{i\theta}, t)| \leq 1.$

Combining (2.4), (2.5) and (2.6) we immediately get $|w(z,t)| < 1$ for all $z \in \mathcal{U}$ and $t \geq 0$. Therefore, the function $p(z,t)$ has an analytic extension in $\mathcal{U}$ and $\Re p(z,t) > 0$ for $(z,t) \in \mathcal{U} \times [0, \infty)$.

Since all the conditions of Theorem 2.1 are satisfied we can conclude that the function $L(z,t)$ has an analytic and univalent extension in $\mathcal{U}$ for all $t \geq 0$. For $t = 0$, we have $L(z,0) = F_{\alpha,\beta}(z)$ and thus, the function $F_{\alpha,\beta}(z)$ given by (1.6) is analytic and univalent in $\mathcal{U}$. With this the proof is complete. □

Remark. The univalence condition (1.3) can be derived from Theorem 2.2 for $\alpha = \beta = 1$.

3. Specific univalence criteria

Many new or known univalence criteria can be generated with Theorem 2.2 and specific choiches of the functions $a(z)$ and $f(z)$. In this section some of these univalence criteria are listed.

1. Consider first

$$a(z) = \beta \frac{f''(z)}{f'(z)}, \quad z \in \mathcal{U}, \quad f \in \mathcal{A}.$$

Then, making use of Theorem 2.2 we immediately obtain the following result.

3.1. Theorem. *Let $f \in \mathcal{A}$ and let α, β, c be complex numbers such that $\Re\alpha > 0, \beta \neq 0$ and $|c| \leq 1, c \neq -1$. If*

(3.1) $$\left| c|z|^{2\alpha} + \frac{\beta}{\alpha}(1 - |z|^{2\alpha}) \frac{zf''(z)}{f'(z)} \right| \leq 1, \quad z \in \mathcal{U}$$

then the integral operator $F_{\alpha,\beta}(z)$ defined by (1.6) is in the class $\mathcal{S}$.

Remark.

(i) For $\beta = 1$, Theorem 3.1 reduces to a result obtained by V. Pescar [13].

(ii) Setting $\alpha = \beta = 1$ in Theorem 3.1, we obtain the univalence criterion given in (1.2).

(iii) With $c = 0$ and $\beta = 1$, inequality (3.1) specializes to

$$\left|\frac{1-|z|^{2\alpha}}{\alpha}\frac{zf''(z)}{f'(z)}\right| \le 1, \quad z \in \mathcal{U}. \tag{3.2}$$

Using the next inequality

$$\left|\frac{1-|z|^{2\alpha}}{\alpha}\right| \le \frac{1-|z|^{2\Re\alpha}}{\Re\alpha} \tag{3.3}$$

in (3.2) we get the univalence condition (1.4) which guarantees the univalence of the integral operator $F_\alpha(z)$ given by (1.5).

Let $g_\nu : \mathcal{U} \to \mathbb{C}$ be the normalized Bessel function of the first kind (see [2]) with Taylor expansion

$$g_\nu(z) = z + \sum_{n=1}^{\infty} \frac{(-1)^n z^{n+1}}{4^n n!(\nu+1)\ldots(\nu+n)}.$$

For $\nu = \frac{1}{2}$ we have $g_{\frac{1}{2}}(z) = \sqrt{z}\sin\sqrt{z}$.

The next result follows from Theorem 3.1 with $f(z) = g_\nu(z)$.

3.2. Corollary. *Let $\nu > 0$ and let α, β, c be complex numbers such that $0 < |\beta| \le \frac{2(4\nu^2+9\nu+3)}{4\nu+9}\Re\alpha$ and $|c| \le 1, c \ne -1$. Then the function*

$$F_{\alpha,\beta,\nu}(z) = \left[\alpha \int_0^z \tau^{\alpha-1}(g_\nu'(\tau))^\beta d\tau\right]^{1/\alpha}, \quad z \in \mathcal{U} \tag{3.4}$$

is in the class $\mathcal{S}$. In particular, if $0 < |\beta| \le \frac{17}{11}\Re\alpha$ and $|c| \le 1, c \ne -1$, then the function

$$F_{\alpha,\beta,\frac{1}{2}}(z) = \left[\alpha \int_0^z \tau^{\alpha-1}\left(\frac{\sin\sqrt{\tau}+\sqrt{\tau}\cos\sqrt{\tau}}{2\sqrt{\tau}}\right)^\beta d\tau\right]^{1/\alpha}$$

is in $\mathcal{S}$.

Proof. Replace $f(z) = g_\nu(z)$ in left-hand side of (3.1). Making use of the triangle inequality and (3.3) we have

$$\begin{aligned} &\left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\frac{zf''(z)}{f'(z)}\right| \\ &= \left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\frac{zg_\nu''(z)}{g_\nu'(z)}\right| \\ &\le |c||z|^{2\Re\alpha} + \frac{|\beta|}{\Re\alpha}(1-|z|^{2\Re\alpha})\left|\frac{zg_\nu''(z)}{g_\nu'(z)}\right|. \end{aligned}$$

Since $0 < |\beta| \le \frac{2(4\nu^2+9\nu+3)}{4\nu+9}\Re\alpha, |c| \le 1, c \ne -1$ and making use of

$$\left|\frac{zg_\nu''(z)}{g_\nu'(z)}\right| \le \frac{4\nu+9}{2(4\nu^2+9\nu+3)}, \quad z \in \mathcal{U}, \quad \nu > 0$$

(see [6]), we obtain that

$$\begin{aligned} &|c||z|^{2\Re\alpha} + \frac{|\beta|}{\Re\alpha}(1-|z|^{2\Re\alpha})\left|\frac{zg_\nu''(z)}{g_\nu'(z)}\right| \\ &\le |z|^{2\Re\alpha} + \frac{|\beta|}{\Re\alpha}(1-|z|^{2\Re\alpha})\frac{4\nu+9}{2(4\nu^2+9\nu+3)} \le |z|^{2\Re\alpha} + 1 - |z|^{2\Re\alpha} = 1. \end{aligned}$$

It follows that inequality (3.1) holds true and therefore, the function $F_{\alpha,\beta,\nu}(z)$ defined by (3.4) is in $\mathcal{S}$. The particular case follows from the first part by setting $\nu = \dfrac{1}{2}$. □

2. Let $g \in \mathcal{A}$. Choosing

$$f(z) = \int_0^z \frac{g(\tau)}{\tau} d\tau, \quad z \in \mathcal{U}$$

in Theorem 2.2 we obtain easily a univalence criterion for another well known integral operator.

3.3. Theorem. *Let $g \in \mathcal{A}$ and let α, β, c be complex numbers such that $\Re\alpha > 0, \beta \neq 0$ and $|c| \leq 1, c \neq -1$. Suppose that*

$$\left|(c+1)\left(\frac{g(z)}{z}\right)^{\beta} e^{-\int_0^z a(\tau)d\tau} - 1\right| \leq 1, \quad z \in \mathcal{U}$$

and

$$\left||z|^{2\alpha}\left[(c+1)\left(\frac{g(z)}{z}\right)^{\beta} e^{-\int_0^z a(\tau)d\tau} - 1\right] + \frac{1-|z|^{2\alpha}}{\alpha} z a(z)\right| \leq 1, \quad z \in \mathcal{U} \setminus \{0\}.$$

Then the integral operator

$$G_{\alpha,\beta}(z) = \left[\alpha \int_0^z \tau^{\alpha-1}\left(\frac{g(\tau)}{\tau}\right)^{\beta} d\tau\right]^{1/\alpha}, \quad z \in \mathcal{U} \tag{3.5}$$

is in the class $\mathcal{S}$.

3. Consider $a(z)$ defined by

$$a(z) = \beta\left(\frac{g'(z)}{g(z)} - \frac{1}{z}\right), \quad z \in \mathcal{U}, \quad g \in \mathcal{A}.$$

Then, making use of Theorem 3.2 we get the following result.

3.4. Corollary. *Let $g \in \mathcal{A}$ and let $\alpha, \beta, c \in \mathbb{C}$ with $\Re\alpha > 0, \beta \neq 0$ and $|c| \leq 1, c \neq -1$. If*

$$\left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\left(\frac{zg'(z)}{g(z)} - 1\right)\right| \leq 1, \quad z \in \mathcal{U} \tag{3.6}$$

then the function $G_{\alpha,\beta}(z)$ defined by (3.5) is in the class $\mathcal{S}$.

Suppose that the function g in Corollary 3.2 is in $\mathcal{S}$. Then we have the following result which shows that the integral operator $G_{\alpha,\beta}(z)$ preserves univalency.

3.5. Corollary. *Let $g \in \mathcal{S}$ and let $\alpha, \beta, c \in \mathbb{C}$ with $c \neq -1, 0 < |\beta| \leq \min\left\{\frac{\Re\alpha}{2}, \frac{1}{4}\right\}$ and $\Re\alpha > 0$. If*

$$|c| \leq \begin{cases} 1 - \dfrac{2|\beta|}{\Re\alpha}, & \Re\alpha \in (0, \dfrac{1}{2}) \\ 1 - 4|\beta|, & \Re\alpha \in [\dfrac{1}{2}, \infty) \end{cases} \tag{3.7}$$

then the function $G_{\alpha,\beta}(z)$ is in $\mathcal{S}$.

Proof. Making use of the triangle inequality in left-hand side of (3.6) we obtain

$$\left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\left(\frac{zg'(z)}{g(z)} - 1\right)\right|$$
$$\leq |c||z|^{2\Re\alpha} + \frac{|\beta|}{\Re\alpha}(1-|z|^{2\Re\alpha})\left[\left|\frac{zg'(z)}{g(z)}\right| + 1\right].$$

Let $g \in \mathcal{S}$. Then

$$\left|\frac{zg'(z)}{g(z)}\right| \leq \frac{1+|z|}{1-|z|}, \quad z \in \mathcal{U}.$$

It follows that

$$\left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\left(\frac{zg'(z)}{g(z)} - 1\right)\right| \leq |c| + \frac{2|\beta|}{\Re\alpha}\frac{1-|z|^{2\Re\alpha}}{1-|z|}. \tag{3.8}$$

Denote $x = |z|$ and $a = \Re\alpha$. Consider the function $\phi : [0,1) \to \mathbb{R}$ defined by

$$\phi(x) = \frac{1-x^{2a}}{1-x}.$$

It is easy to check that

$$\phi(x) \leq \begin{cases} 1, & a \in (0, \frac{1}{2}) \\ 2a, & a \in [\frac{1}{2}, \infty). \end{cases} \tag{3.9}$$

Combining (3.8) and (3.9) we have

$$\left|c|z|^{2\alpha} + \frac{\beta}{\alpha}(1-|z|^{2\alpha})\left(\frac{zg'(z)}{g(z)} - 1\right)\right| \leq \begin{cases} |c| + \frac{2|\beta|}{\Re\alpha}, & \Re\alpha \in (0, \frac{1}{2}) \\ |c| + 4|\beta|, & \Re\alpha \in [\frac{1}{2}, \infty) \end{cases}$$

Inequality (3.7) from hypothesis shows that the condition (3.6) is satisfied and thus, making use of Corollary 3.2 we obtain that the function $G_{\alpha,\beta}(z)$ is in $\mathcal{S}$. With this the proof is complete. □

3.6. Corollary. *Let $\alpha, \beta, c \in \mathbb{C}$ with $c \neq -1, 0 < |\beta| \leq \min\left\{\frac{\Re\alpha}{2}, \frac{1}{4}\right\}, \Re\alpha > 0$. If inequality (3.7) holds true, then the function $K_{\alpha,\beta}(z) = z\left[{}_2F_1(\alpha, 2\beta; 1+\alpha; z)\right]^{1/\alpha}$ is in the class $\mathcal{S}$. The symbol ${}_2F_1(a,b;c;z)$ denotes the well known hypergeometric function.*

Proof. The Koebe function $k(z) = \dfrac{z}{(1-z)^2}$ is in $\mathcal{S}$. Applying Corollary 3.3 we obtain that the function

$$K_{\alpha,\beta}(z) := \left[\alpha \int_0^z \tau^{\alpha-1}\left(\frac{k(\tau)}{\tau}\right)^{\beta} d\tau\right]^{1/\alpha} = \left[\alpha \int_0^z \tau^{\alpha-1}(1-\tau)^{-2\beta}\, d\tau\right]^{1/\alpha}$$

is also in $\mathcal{S}$. With the substitution $\tau = uz$ the function $K_{\alpha,\beta}(z)$ becomes

$$K_{\alpha,\beta}(z) = z\left[\alpha \int_0^1 u^{\alpha-1}(1-uz)^{-2\beta} du\right]^{1/\alpha} = z\left[{}_2F_1(\alpha, 2\beta; 1+\alpha; z)\right]^{1/\alpha}.$$

Thus, the proof is completed. □

Remark. Similar results with the one given in Corollary 3.3 can be found in [9], [14].

4. Let $g_1, \ldots, g_m \in \mathcal{A}$ and $\delta_1, \ldots, \delta_m \in \mathbb{C} \setminus \{0\}$. Setting

$$f(z) = \int_0^z \prod_{k=1}^{m} \left(\frac{g_k(\tau)}{\tau}\right)^{\frac{\delta_k}{\beta}} d\tau$$

in Theorem 2.2 or Theorem 3.1 we can easily obtain various univalence criteria for the integral operator

$$G_{\delta_1,\ldots,\delta_m}(z) = \left[\alpha \int_0^z \tau^{\alpha-1} \prod_{k=1}^{m} \left(\frac{g_k(\tau)}{\tau}\right)^{\delta_k} d\tau\right]^{1/\alpha}$$

which has been studied by many authors (see [2], [5], [6], [8], [18], etc.)

From the previous examples, it is clear that one can generate many univalence criteria with Theorem 2.2 and suitable choices of the functions $a(z)$ and $f(z)$.

References

[1] L. V. Ahlfors,*Sufficient conditions for quasiconformal extension*, Ann. Math. Studies., **79** (1974), 23-29.

[2] A. Baricz, B. A. Frasin ,*Univalence of integral operator involving Bessel functions*, Appl. Math. Lett., **23**(4)(2010), 371-376.

[3] J. Becker, *Löwnersche differential gleichung und quasikonform fortsetzbare schlichte functionen*, J. Reine Angew. Math.**255** (1972), 23-43.

[4] J. Becker, *Conformal mappings with quasiconformal extensions*, Aspects of Contemporary Complex Analysis, Ed. by D. A. Brannan and J. G. Clunie, Acad. Press, 1980, 37-77.

[5] D. Breaz, N. Breaz, H. M. Srivastava, *An extension of the univalent condition for a family of integral operators*, Appl. Math. Lett., **22**(2009), 41-44.

[6] E. Deniz, H. Orhan, H. M. Srivastava, *Some sufficient conditions for univalence of certain families of integral operators involving generalized Bessel functions*, Taiwanese J. Math., **15**(2)(2011), 883-971.

[7] E. Deniz, D. Răducanu, H. Orhan, *On the univalence of an integral operator defined by Hadamard product*, Appl. Math. Lett., **25**(2012), 179-184.

[8] B. A. Frasin, *Certain sufficient conditions for univalence of two integral operators*, European J. Pure Appl. Math., **3**(6)(2010), 1141-1149.

[9] S. Moldoveanu, N. N. Pascu, *Integral operators which preserve the univalence*, Mathematica (Cluj), **32**(55)(1990), 159-166.

[10] H. Ovesea, *An extension of Lewandowski's univalence criterion*, Demonstratio Math., XXIX, **4**(1996), 699-706.

[11] H. Ovesea, *A generalization of Ruscheweyh's univalence criterion*, J. Math. Anal. Appl., **258** (2001), 102-109.

[12] N. N. Pascu, *An improvement of Becker's univalence criterion*, Commemorative Session Simion Stoilov, Univ. of Braşov, (Preprint) (1987), 43-48.

[13] V. Pescar, *A new generalization of Ahlfors and Becker's criterion of univalence*, Bull. Math. Malaysian Math. Soc. (Second Series), **19** (1996), 53-54.

[14] V. Pescar, D. Breaz, *The Univalence of Integrals Operators*, Prof. M. Drinov Acad. Publish. House, Sofia, 2008.

[15] Ch. Pommerenke, *Über die Subordination analytischer Funktionen*, J. Reine Angew Math., **218** (1965), 159-173.

[16] Ch. Pommerenke, *Univalent Functions*, Vandenhoeck Ruprecht in Göttingen, 1975.

[17] D. Răducanu, H. Orhan, E. Deniz, *On some sufficient conditions for univalence*, An. Şt. Univ. Ovidius, Constanţa, Ser. Math., **18**(2)(2010), 217-222.

[18] H. M. Srivastava, E. Deniz, H. Orhan, *Some general univalence criteria for a family of integral operators*, Appl. Math. Comp., **215**(2010), 3696-3701.

[19] D. Tan, *Quasiconformal extension and univalency criteria*, Mich. Math. J., **39**(1992), 163-172.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 633 – 640

GCED and reciprocal GCED matrices

Zahid Raza * † Seemal Abdul Waheed ‡

Abstract

We have given structure theorems for a GCED (greatest common exponential divisor) and Reciprocal GCED matrix. We have also calculated the value of the determinant of these matrices. The formulae for the inverse and determinant of GCED and Reciprocal GCED matrices defined on an exponential divisor closed set have been determined.

2000 AMS Classification: 05-xx,15-xx, 15B36, 15A36

Keywords: GCED matrix, Reciprocal GCED matrix, exponential divisor, exponential divisor closed set

Received 29/12/2011 : Accepted 15/02/2014 Doi : 10.15672/HJMS.2015449417

1. Introduction

Let $S = \{x_1, x_2, \ldots, x_n\}$ be a finite ordered set of distinct positive integers. The matrix (S) where $s_{ij} = (x_i, x_j)$ =greatest common divisor of x_i and x_j, is called the greatest common divisor(GCD) matrix on the set S. A set $S = \{x_1, x_2, \ldots, x_n\}$ is said to be factor closed if for every $x_i \in S$, and $d \mid x_i$ then $d \in S$.

In 1876, H.J. Smith [7] proved that the determinant of a GCD matrix on $S = \{1, 2, \ldots, n\}$ is equal to $\varphi(1)\varphi(2)\cdots\varphi(n)$ where φ is Euler's totient function. The result holds if S is a factor closed set. The structure theorems for Reciprocal GCD matrices and LCM (least common multiple) matrices were determined by S.J. Beslin [2]. The structures of Power GCD matrix, Power LCM matrix, Reciprocal LCM matrix, GCD Reciprocal LCM matrix, GCUD (greatest common unitary divisor) Reciprocal LCUM (least common unitary multiple) matrices have been determined [1, 3, 5, 9]. Research has also been extended to divisibility properties of such matrices and their applications [4, 6]. It is worth to note that the structures of most of the above mentioned matrices have been determined on factor closed sets, gcd closed sets, lcm closed sets or unitary divisor closed sets or on sets contained in factor closed sets. This has motivated the authors to follow the same direction.

*Department of Mathematics, National University of Computer and Emerging Sciences, B-Block, Faisal Town, Lahore, Pakistan. Email: zahid.raza@nu.edu.pk

†Department of Mathematics, College of Sciences, The University of Sharjah, P.O.Box: 27272, Sharjah, United Arab Emirates.

‡Department of Mathematics, National University of Computer and Emerging Sciences, B-Block, Faisal Town, Lahore, Pakistan. Email:seemalw11@gmail.com

We recall that an integer $d = \prod_{i=1}^{t} {p_i}^{a_i}$ is said to be an exponential divisor of $m = \prod_{i=1}^{t} {p_i}^{b_i}$, if $a_i | b_i$ for every $1 \le i \le t$ and is denoted by $d|_e m$. This notion was introduced by M. V. Subrarao [8]. Note that unlike divisor and unitary divisor, 1 is not an exponential divisor for every $m > 1$. By convention $1|_e 1$. The smallest exponential divisor of $m > 1$ is its square free kernel $\kappa(m) = \prod_{i=1}^{r} p_i$ [10].

Two integers n and m have common exponential divisor if and only if they have the same prime factors. Two integers $m = \prod_{i=1}^{r} {p_i}^{b_i}$ and $n = \prod_{i=1}^{r} {p_i}^{c_i}$ are exponentially co-prime if $(b_i, c_i) = 1$ for every $1 \le i \le r$. We denote the GCED (greatest common exponential divisor) of two integers m and n by $(m, n)_e$. By convention $(1, 1)_{(e)} = 1$ and $(1, m)_{(e)}$ does not exist for every $m > 1$.

A set $S = \{x_1, x_2, x_3, \ldots, x_n\}$ is said to be an exponential divisor closed set if the exponential divisor of every element of S belongs to S. For example the set $\{12, 18, 36\}$ is not an exponential divisor closed set. But, $\{6, 12, 18, 36\}$ is an exponential divisor closed set.

Similarly, a set $S = \{x_1, x_2, x_3, \ldots, x_n\}$ is said to be GCED closed if $(x_i, x_j)_{(e)} \in S$ for every $x_i, x_j \in S$. Note that $\{6, 12, 18, 36\}$ is also a GCED closed set.

The exponential convolution of two arithmetic functions f and g is given as

$$(f \odot g)(n) = \sum_{k_1 l_1 = m_1} \cdots \sum_{k_r l_r = m_r} f({p_1}^{k_1} {p_2}^{k_2} \ldots {p_r}^{k_r}) g({p_1}^{l_1} {p_2}^{l_2} \ldots {p_r}^{l_r}),$$

where $n = {p_1}^{m_1} {p_2}^{m_2} \ldots {p_r}^{m_r}$.

The inverse with respect to $\odot$ of the constant function 1 is called the exponential analogue of Möbius function and is denoted by $\mu^{(e)}$. It should be noted that the sets considered in section 2 are such that the GCED of every two elements exists.

2. Structure of GCED matrix

Let $T = \{x_1, x_2, x_3, \ldots, x_n\}$ be an ordered set of distinct positive integers greater than 1. The $n \times n$ matrix $T_{(e)} = (t_{ij})_{(e)}$ having $t_{ij} = (x_i, x_j)_{(e)}$ as its ij^{th} entry is referred as the GCED (greatest common exponential divisor) matrix on the set T, where $(x_i, x_j)_{(e)}$ is the greatest common exponential divisor of x_i and x_j. Let $R = \{y_1, y_2, y_3, \ldots, y_m\}$ which is ordered by $y_1 < y_2 < y_3 < \ldots < y_m$ be a minimal exponential divisor-closed set containing T. We refer R the exponential closure of the set T. It is easy to see that GCED matrices are symmetric. We always assume that $x_1 < x_2 < x_3 < \cdots < x_n$ in T.

We define arithmetic function $g(n)$ as follows:

$$g(n) = \sum_{a_1 b_1 = c_1} \sum_{a_2 b_2 = c_2} \cdots \sum_{a_r b_r = c_r} {p_1}^{a_1} {p_2}^{a_2} \cdots {p_r}^{a_r} \mu^{(e)}({p_1}^{b_1} {p_2}^{b_2} \cdots {p_r}^{b_r}), \tag{2.1}$$

where $n = {p_1}^{c_1} {p_2}^{c_2} \cdots {p_r}^{c_r}$.

2.1. Theorem. *Let $R = \{y_1, y_2, \ldots, y_m\}$ be the exponential closure of the set $T = \{x_1, x_2, \ldots, x_n\}$, where $y_1 < y_2 < y_3 < \cdots < y_m$ and $x_1 < x_2 < x_3 < \cdots < x_n$.*

Define the $n \times m$ matrix $C = (c_{ij})$ by

$$c_{ij} = \begin{cases} 1, & y_j |_e x_i \\ 0, & \textit{otherwise} \end{cases}$$

and the $m \times m$ diagonal matrix by

$$\Psi = diag(g(y_1), g(y_2), \ldots, g(y_m)).$$

Then,

$$T_{(e)} = C \Psi C^t.$$

Proof. The ij^{th} entry of $C\Psi C^t$ is equal to

$$(C\Psi C^t)_{ij} = \sum_{k=1}^{n} c_{ik} g(y_k) c_{jk} = \sum_{y_k |_e x_i, y_k |_e x_j} g(y_k) = \sum_{y_k |_e (x_i, x_j)_{(e)}} g(y_k),$$

where the function g is defined in Equation 2.1.
By Möbius Inversion Exponential formula, we have,

$$\sum_{d|_e n} g(d) = n.$$

Finally, we get,

$$(C\Psi C^t)_{ij} = (x_i, x_j)_{(e)}.$$

2.2. Theorem. *Let $R = \{y_1, y_2, \dots, y_m\}$ be the exponential closure of the set $T = \{x_1, x_2, \dots, x_n\}$ where $y_1 < y_2 < y_3 < \dots < y_m$ and $x_1 < x_2 < x_3 < \dots < x_n$. Then*

$$\det T_{(e)} = \sum_{1 \le k_1 < k_2 < \dots < k_n \le m} (\det C_{(k_1,k_2,\dots,k_n)})^2 g(y_{k_1}) g(y_{k_2}) \dots g(y_{k_n}),$$

where $C_{(k_1,k_2,\dots,k_n)}$ is the sub matrix of C consisting of the ${k_1}^{th}$, ${k_2}^{th}, \dots, {k_n}^{th}$ columns of C.

Proof. By Theorem 2.1, we have, $T_{(e)} = (C\Psi^{\frac{1}{2}})(C\Psi^{\frac{1}{2}})^t$. Thus we can write $E = C\Psi^{\frac{1}{2}}$ which leads to $T_{(e)} = EE^t$. By applying Cauchy-Binet formula, we get

$$\begin{aligned} \det(T)_{(e)} &= \sum_{1 \le k_1 < k_2 < \dots < k_n \le m} \det E_{(k_1,k_2,\dots,k_n)} \det {E^t}_{(k_1,k_2,\dots,k_n)} \\ &= \sum_{1 \le k_1 < k_2 < \dots < k_n \le m} (\det E_{(k_1,k_2,\dots,k_n)})^2, \end{aligned}$$

where $E_{(k_1,k_2,\dots,k_n)}$ is the sub matrix of E consisting of the ${k_1}^{th}, {k_2}^{th}, \dots, {k_n}^{th}$ columns of E.

$$\det E_{(k_1,k_2,\dots,k_n)} = \sqrt{g(y_{k_1}) g(y_{k_2}) \dots g(y_{k_n})} \det C_{(k_1,k_2,\dots,k_n)}.$$

Hence,

$$\det T_{(e)} = \sum_{1 \le k_1 < k_2 < \dots < k_n \le m} (\det C_{(k_1,k_2,\dots,k_n)})^2 g(y_{k_1}) g(y_{k_2}) \dots g(y_{k_n}).$$

2.3. Corollary. *Let $T = \{x_1, x_2, \dots, x_n\}$ be a finite ordered set of distinct positive integers. If $T = R$, then the determinant of GCED matrix $T_{(e)}$ defined on T is given as:*

$$\det T_{(e)} = \prod_{k=1}^{n} g(x_k).$$

Proof. Note that C is a lower triangular matrix with diagonal $(1, 1, \dots, 1)_n$. This implies that $\det C = 1$. Since the determinant of a diagonal matrix is equal to the product of its diagonal entries, hence the desired outcome achieved.

2.4. Corollary. *If $T_{(e)}$ is an $n \times n$ GCED matrix on a finite ordered set of distinct integers denoted by $T = \{x_1, x_2, \dots, x_n\}$, then the trace is given as:*

$$trT_{(e)} = \sum_{k=1}^{n} x_i.$$

2.5. Lemma. *Let $T_{(e)} = (t_{ij})_{(e)}$ is an $n \times n$ GCED matrix defined on an exponential divisor closed set T. Consider $n \times n$ matrix $C = (c_{ij})$ as defined in Theorem 2.1. Then, the $n \times n$ matrix $W = (w_{ij})$ defined by*

$$w_{ij} = \begin{cases} \mu^{(e)}(\frac{x_i}{x_j}), & x_j|_e x_i \\ 0, & otherwise \end{cases}$$

is the inverse of C.

Proof. The ij^{th} entry of CW is given by

$$(CW)_{ij} = \sum_{k=1}^{n} c_{ik} w_{kj} = \sum_{x_k|_e x_i, x_j|_e x_k} \mu^{(e)}(\frac{x_k}{x_j}) = \sum_{x_d|_e \frac{x_i}{x_j}} \mu^{(e)}(x_d) = \begin{cases} 1, & if \quad x_i = x_j \\ 0, & otherwise \end{cases}$$

If $\frac{x_i}{x_j}$ is not an integer then no x_d divides $\frac{x_i}{x_j}$. If $x_i = x_j$ then, $1|_e 1$ and $\mu^{(e)}(1) = 1$.

2.6. Theorem. *Let $T_{(e)}$ be an $n \times n$ GCED matrix on an exponential divisor closed set. Then, its inverse matrix $(A)_{(e)} = (a_{ij})_{(e)}$ is given as*

$$(a_{ij})_{(e)} = \sum_{x_i|_{(e)} x_k, x_j|_{(e)} x_k} \frac{\mu^{(e)} \frac{x_d}{x_i} \mu^{(e)} \frac{x_d}{x_j}}{g(x_d)}.$$

Proof. Since $T_{(e)} = (C\Psi C^t)$ and Lemma 2.5 suggests that, $C^{-1} = W$, therefore

$$(T)^{-1}_{(e)} = (C\Psi C^t)^{-1} = W^t \Psi^{-1} W,$$

where ij^{th} entry of $(T)^{-1}_{(e)}$ is given as

$$(a_{ij})_{(e)} = \sum_{x_i|_{(e)} x_d, x_j|_{(e)} x_d} \frac{\mu^{(e)} \frac{x_d}{x_i} \mu^{(e)} \frac{x_d}{x_j}}{g(x_d)}.$$

Hence, the required result.

3. Structure of Reciprocal GCED matrix

Let $T = \{x_1, x_2, x_3, \ldots, x_n\}$ be an ordered set of positive integers greater than 1. The $n \times n$ matrix $\overline{T}_{(e)} = (t_{ij})_{(e)}$ having $t_{ij} = \frac{1}{(x_i, x_j)_{(e)}}$ as its ij^{th} entry on T is called a Reciprocal GCED matrix. It is easy to note that Reciprocal GCED matrices are symmetric. We always assume that $x_1 < x_2 < x_3 < \cdots < x_n$.

We define arithmetic function $f(n)$ as follows:

$$f(n) = \sum_{a_1 b_1 = c_1} \sum_{a_2 b_2 = c_2} \cdots \sum_{a_r b_r = c_r} \frac{1}{{p_1}^{a_1} {p_2}^{a_2} \cdots {p_r}^{a_r}} \mu^{(e)}({p_1}^{b_1} {p_2}^{b_2} \cdots {p_r}^{b_r}), \tag{3.1}$$

where $n = {p_1}^{c_1} {p_2}^{c_2} \cdots {p_r}^{c_r}$.

3.1. Theorem. *Let $R = \{y_1, y_2, \ldots, y_m\}$ be an exponential closure of the set $T = \{x_1, x_2, \ldots, x_n\}$, where $y_1 < y_2 < y_3 < \cdots < y_m$ and $x_1 < x_2 < x_3 < \cdots < x_n$. Define the $n \times m$ matrix $C = (c_{ij})$ by*

$$c_{ij} = \begin{cases} 1, & y_j|_e x_i \\ 0, & otherwise \end{cases}$$

and the $m \times m$ diagonal matrix by

$$\Xi = diag(f(y_1), f(y_2), \ldots, f(y_m)).$$

Then,

$$\overline{T}_{(e)} = C \Xi C^t.$$

Proof. The ij^{th} entry of $C\Xi C^t$ is equal to

$$(C\Xi C^t)_{ij} = \sum_{k=1}^{n} c_{ik} f(y_k) c_{jk} = \sum_{y_k|_e x_i, y_k|_e x_j} f(y_k) = \sum_{y_k|_e (x_i, x_j)_{(e)}} f(y_k),$$

where f is defined in Equation 3.1. By Möbius Inversion Exponential formula,

$$\sum_{d|_e n} g(d) = \frac{1}{n}.$$

Finally we get,

$$(C\Xi C^t)_{ij} = \frac{1}{(x_i, x_j)_{(e)}}.$$

3.2. Theorem. *Let $R = \{y_1, y_2, \ldots, y_m\}$ be an exponential closure of the set $T = \{x_1, x_2, \ldots, x_n\}$, where $y_1 < y_2 < y_3 < \cdots < y_m$ and $x_1 < x_2 < x_3 < \cdots < x_n$. Then*

$$\det \overline{T}_{(e)} = \sum_{1 \le k_1 < k_2 < \cdots < k_n \le m} (\det C_{(k_1,k_2,\ldots,k_n)})^2 f(y_{k_1}) f(y_{k_2}) \ldots f(y_{k_n}),$$

where $C_{(k_1,k_2,\ldots,k_n)}$ is the sub matrix of C consisting of the ${k_1}^{th}, {k_2}^{th}, \ldots, {k_n}^{th}$ columns of C.

Proof. The proof can be done on similar lines as Theorem 2.2.

3.3. Corollary. *Let $T = \{x_1, x_2, \ldots, x_n\}$ be a finite ordered set of distinct positive integers. If $T = R$, then the determinant of Reciprocal GCED matrix $\overline{T}_{(e)}$ defined on T is given as:*

$$\det \overline{T}_{(e)} = \prod_{k=1}^{n} f(x_k).$$

Proof. Note that C is a lower triangular matrix with diagonal $(1, 1, \ldots, 1)_n$. This implies that $\det C = 1$. The result is further proved by using the fact that the determinant of a diagonal matrix is equal to the product of its diagonal entries.

3.4. Corollary. *If $\overline{T}_{(e)}$ is an $n \times n$ Reciprocal GCED matrix on a set $T = \{x_1, x_2, \ldots, x_n\}$, then the trace is given as:*

$$tr\overline{T}_{(e)} = \sum_{k=1}^{n} \frac{1}{x_i}.$$

3.5. Theorem. *Let $\overline{T}_{(e)}$ be an $n \times n$ Reciprocal GCED matrix on an exponential divisor closed set T. Then, its inverse matrix $\overline{A}_{(e)} = (a_{ij})_{(e)}$ is given as:*

$$(a_{ij})_{(e)} = \sum_{x_i|_{(e)} x_k, x_j|_{(e)} x_k} \frac{\mu^{(e)} \frac{x_d}{x_i} \mu^{(e)} \frac{x_d}{x_j}}{f(x_d)}.$$

Proof. Since $\overline{T}_{(e)} = (C\Xi C^t)$ and by Lemma 2.5, $C^{-1} = W$, therefore

$$(T)^{-1}_{(e)} = (C\Xi C^t)^{-1} = W^t \Xi^{-1} W,$$

where ij^{th} entry of $(T)^{-1}_{(e)}$ is given as

$$(a_{ij})_{(e)} = \sum_{x_i|_{(e)} x_k, x_j|_{(e)} x_k} \frac{\mu^{(e)} \frac{x_d}{x_i} \mu^{(e)} \frac{x_d}{x_j}}{f(x_d)}.$$

Hence, the required result.

4. Examples

4.1. Example. Let $T = \{12, 18, 36\}$. The GCED matrix $T_{(e)}$ on T is given as:

$$T_{(e)} = \begin{bmatrix} 12 & 6 & 12 \\ 6 & 18 & 18 \\ 12 & 18 & 36 \end{bmatrix}.$$

Note that $T = \{12, 18, 36\}$ is not an exponential divisor closed set. Its exponential closure is $R = \{6, 12, 18, 36\}$. The 3×4 matrix $(C)_{(e)}$ is

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

By Theorem 2.2, we know that,

$$\det T_{(e)} = \sum_{1 \le k_1 < k_2 < \ldots < k_n \le m} (\det C_{k_1,k_2,\ldots,k_n})^2 g(y_{k_1}) g(y_{k_2}) \ldots g(y_{k_n}).$$

So,

$$\det T_{(e)} = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{vmatrix}^2 g(6)g(12)g(18) + \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 g(6)g(12)g(36) +$$

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 g(6)g(18)g(36) + \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 g(12)g(18)g(36)$$

where, $g(6) = 6, g(12) = 6, g(18) = 12$ and $g(36) = 12$.
Hence, the determinant is given as:

$$\det T_{(e)} = (6)(6)(12) + (6)(6)(12) + (6)(12)(12) + (6)(12)(12) = 2592.$$

The Reciprocal GCED matrix $\overline{T}_{(e)}$ on T is given as:

$$\overline{T}_{(e)} = \begin{bmatrix} \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{6} & \frac{1}{18} & \frac{1}{18} \\ \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \end{bmatrix}.$$

By Theorem 3.2,

$$\det \overline{T}_{(e)} = \sum_{1 \le k_1 < k_2 < \ldots < k_n \le m} (\det C_{k_1,k_2,\ldots,k_n})^2 f(y_{k_1}) f(y_{k_2}) \ldots f(y_{k_n}).$$

So,

$$\det \overline{T}_{(e)} = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{vmatrix}^2 f(6)f(12)f(18) + \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 f(6)f(12)f(36) +$$

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 f(6)f(18)f(36) + \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix}^2 f(12)f(18)f(36),$$

where,
$f(6) = \frac{1}{6}\mu^{(e)}((2)(3)) = \frac{1}{6}$, $f(12) = \frac{1}{(2)(3)}\mu^{(e)}((2^2)(3)) + \frac{1}{(2^2)(3)}\mu^{(e)}((2)(3)) = \frac{-1}{12}$
$f(18) = \frac{1}{(2)(3)}\mu^{(e)}((3^2)(2)) + \frac{1}{(2)(3^2)}\mu^{(e)}((2)(3)) = \frac{-1}{9}$ and
$f(36) = \frac{1}{(2)(3)}\mu^{(e)}((2^2)(3^2)) + \frac{1}{(2^2)(3)}\mu^{(e)}((3^2)(2)) + \frac{1}{(2)(3^2)}\mu^{(e)}((2^2)(3)) + \frac{1}{(2^2)(3^2)}\mu^{(e)}((2)(3)) = \frac{1}{18}$. Hence,

$$\det \overline{T}_{(e)} = \frac{1}{3888}.$$

4.2. Example. Let $T = \{2, 4, 16\}$. This set is an exponential divisor closed, so we apply the Corollary to Theorem 2.2 directly to calculate the determinant. The GCED matrix defined on T is

$$T_{(e)} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 16 \end{bmatrix},$$

where, $g(2) = 2\mu^{(e)}(2) = 2, g(4) = 2\mu^{(e)}(2^2) + 2^2\mu^{(e)}(2) = 2$, and $g(16) = 2\mu^{(e)}(2^4) + 2^2\mu^{(e)}(2^2) + 2^4\mu^{(e)}(2) = 2(0) + 4(-1) + 16 = 12$. Thus,

$$\det T_{(e)} = \prod_{k=1}^{3} g(x_k) = g(2)g(4)g(16) = (2)(12)(12) = 48.$$

The Reciprocal GCED matrix $\overline{T}_{(e)}$ on T is given as

$$\overline{T}_{(e)} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{16} \end{bmatrix},$$

where, $f(2) = \frac{1}{2}, f(4) = \frac{-1}{4}$ and $f(16) = \frac{-3}{16}$. Thus,

$$\det \overline{T}_{(e)} = \prod_{k=1}^{3} f(x_k) = (\frac{1}{2})(\frac{-1}{4})(\frac{-3}{16}) = \frac{3}{128}.$$

4.3. Example. Let $T = \{2, 4, 16\}$. The 3×3 GCED matrix $T_{(e)}$ defined on T is

$$T_{(e)} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 16 \end{bmatrix}.$$

By Theorem 2.6, we know that ${(T)^{-1}}_{(e)} = (a_{ij})$ where,
$a_{11} = \sum_{2|_e x_k} \frac{\mu^{(e)}(\frac{x_k}{2})\mu^{(e)}(\frac{x_k}{2})}{g(x_k)} = \frac{\mu^{(e)}(2^2)\mu^{(e)}(2^2)}{g(2)} + \frac{\mu^{(e)}(2^2)\mu^{(e)}(2^2)}{g(4)} + \frac{\mu^{(e)}(2^4)\mu^{(e)}(2^4)}{g(16)} = 1,$
$a_{12} = \frac{\mu^{(e)}(2^2)\mu^{(e)}(2)}{g(4)} + \frac{\mu^{(e)}(2^4)\mu^{(e)}(2^2)}{g(16)} = \frac{-1}{2}$, and $a_{13} = \frac{\mu^{(e)}(2^4)\mu^{(e)}(2)}{g(16)} = 0.$
Similarly, one can calculate and verify the following values
$a_{22} = \frac{7}{12}, a_{23} = \frac{-1}{12}$ and $a_{33} = \frac{1}{12}$. So, the inverse of the GCED matrix $T_{(e)}$ is

$${(T)^{-1}}_{(e)} = \begin{bmatrix} 1 & \frac{-1}{2} & 0 \\ \frac{-1}{2} & \frac{7}{12} & \frac{-1}{12} \\ 0 & \frac{-1}{12} & \frac{1}{12} \end{bmatrix}.$$

The 3×3 Reciprocal GCED matrix on T is given as

$$\overline{T}_{(e)} = \begin{bmatrix} \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{6} & \frac{1}{18} & \frac{1}{18} \\ \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \end{bmatrix}.$$

The inverse of the Reciprocal GCED matrix $\overline{T}_{(e)}$ is calculated to be

$$(\overline{T})^{-1}_{(e)} = \begin{bmatrix} -2 & 4 & 0 \\ 4 & \frac{-28}{3} & \frac{16}{3} \\ 0 & \frac{16}{3} & \frac{-16}{3} \end{bmatrix}.$$

References

[1] E. Altinisk, D. Tasci, On a generalization of LCM matrix, Common. Fac. Sci. Univ. Ank. Series A1, **51**(No. 2) (2002), 37-46.

[2] S.J. Beslin, S. Ligh, Greatest common divisor matrices, Linear Algebra Appl., **118** (1989), 69-76.

[3] S. Beslin, Reciprocal GCD matrices and LCM matrices, Fibonacci Quart. **29** (1991), 271-274.

[4] W. Feng, S. Hong, J. Zhao, Divisibilty of power LCM matrices by power GCD matrices on gcd-closed sets, Discrete Mathematics, **309** (2009), 2627-2639.

[5] A. Nalli, D. Tasci, The GCD-Reciprocal LCM matrices on gcd closed sets, Mathematical and Comp. Appl., **9** (2004), 101-106.

[6] B. V. Rajarama Bhat, On greatest common divisor matrices and their applications, Linear Algebra Appl., **158** (1991), 77-97.

[7] H. J. Smith, On the value of a certain arithmetical determinant, Proc. London Math Soc., **7** (1875/1876), 208-212.

[8] M. V. Subrarao, On some arithmetical convolutions, The theory of arithmetic functions, Lecture notes in Mathematics, Springer, No. 251, (1972), 247-271.

[9] N. Tuglu, D. Tasci, On the LCUM-Reciprocal GCUD matrices, Competes rendus de l' academie bulgare des Sciences, **55** (No.12) (2002), 17-21.

[10] L. Tòth, On certain arithmetical functions involving exponential divisors, Annales Univ. Sci. Budapest, **24**(2004), 285-294.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 641 – 649

Approximation of generalized left derivations in modular spaces

Tayebe Lal Shateri *

Abstract

In this paper, we define modular spaces, and introduce some properties of them. Moreover, we present a fixed point method to prove superstability of generalized left derivations from an algebra into a modular space.

2000 AMS Classification: Primary 54H25; Secondary 47H10, 47B48.

Keywords: superstability, left derivation, generalized left derivation, fixed point, modular space.

Received 13/12/2013 : Accepted 21/02/2014 Doi : 10.15672/HJMS.2015449420

1. Introduction

Let $\mathcal{A}$ be an algebra over the real or complex field $\mathbb{F}$ and let $\mathcal{X}$ be an $\mathcal{A}$-module. An additive mapping $d : \mathcal{A} \to \mathcal{X}$ is said to be a left derivation if the functional equation $d(xy) = xd(y) + yd(x)$ holds for all $x, y \in \mathcal{A}$. Moreover, if $d(\alpha x) = \alpha d(x)$ is valid for all $x \in \mathcal{A}$ and for all $\alpha \in \mathbb{F}$, then d is called a linear left derivation. An additive mapping $D : \mathcal{A} \to \mathcal{X}$ is said to be a generalized left derivation if there exists a left derivation $d : \mathcal{A} \to \mathcal{X}$ such that $D(xy) = xD(y) + yd(x)$ holds for all $x, y \in \mathcal{A}$. Furthermore, if $D(\alpha x) = \alpha D(x)$ is valid for all $x \in \mathcal{A}$ and for all $\alpha \in \mathbb{F}$, then D is called a linear generalized left derivation.

In 1940, Ulam [21] posed the first stability problem of functional equations, concerning the stability of group homomorphisms, was solved in the case of the additive mapping by Hyers [4] in the next year. Subsequently, Aoki [1] extended Hyers' theorem for approximately additive mappings and for approximately linear mappings was presented by Rassias [18]. The stability result concerning derivations between operator algebras was first obtained by Semrl [20]. Also Badora [2] present the Hyers-Ulam stability and the superstability of derivations. The equation is called *superstable* if each its approximate solution is an exact solution. Various stability and superstability results for derivations

*Department of Mathematics and Computer Sciences, Hakim Sabzevari University, Sabzevar, P.O. Box 397, IRAN
Email: t.shateri@gmail.com, t.shateri@hsu.ac.ir

have been investigated by a number of mathematicians [3, 5, 11, 12, 16, 17, 19]. In this paper, we define modular spaces, and introduce some properties of them. Moreover, we prove the superstability of generalized left derivations from an algebra with unit into a modular space by using a fixed point method. The theory of modular spaces were founded by Nakano [14] and were intensively developed by Luxemburg [9], Koshi and Shimogaki [7] and Yamamuro [22] and their collaborators. In the present time the theory of modulars and modular spaces is extensively applied, in particular, in the study of various Orlicz spaces [15] and interpolation theory [8, 10], which in their turn have broad applications [13].

1.1. Definition. Let $\mathcal{X}$ be an arbitrary vector space.
(a) A functional $\rho : \mathcal{X} \to [0, \infty]$ is called a modular if for arbitrary $x, y \in \mathcal{X}$,
(i) $\rho(x) = 0$ if and only if $x = 0$,
(ii) $\rho(\alpha x) = \rho(x)$ for every scaler α with $|\alpha| = 1$,
(iii) $\rho(\alpha x + \beta y) \leq \rho(x) + \rho(y)$ if and only if $\alpha + \beta = 1$ and $\alpha, \beta \geq 0$,
(b) if (iii) is replaced by
(iii)$'$ $\rho(\alpha x + \beta y) \leq \alpha\rho(x) + \beta\rho(y)$ if and only if $\alpha + \beta = 1$ and $\alpha, \beta \geq 0$,
then we say that ρ is a convex modular.

If ρ is a modular, the corresponding modular space is the vector space $\mathcal{X}_\rho$ given by

$$\mathcal{X}_\rho = \{x \in \mathcal{X} : \quad \rho(\lambda x) \to 0 \text{ as } \lambda \to 0\}.$$

Let ρ be a convex modular, the modular space $\mathcal{X}_\rho$ can be equipped with a norm called the Luxemburg norm, defined by

$$\|x\|_\rho = \inf\left\{\lambda > 0 : \quad \rho\left(\frac{x}{\lambda}\right) \leq 1\right\}.$$

A function modular is said to be satisfy the Δ_2–condition if there exists $\kappa > 0$ such that $\rho(2x) \leq \kappa\rho(x)$ for all $x \in \mathcal{X}_\rho$.

1.2. Definition. Let $\{x_n\}$ and x be in $\mathcal{X}_\rho$. Then
(i) the sequence $\{x_n\}$, with $x_n \in \mathcal{X}_\rho$, is ρ–convergent to x and we write $x_n \xrightarrow{\rho} x$ if $\rho(x_n - x) \to 0$ as $n \to \infty$.
(ii) The sequence $\{x_n\}$, with $x_n \in \mathcal{X}_\rho$, is called ρ–Cauchy if $\rho(x_n - x_m) \to 0$ as $n, m \to \infty$.
(iii) A subset $\mathcal{S}$ of $\mathcal{X}_\rho$ is called ρ–complete if and only if any ρ–Cauchy sequence is ρ–convergent to an element of $\mathcal{S}$.

We call the modular ρ has the Fatou property if $\rho(x) \leq \liminf_{n\to\infty} \rho(x_n)$ whenever the sequence $\{x_n\}$ is ρ–convergent to x.

1.3. Remark. Note that $\rho(.x)$ is an increasing function for each $x \in \mathcal{X}$. Suppose $0 < a < b$, and put $y = 0$ in property (iii) of Definition 1.1, then $\rho(ax) = \rho\left(\frac{a}{b}bx\right) \leq \rho(bx)$ for all $x \in \mathcal{X}$. Moreover, if ρ is a convex modular on $\mathcal{X}$ and $|\alpha| \leq 1$, then $\rho(\alpha x) \leq \alpha\rho(x)$ and also $\rho(x) \leq \frac{1}{2}\rho(2x)$ for all $x \in \mathcal{X}$.

1.4. Example. An example of a modular space with Δ_2–condition is the Orlicz space. Let τ be a function defined on the interval $[0, \infty)$ such that $\tau(0) = 0, \tau(\alpha) > 0$ for $\alpha > 0$ and $\tau(\alpha) \to \infty$ as $\alpha \to \infty$. Also assume that τ is convex, nondecreasing and continuous. The function τ is called an Orlicz function. The Orlicz function τ satisfies the Δ_2–condition if there exists $\kappa > 0$ such that $\tau(2\alpha) \leq \kappa\tau(\alpha)$ for all $\alpha > 0$. Let $(\Omega, \mathfrak{M}, \mu)$ be a measure space. Let $L^0(\mu)$ be the space of all measurable real–valued (or complex–valued) functions on Ω. For every $f \in L^0(\mu)$, we define the Orlicz modular $\rho_\tau(f)$ as

$$\rho_\tau(f) = \int_\Omega \tau(|f|)d\mu.$$

The associated modular function space with respect to this modular is called an Orlicz space, and will be denoted by $L^\tau(\Omega,\mu)$ or briefly L^τ. In other words,

$$L^\tau = \{f \in L^0(\mu) | \quad \rho_\tau(\lambda f) \to 0 \text{ as } \lambda \to 0\}$$

or equivalently as

$$L^\tau = \{f \in L^0(\mu) | \quad \rho_\tau(\lambda f) < \infty \text{ for some } \lambda > 0\}.$$

It is known that the Orlicz space L^τ is ρ_τ–complete. Moreover, $(L^\tau, \|.\|_{\rho_\tau})$ is a Banach space, where the Luxemburg norm $\|.\|_{\rho_\tau}$ is defined as follows

$$\|f\|_{\rho_\tau} = \inf\left\{\lambda > 0 : \quad \int_\Omega \tau\left(\frac{|f|}{\lambda}\right) d\mu \le 1\right\}.$$

2. Main results

Throughout this paper, $\mathcal{A}$ and $\mathcal{X}$ denote a Banach algebra with unit and a unital $\mathcal{A}$-module respectively. Also $\mathcal{X}_\rho$ denotes a ρ–complete modular space where ρ is a convex modular on $\mathcal{X}$ with the Fatou property such that satisfies the Δ_2–condition with $0 < \kappa \le 2$. In this section, we present the superstability of generalized left derivations from a Banach algebra into a complete modular space.

2.1. Theorem. *Let $d : \mathcal{A} \to \mathcal{X}_\rho$ be a mapping with $d(0) = 0$ such that*

$$\rho\left(d(x+y) - d(x) - d(y)\right) \le \varphi(x,y) \tag{2.1}$$

for all $x, y \in \mathcal{A}$, where $\varphi : \mathcal{A} \times \mathcal{A} \to [0,\infty)$ is a given mapping that

$$\varphi(2x, 2x) \le 2L\varphi(x,x)$$

and

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y)}{2^n} = 0 \tag{2.2}$$

for all $x, y \in \mathcal{A}$ and a constant $0 < L < 1$. Then there exist a unique additive mapping $D : \mathcal{A} \to \mathcal{X}_\rho$ and a convex modular function $\widetilde{\rho}$ such that

$$\widetilde{\rho}(D - d) \le \frac{1}{2(1-L)}. \tag{2.3}$$

Proof. Consider the set

$$\mathfrak{B} = \{\delta : \mathcal{A} \to \mathcal{X}_\rho, \quad \delta(0) = 0\}$$

we define the function $\widetilde{\rho}$ on $\mathfrak{B}$ as follows,

$$\widetilde{\rho}(\delta) = \inf\{c > 0 : \rho(\delta(x)) \le c\varphi(x,x)\}. \tag{2.4}$$

Then $\widetilde{\rho}$ is convex modular. It is enough to show that $\widetilde{\rho}$ satisfies the following condition

$$\widetilde{\rho}(\alpha\delta + \beta\gamma) \le \alpha\widetilde{\rho}(\delta) + \beta\widetilde{\rho}(\gamma) \quad (\alpha, \beta \ge 0,\ \alpha + \beta = 1).$$

Given $\varepsilon > 0$, then there exist $c_1 > 0$ and $c_2 > 0$ such that

$$c_1 \le \widetilde{\rho}(\delta) + \varepsilon, \quad \rho(\delta(x)) \le c_1\varphi(x,x)$$

and

$$c_2 \le \widetilde{\rho}(\gamma) + \varepsilon, \quad \rho(\gamma(x)) \le c_2\varphi(x,x).$$

For $\alpha, \beta \ge 0$ such that $\alpha + \beta = 1$, we get

$$\rho(\alpha\delta(x) + \beta\gamma(x)) \le \alpha\rho(\delta(x)) + \beta\rho(\gamma(x)) \le (\alpha c_1 + \beta c_2)\varphi(x,x),$$

hence

$$\widetilde{\rho}(\alpha\delta + \beta\gamma) \le \alpha\widetilde{\rho}(\delta) + \beta\widetilde{\rho}(\gamma) + (\alpha + \beta)\varepsilon.$$

Consequently $\widetilde{\rho}(\alpha\delta+\beta\gamma)\leq\alpha\widetilde{\rho}(\delta)+\beta\widetilde{\rho}(\gamma)$. Moreover, $\widetilde{\rho}$ satisfies the Δ_2–condition with $0<\kappa<2$. For this, let $\{\delta_n\}$ be a $\widetilde{\rho}$–Cauchy sequence in $\mathcal{E}_{\widetilde{\rho}}$ and given $\varepsilon>0$. There exists a positive integer $n_0\in\mathbb{N}$ such that $\widetilde{\rho}(\delta_n-\delta_m)\leq\varepsilon$ for all $n,m\geq n_0$. Then by definition of the modular $\widetilde{\rho}$, we have

$$\rho\left(\delta_n(x)-\delta_m(x)\right)\leq\varepsilon\varphi(x,x) \tag{2.5}$$

for all $x\in\mathcal{A}$ and $n,m\geq n_0$. Let x be a point of $\mathcal{A}$, (2.5) implies that $\{\delta_n(x)\}$ is a ρ–Cauchy sequence in $\mathcal{X}_\rho$. Since $\mathcal{X}_\rho$ is ρ–complete, so $\{\delta_n(x)\}$ is ρ–convergent in $\mathcal{X}_\rho$, for each $x\in\mathcal{A}$. Therefore we can define a function $\delta:\mathcal{A}\to\mathcal{X}_\rho$ by

$$\delta(x)=\lim_{n\to\infty}\delta_n(x)$$

for any $x\in\mathcal{A}$. Letting $m\to\infty$, then (2.5) implies that

$$\widetilde{\rho}(\delta_n-\delta)\leq\varepsilon$$

for all $n\geq n_0$. Since ρ has the Fatou property, thus $\{\delta_n\}$ is $\widetilde{\rho}$–convergent sequence in $\mathfrak{B}_{\widetilde{\rho}}$. Therefore $\mathcal{E}_{\widetilde{\rho}}$ is $\widetilde{\rho}$–complete.

Now, we define the function $\mathcal{T}:\mathcal{E}_{\widetilde{\rho}}\to\mathfrak{B}_{\widetilde{\rho}}$ as follows

$$\mathcal{T}\delta(x):=\frac{1}{2}\delta(2x)$$

for all $\delta\in\mathfrak{B}_{\widetilde{\rho}}$. Let $\delta,\gamma\in\mathfrak{B}_{\widetilde{\rho}}$ and let $c\in[0,\infty]$ be an arbitrary constant with $\widetilde{\rho}(\delta-\gamma)\leq c$. We have

$$\rho(\delta(x)-\gamma(x))\leq c\varphi(x,x)$$

for all $x\in\mathcal{A}$. The last inequality implies that

$$\rho\left(\frac{\delta(2x)}{2}-\frac{\gamma(2x)}{2}\right)\leq\frac{1}{2}\rho(\delta(2x)-\gamma(2x))\leq\frac{1}{2}c\varphi(2x,2x)\leq Lc\varphi(x,x)$$

for all $x\in\mathcal{A}$. Hence, $\widetilde{\rho}(\mathcal{T}\delta-\mathcal{T}\gamma)\leq L\widetilde{\rho}(\delta-\gamma)$, for all $\delta,\gamma\in\mathfrak{B}_{\widetilde{\rho}}$. Therefore $\mathcal{T}$ is a $\widetilde{\rho}$–strict contraction. We show that the $\widetilde{\rho}$–strict mapping $\mathcal{T}$ satisfies the conditions of Theorem 3.4 of [6]. Letting $x=y$ in (2.12), we get

$$\rho(d(2x)-2d(x))\leq\varphi(x,x) \tag{2.6}$$

for all $x\in\mathcal{A}$. Replacing x by $2x$ in (2.6) we get

$$\rho(d(4x)-2d(2x))\leq\varphi(2x,2x)$$

for all $x\in\mathcal{A}$. Since ρ is convex modular and satisfies the Δ_2–condition, for all $x\in\mathcal{A}$ we have

$$\begin{aligned}\rho\left(\frac{d(4x)}{2}-2d(x)\right)&\leq\frac{1}{2}\rho(d(4x)-2d(2x))+\frac{1}{2}\rho(2d(2x)-4d(x))\\&\leq\frac{1}{2}\varphi(2x,2x)+\frac{\kappa}{2}\varphi(x,x).\end{aligned}$$

Moreover,

$$\rho\left(\frac{d(2^2x)}{2^2}-d(x)\right)\leq\frac{1}{2}\rho\left(2\frac{d(4x)}{2^2}-2d(x)\right)\leq\frac{1}{2^2}\varphi(2x,2x)+\frac{\kappa}{2^2}\varphi(x,x).$$

for all $x\in\mathcal{A}$. By induction we obtain

$$\rho\left(\frac{d(2^nx)}{2^n}-d(x)\right)\leq\frac{1}{2^n}\sum_{i=1}^{n}\kappa^{n-i}\varphi(2^{i-1}x,2^{i-1}x)\leq\frac{1}{2(1-L)}\varphi(x,x) \tag{2.7}$$

for all $x \in \mathcal{A}$. Now we claim that $\delta_{\widetilde{\rho}}(d) = \sup\{\widetilde{\rho}\,(\mathcal{T}^n(d) - \mathcal{T}^m(d))\,; n, m \in \mathbb{N})\} < \infty$. It follows from (2.7) that

$$\begin{aligned}\rho\left(\frac{d(2^n x)}{2^n} - \frac{d(2^m x)}{2^m}\right) &\leq \frac{1}{2}\rho\left(2\frac{d(2^n x)}{2^n} - 2d(x)\right) + \frac{1}{2}\rho\left(2\frac{d(2^m x)}{2^m} - 2d(x)\right)\\ &\leq \frac{\kappa}{2}\rho\left(\frac{d(2^n x)}{2^n} - d(x)\right) + \frac{\kappa}{2}\rho\left(\frac{d(2^m x)}{2^m} - d(x)\right)\\ &\leq \frac{1}{1-L}\varphi(x,x),\end{aligned}$$

for every $x \in \mathcal{A}$ and $n, m \in \mathbb{N}$, which implies that

$$\widetilde{\rho}\,(\mathcal{T}^n(d) - \mathcal{T}^m(d)) \leq \frac{1}{1-L},$$

for all $n, m \in \mathbb{N}$. Therefore $\delta_{\widetilde{\rho}}(d) < \infty$. [6, Lemma 3.3] shows that $\{\mathcal{T}^n(d)\}$ is $\widetilde{\rho}$–convergent to $D \in \mathfrak{B}_{\widetilde{\rho}}$. Since ρ has the Fatou property, (2.7) gives $\widetilde{\rho}(\mathcal{T}D - d) < \infty$.

If we replace x by $2^n x$ in (2.6), then

$$\widetilde{\rho}\left(d(2^{n+1}x) - 2d(2^n x)\right) \leq \varphi(2^n x, 2^n x),$$

for all $x \in \mathcal{A}$. Hence

$$\begin{aligned}\rho\left(\frac{d(2^{n+1}x)}{2^{n+1}} - \frac{d(2^n x)}{2^n}\right) &\leq \frac{1}{2^{n+1}}\rho\left(d(2^{n+1}x) - 2d(2^n x)\right) \leq \frac{1}{2^{n+1}}\varphi(2^n, 2^n x)\\ &\leq \frac{1}{2^{n+1}}2^n L^n \varphi(x,x) \leq \frac{L^n}{2}\varphi(x,x) \leq \varphi(x,x)\end{aligned}$$

for all $x \in \mathcal{A}$, therefore $\widetilde{\rho}(\mathcal{T}(D) - D) < \infty$. It follows from [6, Theorem 3.4] that $\widetilde{\rho}$–limit D of $\{\mathcal{T}^n(d)\}$ is fixed point of map $\mathcal{T}$. If we replace x by $2^n x$ and y by $2^n y$ in (2.12), then we obtain

$$\rho\left(d(2^n(x+y)) - d(2^n x) - d(2^n y)\right) \leq \varphi(2^n x, 2^n y)$$

for all $x, y \in \mathcal{A}$. Hence,

$$\begin{aligned}\rho\left(\frac{d(2^n(x+y))}{2^n} - \frac{d(2^n x)}{2^n} - \frac{d(2^n y)}{2^n}\right) &\leq \frac{1}{2^n}\rho\left(d(2^n(x+y)) - d(2^n x) - d(2^n y)\right)\\ &\leq \frac{\varphi(2^n x, 2^n y)}{2^n}\end{aligned}$$

for all $x, y \in \mathcal{A}$. Taking the limit, we deduce that $D(x+y) = D(x) + D(y)$ for all $x, y \in \mathcal{A}$, that is, D is additive. Now, let D^* be another fixed point of $\mathcal{T}$, then

$$\begin{aligned}\widetilde{\rho}(D - D^*) &\leq \frac{1}{2}\widetilde{\rho}(2\mathcal{T}(D) - 2d) + \frac{1}{2}\widetilde{\rho}(2\mathcal{T}(D^*) - 2d)\\ &\leq \frac{\kappa}{2}\widetilde{\rho}(\mathcal{T}(D) - d) + \frac{\kappa}{2}\widetilde{\rho}(\mathcal{T}(D^*) - d) \leq \frac{\kappa}{2(1-L)} < \infty.\end{aligned}$$

Since $\mathcal{T}$ is $\widetilde{\rho}$–strict contraction, we get

$$\widetilde{\rho}(D - D^*) = \widetilde{\rho}(\mathcal{T}(D) - \mathcal{T}(D^*)) \leq L\widetilde{\rho}(D - D^*),$$

which implies that $\widetilde{\rho}(D - D^*) = 0$ or $D = D^*$, since $\widetilde{\rho}(D - D^*) < \infty$. This proves the uniqueness of D. Also it follows from inequality (2.7) that

$$\widetilde{\rho}(D - d) \leq \frac{1}{2(1-L)}.$$

This completes the proof. □

We now investigate the superstability of a generalized left derivation from a unital algebra into a modular space.

2.2. Theorem. *Let $d : \mathcal{A} \to \mathcal{X}_\rho$ be a mapping with $d(0) = 0$. If there exists a mapping $g : \mathcal{A} \to \mathcal{X}_\rho$ such that*

$$\rho\left(d(x + y + zw) - d(x) - d(y) - zd(w) - wg(z)\right) \leq \varphi(x, y, z, w) \tag{2.8}$$

for all $x, y, z, w \in \mathcal{A}$, where $\varphi : \mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \mathcal{A} \to [0, \infty)$ is a given mapping such that

$$\varphi(2x, 2x, 0, 0) \leq 2L\varphi(x, x, 0, 0)$$

and

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y, 0, 0)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0, 0, 2^n z, w)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0, 0, z, 2^n w)}{2^n} = 0 \tag{2.9}$$

for all $x, y \in \mathcal{A}$ and a constant $0 < L < 1$, then d is a generalized left derivation and g is a left derivation.

Proof. Letting $z = w = 0$ in (2.8), then d satisfies (2.12) and so the Theorem 2.1 shows that there exists a unique additive mapping $D : \mathcal{A} \to \mathcal{X}_\rho$ for which satisfies

$$\widetilde{\rho}(D - d) \leq \frac{1}{2(1 - L)},$$

where $\widetilde{\rho}$ is the convex modular defined in (2.4). Now, we prove that d is a generalized left derivation and g is a left derivation. Substituting $x = y = 0$ in (2.8), we get

$$\rho\left(d(zw) - zd(w) - wg(z)\right) \leq \varphi(0, 0, z, w), \tag{2.10}$$

for all $z, w \in \mathcal{A}$. Moreover, if we replace z and w with $2^n z$ and $2^n w$ in (2.10), respectively, and then divide both sides by 2^{2n}, we deduced that

$$\rho\left(\frac{d(2^{2n} zw)}{2^{2n}} - z\frac{d(2^n w)}{2^n} - w\frac{g(2^n z)}{2^n}\right) \leq \frac{\varphi(0, 0, 2^n z, 2^n w)}{2^{2n}},$$

for all $z, w \in \mathcal{A}$. Letting $n \to \infty$, we obtain

$$D(zw) - zD(w) = \lim_{n\to\infty} w\frac{g(2^n z)}{2^n},$$

for all $z, w \in \mathcal{A}$. Suppose that $w = e$, hence it follows

$$\lim_{n\to\infty} \frac{g(2^n z)}{2^n} = D(z) - zD(e),$$

for all $z \in \mathcal{A}$. If $\gamma(z) = D(z) - zD(e)$, then by the additivity of D, we get

$$\gamma(z+w) = D(z+w)-(z+w)D(e) = (D(z)-zD(e))+(D(w)-wD(e)) = \gamma(z)+\gamma(w),$$

for all $z, w \in \mathcal{A}$. Therefore γ is additive.

Suppose $\Delta(z, w) = d(zw) - zd(w) - wg(z)$, for all $z, w \in \mathcal{A}$. The inequality given in (2.10) implies that

$$\lim_{n\to\infty} \frac{\Delta(2^n z, w)}{2^n} = 0,$$

for all $z, w \in \mathcal{A}$. Thus we get

$$\begin{aligned} D(zw) = \widetilde{\rho} \lim_{n\to\infty} \frac{d(2^{2n} zw)}{2^n} &= \lim_{n\to\infty} \frac{2^n zd(w) + wg(2^n z) + \Delta(2^n z, w)}{2^n} \\ &= zd(w) + \lim_{n\to\infty} \frac{wg(2^n z)}{2^n} = zd(w) + w\gamma(z), \end{aligned}$$

for all $z, w \in \mathcal{A}$. Since γ is additive, we have

$$2^n zd(w) + 2^n w\gamma(z) = D(2^n z.w) = D(z.2^n w) = zd(2^n w) + 2^n w\gamma(z),$$

for all $z, w \in \mathcal{A}$. Therefore $zd(w) = z\frac{1}{2^n}d(2^n w)$, for all $z, w \in \mathcal{A}$. By letting $n \to \infty$, we obtain $zd(w) = zD(w)$. If $z = e$, we have $d = D$. Consequently we get

$$d(zw) = zd(w) + w\gamma(z), \tag{2.11}$$

for all $z, w \in \mathcal{A}$. Now, we verify that γ is a left derivation. Using the fact that d satisfies (2.11), we have

$$\begin{aligned}\gamma(xy) &= d(xy) - xyd(e) = xd(y) + y\gamma(x) - xyd(e)\\ &= x(d(y) - yd(e)) + y\gamma(x) = x\gamma(y) + y\gamma(x),\end{aligned}$$

for all $x, y \in \mathcal{A}$, which means that γ is a derivation and hence d is a generalized left derivation.

Finally, we show that g is a left derivation. If we replace w by $2^n w$ in (2.10) and then divide both sides by 2^{2n}, we obtain

$$\rho\left(\frac{d(2^n zw)}{2^n} - z\frac{d(2^n w)}{2^n} - 2^n w\frac{g(z)}{2^n}\right) \le \frac{\varphi(0,0,2^n z,w)}{2^n},$$

for all $z, w \in \mathcal{A}$. Passing the limit as $n \to \infty$, we get

$$d(zw) - zd(w) - wg(z) = 0,$$

for all $z, w \in \mathcal{A}$. Therefore $d(zw) = zd(w) + wg(z)$, for all $z, w \in \mathcal{A}$, and hence if $w = e$, then $g(z) = d(z) - zd(e) = \gamma(z)$, for all $z \in \mathcal{A}$. Since γ is a left derivation, hence g is a left derivation and this completes the proof. □

The similar way as in the proof of Theorem 2.2, we get the following result for a generalized derivation.

2.3. Theorem. *Let $d : \mathcal{A} \to \mathcal{X}_\rho$ be a mapping with $d(0) = 0$. If there exists a mapping $g : \mathcal{A} \to \mathcal{X}_\rho$ such that*

$$\rho\left(d(x + y + zw) - d(x) - d(y) - zd(w) - g(z)w\right) \le \varphi(x, y, z, w) \tag{2.12}$$

for all $x, y, z, w \in \mathcal{A}$, where $\varphi : \mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \mathcal{A} \to [0, \infty)$ is a given mapping such that

$$\varphi(2x, 2x, 0, 0) \le 2L\varphi(x, x, 0, 0)$$

and

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y, 0, 0)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0,0,2^n z, w)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0,0,z,2^n w)}{2^n} = 0 \tag{2.13}$$

for all $x, y \in \mathcal{A}$ and a constant $0 < L < 1$, then d is a generalized derivation and g is a derivation.

With the help of Theorem 2.1, the following result can be derived for a linear generalized left derivation.

2.4. Theorem. *Let $\mathcal{A}$ be a unital algebra and let $\mathcal{X}$ be a unital $\mathcal{A}$-module and $\mathcal{X}_\rho$ a ρ–complete modular space. Suppose $d : \mathcal{A} \to \mathcal{X}_\rho$ satisfies the condition $d(0) = 0$ and an inequality of the form*

$$\rho\left(d(\alpha x + \beta y + zw) - \alpha d(x) - \beta d(y) - zd(w) - wg(z)\right) \le \varphi(x, y, z, w) \tag{2.14}$$

for all $x, y, z, w \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$, where $\varphi : \mathcal{A}\times\mathcal{A}\times\mathcal{A}\times\mathcal{A} \to [0, \infty)$ is a given mapping such that

$$\varphi(2x, 2x, 0, 0) \le 2L\varphi(x, x, 0, 0)$$

and

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y, 0, 0)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0,0,2^n z, w)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0,0,z,2^n w)}{2^n} = 0 \tag{2.15}$$

for all $x, y \in \mathcal{A}$ and a constant $0 < L < 1$. Then d is a linear generalized left derivation and g is a linear left derivation.

Proof. We consider $\alpha = \beta = 1 \in \mathbb{U}$ in (2.14) and then d satisfies the inequality (2.8). It follows from Theorem 2.3 that d is a generalized left derivation and g is a left derivation. It is enough to prove that d and g are linear. By the proof of Theorem 2.2 we know that

$$d(x) = \widetilde{\rho} - \lim_{n\to\infty} \mathfrak{T}^n(d)(x) = \widetilde{\rho} - \lim_{n\to\infty} \frac{1}{2^n} d(2^n x). \tag{2.16}$$

Letting $w = 0$ in (2.14), we have

$$\rho\left(d(\alpha x + \beta y) - \alpha d(x) - \beta d(y)\right) \leq \varphi(x, y, 0, 0), \tag{2.17}$$

for all $x, y \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{U}$. If we replace x and y with $2^n x$ and $2^n y$ in (2.16), respectively, and then divide both sides by 2^n, we see that

$$\rho\left(\frac{1}{2^n} d(\alpha 2^n x + \beta 2^n y) - \frac{1}{2^n}\alpha d(2^n x) - \frac{1}{2^n}\beta d(2^n y)\right) \leq \frac{1}{2^n}\varphi(2^n x, 2^n y, 0, 0) \to 0, \tag{2.18}$$

for all $x, y \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{U}$, as $n \to \infty$. Hence, we get

$$d(\alpha x + \beta y) = \alpha d(x) + \beta d(y), \tag{2.19}$$

for all $x, y \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{U}$. Now the proof of [5, Theorem 2.3] implies that

$$d(\alpha x + \beta y) = \alpha d(x) + \beta d(y), \tag{2.20}$$

for all $x, y \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{C}$. □

Employing the similar way as in the proof of Theorem 2.3 and Theorem 2.4, we get the next corollary for a linear generalized derivation.

2.5. Corollary. *Let $\mathcal{A}$ be a unital algebra and let $\mathcal{X}$ be a unital $\mathcal{A}$-module and $\mathcal{X}_\rho$ a ρ–complete modular space. Suppose $d : \mathcal{A} \to \mathcal{X}_\rho$ satisfies the condition $d(0) = 0$ and an inequality of the form*

$$\rho\left(d(\alpha x + \beta y + zw) - \alpha d(x) - \beta d(y) - zd(w) - g(z)w\right) \leq \varphi(x, y, z, w) \tag{2.21}$$

for all $x, y, z, w \in \mathcal{A}$ and all $\alpha, \beta \in \mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$, where $\varphi : \mathcal{A}\times\mathcal{A}\times\mathcal{A}\times\mathcal{A} \to [0, \infty)$ is a given mapping such that

$$\varphi(2x, 2x, 0, 0) \leq 2L\varphi(x, x, 0, 0)$$

and

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y, 0, 0)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0, 0, 2^n z, w)}{2^n} = \lim_{n\to\infty} \frac{\varphi(0, 0, z, 2^n w)}{2^n} = 0 \tag{2.22}$$

for all $x, y \in \mathcal{A}$ and a constant $0 < L < 1$. Then d is a linear generalized derivation and g is a linear derivation.

2.6. Remark. Let $\mathcal{A}$ be a normed algebra and let $\mathfrak{B}$ be a Banach algebra. It is known that every normed space is modular space with the modular $\rho(x) = \|x\|$ and $\kappa = 2$. A typical example of φ in the above results is $\varphi(x, y) = \varepsilon + \theta(\|x\|^p + \|y\|^p + \|z\|^p + \|w\|^p)$, such that $\varepsilon, \theta \geq 0$ and $p \in [0, 1)$.

References

- [1] T. Aoki, *On the stability of the linear transformation in Banach spaces*, J. Math. Soc. Japan **2**, 64–66, 1950.
- [2] R. Badora, *On approximate derivations*, Math. Ineq. Appl. **9** (1), 167–173, 2006.
- [3] M. BreŠar and J. Vukman, *On left derivations and related mappings* , Proc. Amer. Math. Soc. **110** (1), 7–16, 1990.
- [4] D. H. Hyers, *On the stability of the linear functional equation*, Proc. Nat. Acad. Sci. U.S.A. **27**, 222–224, 1941.
- [5] S.-Y. Kang and I.-S. Chang, *Approximation of generalized left derivations*, J. Abst. Appl. Anal. **2008**, 1–8, 2008.
- [6] M. A. Khamsi, *Quasicontraction Mapping in modular spaces without* Δ_2*–condition*, Fixed Point Theory and Applications, **2008**, Artical ID 916187, 2008.
- [7] S. Koshi, T. Shimogaki, *On F–norms of quasi–modular spaces* , J. Fac. Sci. Hokkaido Univ. Ser. I **15** (3), 202–218, 1961.
- [8] M. Krbec, *Modular interpolation spaces*, Z. Anal. Anwendungen, **1**, 25–40, 1982.
- [9] W. A. Luxemburg, *Banach function spaces*, Ph. D. thesis, Delft Univrsity of technology, Delft, The Netherlands, 1959.
- [10] L. Maligranda, *Orlicz Spaces and Interpolation*, Seminars in Math. **5** (Univ. of Campinas, Brazil, 1989).
- [11] T. Miura, G. Hirasawa, and S.-E. Takahasi, *A perturbation of ring derivations on Banach algebras*, J. Math. Anal. Appl. **319** (2), 223–229, 2006.
- [12] M. S. Moslehain, *Hyers-Ulam-Rassias stability of generalized derivations*, Inter. J. Math. Sci. **2006**, Article ID 93942, 2006.
- [13] J. Musielak, *Orlicz Spaces and Modular Spaces*, Lecture Notes in Math. **1034** (Springer–verlag, Berlin, 1983).
- [14] H. Nakano, *Modulared Semi-Ordered Linear Spaces*, Tokyo Math. Book Ser. **1** (Maruzen Co., Tokyo, 1950).
- [15] W. Orlicz, *Collected Papers*, **I, II** (PWN, Warszawa, 1988).
- [16] C.-G. Park, *linear derivations on Banach algebras*, Nonlinear Funct. Anal. Appl. **9**, 359–368, 2004.
- [17] C.-G. Park, *Lie *-homomorphisms between Lie C*-algebras and Lie *- derivations on Lie C*-algebras*, J. Math. Anal. Appl. **293**, 419–434, 2004.
- [18] Th. M. Rassias, *On the stability of the linear mapping in Banach spaces*, Proc. Amer. Math. Soc. **72**, 297–300, 1978.
- [19] Gh. Sadeghi and T.L. Shateri, *Stability of derivation in modular spaces*, Mediterr. J. Math. (to appear).
- [20] P. *Šemrl*, *The functional equation of multiplicative derivation is superstable on standard operator algebras*, Integr. equ. oper. theory, **18**(1), 118–122, 1994.
- [21] S. M. Ulam, *Problems in Modern Mathematics*, Science Editions, **IV** (Wiley, New York, 1960).
- [22] S. Yamamuro, *On conjugate spaces of Nakano spaces*, Trans. Amer. Math. Soc. **90**, 291–311, 1959.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 651 – 657

Is homotopy perturbation method the traditional Taylor series expansion

M. Turkyilmazoglu *

Abstract

The present paper deals with the homotopy perturbation method. The question of whether the homotopy perturbation method is simply the conventional Taylor series expansion is examined. It is proven that under particular choices of the auxiliary parameters the homotopy perturbation method is indeed the Taylor series expansion of the sought solution of nonlinear equations.

Keywords: Nonlinearity, Analytic solution, Homotopy perturbation method.

Received 07/03/2012 : Accepted 30/07/2013 Doi : 10.15672/HJMS.2015449416

1. Introduction

Most of the real-life phenomena is governed by nonlinear equations whose solutions are difficult to find. Therefore, friendly tools have been the focus of past two decade's research.

Recently, investigators have proposed plenty of techniques to find approximate solutions. One of the most recent popular technique is the homotopy perturbation method based on the concept of topology. This method is quite distinct from the classical perturbation technique and does not require a small parameter or a linear term in a differential equation. Essentially, a homotopy with an embedding parameter $p \in [0, 1]$ is constructed. The basic details of homotopy perturbation method for solving nonlinear differential equations were outlined in [1], see also [2, 3, 4]. A numerous nonlinear problems were recently treated by the method, see for instance [5]. The recent works highlight clearly the fact that there is a close relationship between the Adomian decomposition and Taylor series methods as well as the homotopy and Taylor series methods [6, 7].

The investigation of current paper focuses on the homotopy perturbation technique. The prime motivation is to examine the method mathematically and to prove that under certain constraints, by particular choice of auxiliary linear operator and initial approximation, the homotopy perturbation method simply collapses onto the classical Taylor series expansion.

*Mathematics Department, University of Hacettepe, 06532-Beytepe, Ankara, Turkey
Email: turkyilm@hacettepe.edu.tr

2. The Homotopy Perturbation Method

The essential idea of this method is to introduce a homotopy parameter, say p, which varies from 0 to 1. Consider the nonlinear initial value problem

$$N(u) = 0, \quad B(u, \frac{du}{dn}) = 0, \tag{2.1}$$

where u is the function to be solved under the boundary constraints in B. He's homotopy perturbation technique [1, 10] defines a homotopy $u(r, p) : R \times [0, 1] \to R$ so that

$$H(u, p) = (1 - p)[L(u) - L(u_0)] + pN(u), \tag{2.2}$$

where L is a suitable auxiliary linear operator, u_0 is an initial approximation of equation (2.1) satisfying exactly the boundary conditions, see also [2] for the rest. It is obvious from equation (2.2) that

$$H(u, 0) = L(u) - L(u_0), \quad H(u, 1) = N(u). \tag{2.3}$$

As p moves from 0 to 1, $u(t, p)$ moves from $u_0(t)$ to $u(t)$. Our basic assumption is that the solution of equation (2.2) when equated to zero can be expressed as a power series in p

$$u(t, p) = u_0(t) + pu_1(t) + p^2 u_2(t) + \cdots = \sum_{k=0}^{\infty} u_k(t)p^k. \tag{2.4}$$

The approximate solution of equation (2.1), therefore, can be readily obtained as

$$u(t) = \lim_{p \to 1} u(t, p) = \sum_{k=0}^{\infty} u_k(t). \tag{2.5}$$

3. Homotopy perturbation and Taylor expansion

To answer the question raised in the title of the paper, let's take into account the first-order initial value problem version of (2.1)

$$u'(t) = F(u), \quad u(0) = \alpha, \tag{3.1}$$

where α is a constant. A straightforward Taylor series representation for the solution $u(t)$ at point $t = 0$ can be given in the form

$$u(t) = u(0) + u'(0)t + \frac{u''(0)}{2!}t^2 + \cdots = \sum_{k=0}^{\infty} a_k t^k, \tag{3.2}$$

where $a_n = \frac{u^{(n)}(0)}{n!}$ can be immediately found from differentiating (3.1) successively and substituting $t = 0$. A few of the coefficients follow

$$\begin{aligned} a_1 &= u'(0) = F(\alpha), \\ a_2 &= \frac{u''(0)}{2!} = \frac{1}{2!}F_u(\alpha)a_1, \\ a_3 &= \frac{u'''(0)}{3!} = \frac{1}{3!}[F_{uu}(\alpha)a_1^2 + 2F_u(\alpha)a_2], \\ a_4 &= \frac{u^{(4)}(0)}{4!} = \frac{1}{4!}[F_{uuu}(\alpha)a_1^3 + 6F_{uu}(\alpha)a_1 a_2 + 6F_u(\alpha)a_3], \\ &\vdots \end{aligned} \tag{3.3}$$

Theorem. If the auxiliary linear operator L and the initial approximation $u_0(t)$ to the solution $u(t)$ of equation (3.1) is taken in the homotopy procedure (2.2) as

$$L = \frac{\partial}{\partial t}, \quad u_0(t) = \alpha, \tag{3.4}$$

then the homotopy series solution (2.4) converges to the Taylor series expansion (3.2) whose coefficients are evaluated in the order given by (3.3).

Proof. Expanding the homotopy solution $u(t,p)$ from (2.2) into Taylor series according to the parameter p at $p=0$, it reads

$$u(t,p) = \sum_{k=0}^{\infty} u_k(t)p^k. \tag{3.5}$$

When (3.5) is substituted into the homotopy equations (2.2) or equivalently differentiating (2.2) successively with respect to p and replacing $p=0$ at the end yields a system of linear ordinary differential equations for the coefficients $u_k(t)$ of (3.5)

$$\begin{aligned} L(u_k - \chi_k u_{k-1}) &= -u'_{k-1} + \frac{1}{(k-1)!}[\frac{\partial^{k-1}F}{\partial p^{k-1}}]|_{p=0}, \\ u_k(0,p) &= \alpha, \end{aligned} \tag{3.6}$$

where $\chi_k = 0$ for $k=1$ and $\chi_k = 1$ for $k>1$. Having solved the equations (3.6) iteratively, the followings result for $u_k(t)$

$$\begin{aligned} u_1(t) &= F(\alpha)t = a_1 t, \\ u_2(t) &= \frac{1}{2!}F_u(\alpha)a_1 t^2 = a_2 t^2, \\ u_3(t) &= \frac{1}{3!}[F_{uu}(\alpha)a_1^2 + 2F_u(\alpha)a_2]t^3 = a_3 t^3, \\ u_4(t) &= \frac{1}{4!}[F_{uuu}(\alpha)a_1^3 + 6F_{uu}(\alpha)a_1 a_2 + 6F_u(\alpha)a_3]t^4 = a_4 t^4, \\ &\vdots \end{aligned} \tag{3.7}$$

which generates the homotopy series

$$u(t,p) = \sum_{k=0}^{\infty} u_k(t)p^k = \sum_{k=0}^{\infty} a_k t^k p^k. \tag{3.8}$$

The convergence assumption of (3.8) at $p=1$ yields the homotopy series solution (2.5) which turns out to be the Taylor series expansion (3.2-3.3) to the solution.

Remark 1. Since the homotopy series (3.8) at $p=1$ is the traditional Taylor series, then the convergence issue of the homotopy series (3.5) is guaranteed for those values t, $|t| < R$ such that $R = \lim_{n\to\infty} |\frac{a_n}{a_{n+1}}|$.

Remark 2. If $u = u(t,r)$ with r denoting space variables and the initial-value problem consists of a partial differential equation of the form

$$\begin{aligned} u_t &= F(u, u_r), \\ u(t=0, r) &= f(r), \end{aligned} \tag{3.9}$$

then by a similar argument to Theorem 1, the homotopy solution to (3.9) will be again the traditional Taylor series expansion at $t=0$, provided that the auxiliary linear operator

L and the initial approximation $u_0(t,r)$ are selected as

$$L = \frac{\partial}{\partial t}, \quad u_0(t,r) = f(r).$$

Remark 3. If higher-order ordinary or partial differential initial-value problems (or systems) are considered, by a particular choice of linear differential operator and initial guess, it can be shown that the homotopy perturbation series solution and the Taylor series solution are the same.

4. Illustrative Examples

To justify the presented analysis, the following examples are given, as also stated in reference [2].

Example 1. The steady free convection flow over a vertical semi-infinite flat plate, see [12] and [13] is given by

$$y' + y^2 = 1, \quad y(0) = 0, \tag{4.1}$$

To comply with the Theorem, $u_0(t) = 0$ and $L = \frac{d}{dt}$ are chosen so that the homotopy (2.2) becomes

$$\frac{\partial u(t,p)}{\partial t} + p\,u(x,p)^2 - p = 0, \quad u(0,p) = 0. \tag{4.2}$$

A few approximate homotopy solutions via the homotopy perturbation (4.2) can be calculated as

$$u_1(t) = t,\ u_2(t) = 0,\ u_3(t) = -\frac{t^3}{3},\ u_4(t) = 0,\ u_5(t) = \frac{2}{15}t^5,$$

which are the same as those generated from the classical Taylor series expansion of (4.1) at $t = 0$, the validity region is determined to be $-\frac{\pi}{2} < t < \frac{\pi}{2}$.

Example 2. The steady mixed convection flow [14] is given by

$$2y'' + y - y^2 = 0, \quad y(0) = 0, \quad y'(0) = \alpha = 1/\sqrt{6}. \tag{4.3}$$

The Taylor series expansion of (4.3) at point $t = 0$ yields

$$y(t) = t\alpha - \frac{t^3\alpha}{12} + \frac{t^4\alpha^2}{24} + \frac{t^5\alpha}{480} - \frac{t^6\alpha^2}{288} + \frac{t^7\alpha\left(-1+40\alpha^2\right)}{40320} + \frac{t^8\alpha^2}{7680} + \cdots . \tag{4.4}$$

which totally corresponds to the homotopy perturbation series solution provided that we choose the auxiliary parameters as $u_0(t) = \alpha t$ and $L = 2\frac{\partial^2}{\partial t^2}$, see [2].

Example 3. The approximate theory of the flow through a shock wave traveling in a viscous fluid [15] is given by

$$u_t + uu_x = u_{xx}, \quad u(x,0) = 2x,\ (x,t) \in R \times [0,1/2), \tag{4.5}$$

which receives an exact solution given by (see [2])

$$u(x,t) = \frac{2x}{1+2t}. \tag{4.6}$$

Exact solution (4.6) is approximated by the auxiliary parameters $u_0(x,t) = 2x$ and $L = \frac{\partial}{\partial t}$. Then, the homotopy (2.2) turns out to be

$$u_t(x,t,p) + p(u(x,t,p)u_x(x,p,t) - u_{xx}(x,t,p) = 0,\ u(x,0,p) = 2x. \tag{4.7}$$

Equation (4.7) produces the below homotopy series for the solution of (4.5)

$$u(x,t) = 2x - 4xt + 8xt^2 - 16xt^3 + 32xt^4 + ... + (-1)^n 2^{n+1} xt^n + \cdots, \tag{4.8}$$

which is the same as the classical Taylor series expansion of (4.6) around $t = 0$. The interval of convergence is easy to identify as $0 \leq t < 1/2$.

Example 4. Consider now the well-known KdV-Burger's equation involving both dispersion and dissipation terms

$$u_t + 2(u^3)_x - u_{xxx} + u_{xx} = 0, \quad u(x,0) = \frac{1}{6}\left(1 + \tanh\left[\frac{x}{6}\right]\right), \tag{4.9}$$

whose exact travelling-wave solution is given by

$$u(x,t) = \frac{1}{6}\left(1 + \tanh\left[\frac{1}{6}\left(x - \frac{2}{9}t\right)\right]\right). \tag{4.10}$$

To approximate the exact solution (4.10), if we choose the auxiliary parameters $u_0(x,t) = \frac{1}{6}\left(1 + \tanh\left[\frac{x}{6}\right]\right)$ and $L = \frac{\partial}{\partial t}$, the homotopy (2.2) turns out to be

$$u_t(x,t,p) + p(2(u(x,t,p)^3)_x - u_{xxx}(x,p,t) + u_{xx}(x,t,p) = 0, \tag{4.11}$$
$$u(x,0,p) = \frac{1}{6}\left(1 + \tanh\left[\frac{x}{6}\right]\right).$$

It is no hard to deduce that the Taylor series and homotopy perturbation series completely coincide again for this specific problem.

Example 5. The transverse vibrations of a uniform flexible beam [16] is given by

$$u_{tt} + (\frac{y+z}{2\cos x} - 1)u_{xxxx} + (\frac{z+x}{2\cos y} - 1)u_{yyyy} + (\frac{x+y}{2\cos z} - 1)u_{zzzz} = 0, \tag{4.12}$$
$$u(x,y,z,0) = -u_t(x,y,z,0) = x + y + z - (\cos x + \cos y + \cos z),$$

admitting an exact solution

$$u(x,t) = (x + y + z - \cos x - \cos y - \cos z)e^{-t}, see[2]. \tag{4.13}$$

This exact solution (4.13) is approximated by selecting the auxiliary parameters respectively, $u_0(x,t) = (x + y + z - \cos x - \cos y - \cos z)(1 - t)$ and $L = \frac{\partial^2}{\partial t^2}$. As a result, we obtain the homotopy series

$$u(x,t) = (x + y + z - \cos x - \cos y - \cos z)(1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots) \tag{4.14}$$

which matches exactly onto the Taylor series expansion of (4.13) around $t = 0$

$$u(x,t) = \sum_{n=0}^{\infty}(x + y + z - \cos x - \cos y - \cos z)\frac{t^n}{n!}, \tag{4.15}$$

that is obviously convergent for all t.

Example 6. As a final example, we consider the linear partial differential equation

$$u_t + u_x - 2u_{xxt} = 0, \quad u(x,0) = e^{-x}, \tag{4.16}$$

having the exact solution

$$u(x,t) = e^{-x-t}, see[2]. \tag{4.17}$$

Choosing the auxiliary parameters $u_0(x,t) = e^{-x}$ and $L = \frac{\partial}{\partial t}$, then the homotopy (2.2) becomes

$$u_t(x,t,p) + p(u_x(x,p,t) - u_{xxt}(x,t,p) = 0, \; u(x,0,p) = e^{-x}. \tag{4.18}$$

The homotopy series solution of (4.13) from (2.1) can be found as

$$(4.19)\; u(x,t) \;=\; \frac{e^{-x}}{720}(720+45360t+46440t^2+13320t^3+1470t^4+66t^5+\cdots),$$

whose radius of convergence is zero, so that the homotopy series (4.19) is convergent only at the point $t=0$. On the other hand, the classical Taylor series expansion applied to (4.16) predicts the exact result (4.17). It should be remarked that this example does not contradict at all with the Theorem, since (4.16) involves mixed partial derivatives. The weakness of the homotopy perturbation method on this example may be overcame by a better choice of auxiliary parameters.

It can be concluded as an answer to the title of the paper that for specific choices of auxiliary homotopy parameters, the homotopy perturbation technique produces exactly the same series as the traditional Taylor series. If this is the case, then there seems no a scientific merit to publish papers regarding the homotopy perturbation technique.

5. Concluding remarks

The homotopy perturbation method is mathematically analyzed in the present work. The theorem presented here proves that under certain special conditions the traditional homotopy perturbation method becomes the well-known Taylor series expansion. An example has also been given to demonstrate the advantage of the Taylor series expansion over the homotopy perturbation method. It can be concluded that a great deal of the papers published under the topic of homotopy perturbation technique is simply the traditional Taylor series expansion, whose contributions to science are questionable.

References

[1] He, J. H.: An approximate solution technique depending on an artificial parameter: A special example, Commun. Nonlinear Sci. Numer. Simulat. 3, 92–96 (1998).

[2] Turkyilmazoglu, M.: Convergence of the homotopy analysis method, arxiv4.library.cornell.edu 1–13 (2010).

[3] Turkyilmazoglu, M.: Some issues on HPM and HAM methods: A convergence scheme, Mathematical and Computer Modelling 53, 1929–1936 (2011).

[4] Turkyilmazoglu, M.: Convergence of the Homotopy Perturbation Method, Int. J. Nonlinear Sci. Numer. Simul 12, 9–14 (2012).

[5] Khan, Y.: Homotopy Perturbation Transform Method for nonlinear equations using He's Polynomials, Computers and Mathematics with Applications 61, 1963–1967 (2011).

[6] Duan, J. S.: Recurrence triangle for adomian polynomials, Appl. Math. Comput. 216, 1235–1241 (2010).

[7] Liu, C. S.: The essence of the homotopy analysis method, Appl. Math. Comput. 216, 1299–1303 (2010).

[8] He, J. H.: Homotopy perturbation technique, Comput. Meth. Appl. Mech. Eng. 178, 257–262 (1999).

[9] Liao, S. J.: The proposed homotopy analysis technique for the solution of nonlinear problems, PhD thesis, Shanghai Jiao Tong University, 1992.

[10] He, J. H.: Non-perturbative methods for strongly nonlinear problems, Berlin, dissertation de-Verlag im Internet GmbH, 2006.

[11] Liao, S. J.: Beyond perturbation: introduction to homotopy analysis method, Chapman & Hall/CRC, 2003.

[12] Crane, L.: Flow past a stretching plate, Z. angew. Math. Phys. 21, 645–647 (1970).

[13] Magyari, E. and Keller, B.: Exact solutions for self-similar boundary-layer flows induced by permeable stretching walls, Eur. J. Mech., B/Fluids 19, 109–122 (2000).

[14] Magyari, E. and Pop, I. and Keller, B.: Exact dual solutions occurring in darcy mixed convection flow, Int. J. Heat Mass Transfer 44, 4563–4566 (2001).

[15] Benton, E.R.: Some new exact, viscous, nonsteady solutions of Burgers' equation, Phys. Fluids. 9, 1247–1248 (1966).
[16] Gorman, D. J.: Free vibrations analysis of beams and shafts, Wiley, 1975.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 659 – 667

Existence and uniqueness of positive solutions for boundary value problems of a fractional differential equation with a parameter

Chen Yang *

Abstract

In this paper, we are concerned with the existence and uniqueness of positive solutions for the following nonlinear fractional two-point boundary value problem

$$\begin{cases} D_{0+}^{\alpha}u(t) + \lambda f(t, u(t), u(t)) = 0, \quad 0 < t < 1,\ 2 < \alpha \leq 3, \\ u(0) = u'(0) = u'(1) = 0, \end{cases}$$

where D_{0+}^{α} is the standard Riemann-Liouville fractional derivative, and λ is a positive parameter. Our analysis relies on a fixed point theorem and some properties of eigenvalue problems for a class of general mixed monotone operators. Our results can not only guarantee the existence of a unique positive solution, but also be applied to construct an iterative scheme for approximating it. An example is given to illustrate the main results.

2000 AMS Classification: 35J60, 35J65.

Keywords: Riemann-Liouville fractional derivative; Fractional differential equation; Positive solution; Existence and uniqueness; Mixed monotone operator.

Received 03/07/2013 : Accepted 20/11/2013 Doi : 10.15672/HJMS.20154510032

*Basic Course Department, Business College of Shanxi University, Taiyuan 030031 Shanxi, P.R.China
Email: yangchen0809@126.com
The research was supported by the Youth Science Foundation of China(11201272), the Science Foundation of Shanxi Province(2013011003-3) and the Science Foundation of Business College of Shanxi University(2012050).

1. Introduction

Fractional differential equations arise in many fields, such as physics, mechanics, chemistry, economics, engineering and biological sciences,etc;see[1-15] for example. In recent years, the study of positive solutions for fractional differential equation boundary value problems has attracted considerable attention, and fruits from research into it emerge continuously. For a small sample of such work, we refer the reader to [16-28] and the references therein. In these papers, many authors have investigated the existence of positive solutions for nonlinear fractional differential equation boundary value problems. Their results are based on Schauder fixed point theorem, Leggett-Williams theorem, fixed point index theorems in cones, Krasnosel'skii fixed point theorem, the method of upper-lower solutions, fixed point theorems in cones and so on. On the other hand, the uniqueness of positive solutions for nonlinear fractional differential equation boundary value problems has been studied by some authors, see [20-22,24,27] for example. The methods used in these papers are fixed point theorems for mixed monotone operators, u_0-concave operators and monotone operators in partially ordered sets.

In [26], by means of Krasnosel'skii fixed point theorem, El-Shahed considered the existence and nonexistence of positive solutions for the nonlinear fractional boundary value problem

$$\begin{cases} D_{0+}^{\alpha}u(t) + \lambda a(t) f(u(t)) = 0, \quad 0 < t < 1,\ 2 < \alpha < 3, \\ u(0) = u'(0) = u'(1) = 0, \end{cases}$$

where D_{0+}^{α} is the standard Riemann-Liouville fractional derivative, $a : (0,1) \to [0,+\infty)$ is continuous with $\int_0^1 a(t)dt > 0, f \in C([0,+\infty),[0,+\infty))$ and λ is a positive parameter.

In [28], by using the properties of the Green function, the method of upper-lower solutions and fixed point theorem, Zhao et al. studied the existence of multiple positive solutions for the nonlinear fractional differential equation boundary value problem

$$\begin{cases} D_{0+}^{\alpha}u(t) + f(t,u(t)) = 0, \quad 0 < t < 1,\ 2 < \alpha \le 3, \\ u(0) = u'(0) = u'(1) = 0. \end{cases}$$

The purpose of this paper is to establish the existence and uniqueness of positive solutions for the following nonlinear fractional two-point boundary value problem

$$\begin{cases} D_{0+}^{\alpha}u(t) + \lambda f(t,u(t),u(t)) = 0, \quad 0 < t < 1,\ 2 < \alpha \le 3, \\ u(0) = u'(0) = u'(1) = 0, \end{cases} \tag{1.1}$$

where D_{0+}^{α} is the standard Riemann-Liouville fractional derivative, λ is a positive parameter and $f : [0,1] \times [0,+\infty) \times [0,+\infty) \to [0,+\infty)$ is continuous.

Different from the above works mentioned, we will use a fixed point theorem and some properties of eigenvalue problems for a class of general mixed monotone operators to show the existence and uniqueness of positive solutions for the problem (1.1). Moreover, we can construct two sequences for approximating the unique solution and we show that the positive solution with respect to λ has some pleasant properties.

2. Preliminaries and previous results

For the convenience of the reader, we present here some definitions, lemmas and basic results that will be used in the proof of our theorem.

Definition 2.1([4, Definition 2.1]). The integral

$$I_{0+}^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x \frac{f(t)}{(x-t)^{1-\alpha}} dt, \quad x > 0$$

is called the Riemann-Liouville fractional integral of order α, where $\alpha > 0$ and $\Gamma(\alpha)$ denotes the gamma function.

Definition 2.2([4, page 36-37]). For a function $f(x)$ given in the interval $[0, \infty)$, the expression

$$D_{0+}^{\alpha} f(x) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dx} \right)^n \int_0^x \frac{f(t)}{(x-t)^{\alpha-n+1}} dt$$

where $n = [\alpha] + 1, [\alpha]$ denotes the integer part of number α, is called the Riemann-Liouville fractional derivative of order α.

Lemma 2.3.([26]). Given $y \in C[0,1]$ and $2 < \alpha \leq 3$, the following boundary value problem

$$\begin{cases} D_{0+}^{\alpha} u(t) + y(t) = 0, \quad 0 < t < 1, \\ u(0) = u'(0) = u'(1) = 0, \end{cases} \tag{2.1}$$

has a unique solution

$$u(t) = \int_0^1 G(t,s) y(s) ds, \ t \in [0,1],$$

where

$$G(t,s) = \frac{1}{\Gamma(\alpha)} \begin{cases} (1-s)^{\alpha-2} t^{\alpha-1}, & 0 \leq t \leq s \leq 1, \\ (1-s)^{\alpha-2} t^{\alpha-1} - (t-s)^{\alpha-1}, & 0 \leq s \leq t \leq 1. \end{cases}$$

Here $G(t,s)$ is called the Green function of boundary value problem (2.1). Evidently, $G(t,s) \geq 0$ for $t, s \in [0,1]$.
The following property of the Green function plays important roles in this paper.

Lemma 2.4. Let $2 < \alpha \leq 3$. Then the Green function $G(t,s)$ in Lemma 2.3 has the following property:

$$\frac{1}{\Gamma(\alpha)} s(1-s)^{\alpha-2} t^{\alpha-1} \leq G(t,s) \leq \frac{1}{\Gamma(\alpha)} (1-s)^{\alpha-2} t^{\alpha-1} \text{ for } \ t, s \in [0,1].$$

Proof. Evidently, the right inequality holds. So we only need to prove the left inequality. If $0 \leq s \leq t \leq 1$, then we have $0 \leq t - s \leq t - ts = (1-s)t$, and thus

$$(t-s)^{\alpha-1} \leq (1-s)^{\alpha-1} t^{\alpha-1}.$$

Hence,

$$\begin{aligned} G(t,s) &= \frac{1}{\Gamma(\alpha)} [(1-s)^{\alpha-2} t^{\alpha-1} - (t-s)^{\alpha-1}] \\ &\geq \frac{1}{\Gamma(\alpha)} [(1-s)^{\alpha-2} t^{\alpha-1} - (1-s)^{\alpha-1} t^{\alpha-1}] \\ &= \frac{1}{\Gamma(\alpha)} s(1-s)^{\alpha-2} t^{\alpha-1}. \end{aligned}$$

If $0 \leq t \leq s \leq 1$, then we have

$$G(t,s) = \frac{1}{\Gamma(\alpha)} (1-s)^{\alpha-2} t^{\alpha-1} \geq \frac{1}{\Gamma(\alpha)} s(1-s)^{\alpha-2} t^{\alpha-1}.$$

So the left inequality also holds.□

In the sequel, we present some basic concepts in ordered Banach spaces for completeness and fixed point theorems which we will be used later. For convenience of readers, we suggest that one refer to [29,30] for details.

Suppose that $(E, \|\cdot\|)$ is a real Banach space which is partially ordered by a cone $P \subset E$, *i.e.*, $x \leq y$ if and only if $y - x \in P$. If $x \leq y$ and $x \neq y$, then we denote $x < y$ or $y > x$. By θ we denote the zero element of E. Recall that a non-empty closed convex set $P \subset E$ is a cone if it satisfies (i) $x \in P, \lambda \geq 0 \Rightarrow \lambda x \in P$; (ii) $x \in P, -x \in P \Rightarrow x = \theta$.

P is called normal if there exists a constant $M > 0$ such that, for all $x, y \in E$, $\theta \leq x \leq y$ implies $\|x\| \leq M\|y\|$; in this case M is called the normality constant of P. If $x_1, x_2 \in E$, the set $[x_1, x_2] = \{x \in E | x_1 \leq x \leq x_2\}$ is called the order interval between x_1 and x_2.

For all $x, y \in E$, the notation $x \sim y$ means that there exist $\lambda > 0$ and $\mu > 0$ such that $\lambda x \leq y \leq \mu x$. Clearly, $\sim$ is an equivalence relation. Given $h > \theta$(*i.e.*, $h \geq \theta$ and $h \neq \theta$), we denote by P_h the set $P_h = \{x \in E |\ x \sim h\}$. It is easy to see that $P_h \subset P$ is convex and $\lambda P_h = P_h$ for all $\lambda > 0$.

Definition 2.5(see [29,30]). $A : P \times P \to P$ is said to be a mixed monotone operator if $A(x, y)$ is increasing in x and decreasing in y, i.e., $u_i, v_i (i = 1, 2) \in P, u_1 \leq u_2, v_1 \geq v_2$ implies $A(u_1, v_1) \leq A(u_2, v_2)$. Element $x \in P$ is called a fixed point of A if $A(x, x) = x$.

In a recent paper [30], Zhai and Zhang considered the following operator equations

$$A(x, x) = x \text{ and } A(x, x) = \lambda x,$$

where $A : P \times P \to P$ is a mixed monotone operator which satisfy the following conditions:

(A_1) there exists $h \in P$ with $h \neq \theta$ such that $A(h, h) \in P_h$.
(A_2) for any $u, v \in P$ and $t \in (0, 1)$, there exists $\varphi(t) \in (t, 1)$ such that $A(tu, t^{-1}v) \geq \varphi(t)A(u, v)$.
They established the existence and uniqueness of positive solutions for the above equations and they present the following interesting results.

Theorem 2.6. Suppose that P is a normal cone of E, and $(A_1), (A_2)$ hold. Then operator A has a unique fixed point x^* in P_h. Moreover, for any initial $x_0, y_0 \in P_h$, constructing successively the sequences

$$x_n = A(x_{n-1}, y_{n-1}), y_n = A(y_{n-1}, x_{n-1}), n = 1, 2, \ldots,$$

we have $\|x_n - x^*\| \to 0$ and $\|y_n - x^*\| \to 0$ as $n \to \infty$.

Theorem 2.7. Suppose that P is a normal cone of E, and $(A_1), (A_2)$ hold. Let $x_\lambda (\lambda > 0)$ denote the unique solution of nonlinear eigenvalue equation $A(x, x) = \lambda x$ in P_h. Then we have the following conclusions:

(R_1) If $\varphi(t) > t^{\frac{1}{2}}$ for $t \in (0, 1)$, then x_λ is strictly decreasing in λ, that is, $0 < \lambda_1 < \lambda_2$ implies $x_{\lambda_1} > x_{\lambda_2}$;
(R_2) If there exists $\beta \in (0, 1)$ such that $\varphi(t) \geq t^\beta$ for $t \in (0, 1)$, then x_λ is continuous in λ, that is, $\lambda \to \lambda_0 (\lambda_0 > 0)$ implies $\|x_\lambda - x_{\lambda_0}\| \to 0$;
(R_3) If there exists $\beta \in (0, \frac{1}{2})$ such that $\varphi(t) \geq t^\beta$ for $t \in (0, 1)$, then $\lim_{\lambda \to \infty} \|x_\lambda\| = 0$, $\lim_{\lambda \to 0^+} \|x_\lambda\| = \infty$.

3. Existence and uniqueness of positive solutions for the problem (1.1)

In this section, we apply Theorem 2.6 and Theorem 2.7 to study the problem (1.1), and we obtain a new result on the existence and uniqueness of positive solutions. Moreover, we show that the positive solution with respect to λ has some pleasant properties. The method used here is new to the literature and so is the existence and uniqueness result to the fractional differential equations.

In our considerations we will work in the Banach space $C[0,1] = \{x : [0,1] \to \mathbf{R}$ is continuous$\}$ with the standard norm $\|x\| = \sup\{|x(t)| : t \in [0,1]\}$. Notice that this space can be equipped with a partial order given by

$$x, y \in C[0,1], x \leq y \Leftrightarrow x(t) \leq y(t) \text{ for } t \in [0,1].$$

Set $P = \{x \in C[0,1] | x(t) \geq 0,\ t \in [0,1]\}$, the standard cone. It is clear that P is a normal cone in $C[0,1]$ and the normality constant is 1. Our main result is summarized in the following theorem.

Theorem 3.1. Assume that

(H_1) $f(t,x,y)$ is nondecreasing in x for each $t \in [0,1]$ and $y \in [0,+\infty)$, nonincreasing in y for each $t \in [0,1]$ and $x \in [0,+\infty)$ with $f(t,0,1) \not\equiv 0$;
(H_2) for any $\gamma \in (0,1)$, there exist constants $\varphi_1(\gamma), \varphi_2(\gamma) \in (0,1)$ with $\varphi_1(\gamma)\varphi_2(\gamma) > \gamma$ such that

$$f(t,\gamma x,y) \geq \varphi_1(\gamma) f(t,x,y), f(t,x,\gamma y) \leq \frac{1}{\varphi_2(\gamma)} f(t,x,y) \text{ for any } \ x, y \in [0,+\infty).$$

Then: (1) the problem (1.1) has a unique positive solution u_λ^* in P_h, where $h(t) = t^{\alpha-1}, t \in [0,1]$. Moreover, for any initial values $u_0, v_0 \in P_h$, constructing successively the sequences

$$u_{n+1}(t) = \lambda \int_0^1 G(t,s) f(s,u_n(s),v_n(s))ds,\ v_{n+1}(t) = \lambda \int_0^1 G(t,s) f(s,v_n(s),u_n(s))ds,\ n = 0,1,2,\ldots,$$

we have $u_n(t) \to u_\lambda^*(t), v_n(t) \to u_\lambda^*(t)$ as $n \to \infty$, where $G(t,s)$ is given as in Lemma 2.3;
(2) if $\varphi_1(t)\varphi_2(t) > t^{\frac{1}{2}}$ for $t \in (0,1)$, then u_λ^* is strictly increasing in λ, that is, $0 < \lambda_1 < \lambda_2$ implies $u_{\lambda_1}^* \leq u_{\lambda_2}^*, u_{\lambda_1}^* \neq u_{\lambda_2}^*$. If there exists $\beta \in (0,1)$ such that $\varphi_1(t)\varphi_2(t) \geq t^\beta$ for $t \in (0,1)$, then u_λ^* is continuous in λ, that is, $\lambda \to \lambda_0 (\lambda_0 > 0)$ implies $\|u_\lambda^* - u_{\lambda_0}^*\| \to 0$. If there exists $\beta \in (0,\frac{1}{2})$ such that $\varphi_1(t)\varphi_2(t) \geq t^\beta$ for $t \in (0,1)$, then $\lim_{\lambda \to 0^+} \|u_\lambda^*\| = 0,\ \lim_{\lambda \to \infty} \|u_\lambda^*\| = \infty$.

Proof. To begin with, from [26] the problem (1.1) has an integral formulation given by

$$u(t) = \lambda \int_0^1 G(t,s) f(s,u(s),u(s))ds,$$

where $G(t,s)$ is given as in Lemma 2.3. For any $u, v \in P$, we define

$$A(u,v)(t) = \int_0^1 G(t,s) f(s,u(s),v(s))ds.$$

Noting that $f(t,x,y) \geq 0$ and $G(t,s) \geq 0$, it is easy to check that $A : P \times P \to P$. In the sequel we check that A satisfies all assumptions of Theorem 2.6.

Firstly, we prove that A is a mixed monotone operator. In fact, for $u_i, v_i \in P, i = 1,2$ with $u_1 \geq u_2, v_1 \leq v_2$, we know that $u_1(t) \geq u_2(t), v_1(t) \leq v_2(t), t \in [0,1]$ and by (H_1)

and Lemma 2.3,

$$A(u_1, v_1)(t) = \int_0^1 G(t,s)f(s, u_1(s), v_1(s))ds \geq \int_0^1 G(t,s)f(s, u_2(s), v_2(s))ds = A(u_2, v_2)(t).$$

That is, $A(u_1, v_1) \geq A(u_2, v_2)$.

Next we show that A satisfies the condition (A_2). From (H_2), for $\gamma \in (0,1)$ we can get $f(t, x, \gamma^{-1}y) \geq \varphi_2(\gamma)f(t, x, y)$ for any $x, y \in [0, +\infty)$. Then for any $\gamma \in (0,1)$ and $u, v \in P$, we obtain

$$\begin{aligned} A(\gamma u, \gamma^{-1}v)(t) &= \int_0^1 G(t,s)f(s, \gamma u(s), \gamma^{-1}v(s))ds \\ &\geq \int_0^1 G(t,s)\varphi_1(\gamma)f(s, u(s), \gamma^{-1}v(s))ds \\ &\geq \int_0^1 G(t,s)\varphi_1(\gamma)\varphi_2(\gamma)f(s, u(s), v(s))ds \\ &= \varphi_1(\gamma)\varphi_2(\gamma)A(u,v)(t), \ t \in [0,1]. \end{aligned}$$

Let $\varphi(t) = \varphi_1(t)\varphi_2(t), t \in (0,1)$. Then $\varphi(t) \in (t, 1)$ for $t \in (0,1)$. Hence, $A(\gamma u, \gamma^{-1}v) \geq \varphi(\gamma)A(u,v), \ \forall\ u, v \in P, \ \gamma \in (0,1)$. So the condition (A_2) in Theorem 2.6 is satisfied. Now we show that $A(h,h) \in P_h$. On one hand, it follows from $(H_1), (H_2)$ and Lemma 2.4 that

$$\begin{aligned} A(h,h)(t) &= \int_0^1 G(t,s)f(s, h(s), h(s))ds \\ &= \int_0^1 G(t,s)f(s, s^{\alpha-1}, s^{\alpha-1})ds \\ &\geq \int_0^1 \frac{1}{\Gamma(\alpha)}s(1-s)^{\alpha-2}t^{\alpha-1}f(s,0,1)ds \\ &= \frac{1}{\Gamma(\alpha)}h(t)\int_0^1 s(1-s)^{\alpha-2}f(s,0,1)ds, \ t \in [0,1]. \end{aligned}$$

On the other hand, also from $(H_1), (H_2)$ and Lemma 2.4, we obtain

$$\begin{aligned} A(h,h)(t) &= \int_0^1 G(t,s)f(s, s^{\alpha-1}, s^{\alpha-1})ds \\ &\leq \int_0^1 \frac{1}{\Gamma(\alpha)}(1-s)^{\alpha-2}t^{\alpha-1}f(s,1,0)ds \\ &= \frac{1}{\Gamma(\alpha)}h(t)\int_0^1 f(s,1,0)ds, \ t \in [0,1]. \end{aligned}$$

Let

$$r_1 = \int_0^1 s(1-s)^{\alpha-2}f(s,0,1)ds, \ r_2 = \int_0^1 f(s,1,0)ds.$$

Since f is continuous and $f(t,0,1) \not\equiv 0$, we can get

$$0 < r_1 = \int_0^1 s(1-s)^{\alpha-2}f(s,0,1)ds \leq \int_0^1 f(s,1,0)ds = r_2.$$

Consequently,

$$A(h,h)(t) \geq \frac{r_1}{\Gamma(\alpha)} \cdot h(t), \ A(h,h)(t) \leq \frac{r_2}{\Gamma(\alpha)} \cdot h(t), \ t \in [0,1].$$

So we have

$$\frac{r_1}{\Gamma(\alpha)} \cdot h \leq A(h,h) \leq \frac{r_2}{\Gamma(\alpha)} \cdot h.$$

Hence $A(h,h) \in P_h$, the condition (A_1) in Theorem 2.6 is satisfied. Therefore, by Theorem 2.7, there exists a unique $u_\lambda^* \in P_h$ such that $A(u_\lambda^*, u_\lambda^*) = \frac{1}{\lambda}u_\lambda^*$. That is, $u_\lambda^* = \lambda A(u_\lambda^*, u_\lambda^*)$. It is easy to check that u_λ^* is a unique positive solution of the problem (1.1) for given $\lambda > 0$. Further, if $\varphi(t) > t^{\frac{1}{2}}$ for $t \in (0,1)$, then Theorem 2.7 (R_1) means that u_λ^* is strictly increasing in λ, that is, $0 < \lambda_1 < \lambda_2$ implies $u_{\lambda_1}^* \leq u_{\lambda_2}^*, u_{\lambda_1}^* \neq u_{\lambda_2}^*$. If there exists $\beta \in (0,1)$ such that $\varphi(t) \geq t^\beta$ for $t \in (0,1)$, then Theorem 2.7 (R_2) means that u_λ^* is continuous in λ, that is, $\lambda \to \lambda_0(\lambda_0 > 0)$ implies $\|u_\lambda^* - u_{\lambda_0}^*\| \to 0$. If there exists $\beta \in (0, \frac{1}{2})$ such that $\varphi(t) \geq t^\beta$ for $t \in (0,1)$, then Theorem 2.7 (R_3) means $\lim\limits_{\lambda\to 0^+} \|u_\lambda^*\| = 0,\ \lim\limits_{\lambda\to\infty} \|u_\lambda^*\| = \infty$.

Let $A_\lambda = \lambda A$, then A_λ also satisfies all the conditions of Theorem 2.6. By Theorem 2.6, for any initial values $u_0, v_0 \in P_h$, constructing successively the sequences $u_{n+1} = A_\lambda(u_n, v_n),\ v_{n+1} = A_\lambda(v_n, u_n), n = 0, 1, 2, \ldots$, we have $u_n \to u_\lambda^*, v_n \to u_\lambda^*$ as $n \to \infty$. That is,

$$u_{n+1}(t) = \lambda \int_0^1 G(t,s) f(s, u_n(s), v_n(s)) ds \to u_\lambda^*(t),$$

$$v_{n+1}(t) = \lambda \int_0^1 G(t,s) f(s, v_n(s), u_n(s)) ds,\ \to u_\lambda^*(t)$$

as $n \to \infty$. □

Remark 3.1. Let $f(t,x,y) \equiv C > 0$. Then the conditions $(H_1), (H_2)$ are satisfied and the problem (1.1) has a unique solution $u_\lambda(t) = \lambda C \int_0^1 G(t,s) ds,\ t \in [0,1]$. From Lemma 2.4, the unique solution u_λ is a positive solution and satisfies $u_\lambda \in P_h = P_{t^{\alpha-1}}$.

Example 3.1. Consider the following problem:

$$\begin{cases} D_{0+}^{\frac{5}{2}} u(t) + \lambda a(t)[u^{\frac{1}{5}}(t) + (u(t)+3)^{-\frac{1}{4}}] = 0, \quad 0 < t < 1, \\ u(0) = u'(0) = u'(1) = 0, \end{cases} \tag{3.1}$$

where $a : [0,1] \to [0, +\infty)$ is continuous with $a \not\equiv 0$.

In this example, we have $\alpha = \frac{5}{2}$. Let $f(t,x,y) = a(t)[x^{\frac{1}{5}} + (y+3)^{-\frac{1}{4}}]$. Evidently, $f(t,x,y)$ is increasing in x for $t \in [0,1], y \geq 0$, decreasing in y for $t \in [0,1], x \geq 0$. Moreover, $f(t,0,1) = a(t)4^{-\frac{1}{4}} \not\equiv 0$. Set $\varphi_1(\gamma) = \gamma^{\frac{1}{5}}, \varphi_2(\gamma) = \gamma^{\frac{1}{4}}, \gamma \in (0,1)$. Then $\varphi_1(\gamma)\varphi_2(\gamma) = \gamma^{\frac{9}{20}} > \gamma$ and

$$f(t,\gamma x,y) = a(t)[\gamma^{\frac{1}{5}} x^{\frac{1}{5}} + (y+3)^{-\frac{1}{4}}] \geq \varphi_1(\gamma) f(t,x,y), f(t,x,\gamma y) = a(t)[x^{\frac{1}{5}} + \frac{1}{\gamma^{\frac{1}{4}}}(y+3)^{-\frac{1}{4}}] \leq \frac{1}{\varphi_2(\gamma)} f(t,x,y),$$

for $t \in [0,1], x, y \geq 0$. Hence, all the conditions of Theorem 3.1 are satisfied. An application of Theorem 3.1 implies that the problem (3.1) has a unique positive solution u_λ^* in $P_h = P_{t^{\alpha-1}}$, and for any initial values $u_0, v_0 \in P_{t^{\alpha-1}}$, constructing successively the sequences

$$u_{n+1}(t) = \lambda \int_0^1 G(t,s) a(s)[u_n^{\frac{1}{5}}(s) + (v_n(s)+3)^{-\frac{1}{4}}] ds,\ v_{n+1}(t) = \lambda \int_0^1 G(t,s) a(s)[v_n^{\frac{1}{5}}(s) + (u_n(s)+3)^{-\frac{1}{4}}] ds,$$

$n = 0, 1, 2, \ldots$, we have $u_n(t) \to u_\lambda^*(t), v_n(t) \to u_\lambda^*(t)$ as $n \to \infty$, where $G(t,s)$ is given as in Lemma 2.3. Moreover, note that $\varphi_1(t)\varphi_2(t) > t^{\frac{1}{2}}$ for $t \in (0,1)$, then from Theorem 3.1, u_λ^* is strictly increasing in λ, that is, $0 < \lambda_1 < \lambda_2$ implies $u_{\lambda_1}^* \leq u_{\lambda_2}^*, u_{\lambda_1}^* \neq u_{\lambda_2}^*$. Take $\beta \in [\frac{9}{20}, \frac{1}{2})$ and applying Theorem 3.1, we know that u_λ^* is continuous in λ and $\lim\limits_{\lambda\to 0^+} \|u_\lambda^*\| = 0,\ \lim\limits_{\lambda\to\infty} \|u_\lambda^*\| = \infty$.

References

[1] K.B. Oldham, J. Spanier, The Fractional Calculus, Academic Press, New York, London, (1974).

[2] L. Gaul, P. Klein, S. Kempffe, Damping description involving fractional operators, Mech. Systems Signal Processing, 5(1991) 81-88.

[3] K.S. Miller, B. Ross, An Introduction to the Fractional Calculus and Fractional Differential Equations, Wiley, New York, 1993.

[4] S.G. Samko, A.A. Kilbas, O.I. Marichev, Fractional Integral and Derivatives (Theory and Applications), Gordon and Breach, Switzerland, 1993.

[5] F. Metzler, W. Schick, H.G. Kilian, T.F. Nonnenmacher, Relaxation in filled polymers: A fractional calculus approach, J. Chem. Phys. 103 (1995) 7180-7186.

[6] W.G. Glockle, T.F. Nonnenmacher, A fractional calculus approach of self-similar protein dynamics, Biophys. J., 68(1995) 46-53.

[7] F. Mainardi, Fractional calculus: Some basic problems in continuum and statistical mechanics, in Fractals and Fractional Calculus in Continuum Mechanics, (C.A. Carpinteri and F. Mainardi, Eds), Springer-Verlag, Wien, 1997, 291-348.

[8] K. Diethelm, A.D. Freed, On the solutions of nonlinear fractional order differential equations used in the modelling of viscoplasticity in Keil, F., Mackens, W., Voss, H., Werthers, J., (Eds), Scientifice Computing in Chemical Engineering II- Computa- tional Fluid Dynamics, Reaction Engineering and Molecular Properties, Springer-Verlag, Heildelberg,(1999), 217-224.

[9] I. Podlubny, Fractional Differential Equations, Mathematics in Science and Engineering, Academic Press, New York, 1999.

[10] A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, Theory and Applications of Fractional Differential Equations, in: North-Holland Mathematics Studies, vol. 204, Elsevier, Amsterdam, 2006.

[11] E. M. Rabei, K. I. Nawa eh, R. S. Hijjawi, S. I. Muslih, D. Baleanu, The Hamilton formalism with fractional derivatives, J. Math. Anal. Appl. 327 (2007) 891-897.

[12] V. Lakshmikantham, Theory of fractional functional differential equations, Nonlinear Anal. 69 (2008) 3337-3343.

[13] Y. Zhou, Existence and uniqueness of solutions for a system of fractional differential equations, J. Frac. Calc. Appl. Anal. 12 (2009) 195-204.

[14] N. Kosmatov, A singular boundary value problem for nonlinear differential equations of fractional order, J. Appl. Math. Comput. 29(2009) 125-135.

[15] C. Lizama, An operator theoretical approach to a class of fractional order differential equations, Appl. Math. Lett. 24(2011) 184-190.

[16] S.Q. Zhang, Existence of positive solution for some class of nonlinear fractional differential equations, J. Math. Anal. Appl. 278 (2003) 136-148.

[17] Z.B. Bai, H.S. L Positive solutions for boundary value problem of nonlinear fractional differential equation, J. Math. Anal. Appl. 311 (2005) 495-505.

[18] E.R. Kaufmann, E. Mboumi, Positive solutions of a boundary value problem for a nonlinear fractional differential equation, Electron. J. Qual. Theory Diff. Equ. 2008 (3) (2008) 1-11.

[19] C.Z. Bai, Triple positive solutions for a boundary value problem of nonlinear fractional differential equation, Electron. J. Qual. Theory Diff. Equ. 2008 (24) (2008) 1-10.

[20] X. Xu, D. Jiang, C. Yuan, Multiple positive solutions for boundary value problem of nonlinear fractional differential equation, Nonlinear Anal. 71(2009) 4676-4688.

[21] J. Caballero Mena, J. Harjani, K. Sadarangani, Existence and uniqueness of positive and nondecreasing solutions for a class of singular fractional boundary value problems, Bound. Value Probl. 2009 (2009) 10 pages. Article ID 421310, doi:10.1155/2009/421310.

[22] S.Q. Zhang, Positive solutions to singular boundary value problem for nonlinear fractional differential equation, Comput. Math. Appl. 59 (2010) 1300-1309.

[23] D. Jiang, C. Yuan, The positive properties of the Green function for Dirichlet-type boundary value problems of nonlinear fractional differential equations and its application, Nonlinear Anal. 72 (2010) 710-719.

[24] L. Yang, H. Chen, Unique positive solutions for fractional differential equation boundary value problems, Appl. Math. Lett. 23(2010) 1095-1098.

[25] Y.Q. Wang, L.S. Liu, Y.H. Wu, Positive solutions for a nonlocal fractional differential equation, Nonlinear Anal. 74(2011) 3599-3605.

[26] El-Shahed M, Positive solutions for boundary value problems of nonlinear fractional differential equation. Abstr. Appl. Anal. 2007(2007) 8 pages.

[27] S.H. Liang, J.H. Zhang, Existence and uniqueness of strictly nondecreasing and positive solution for a fractional three-point boundary value problem, Comput. Math. Appl. 62(2011) 1333-1340.

[28] Y.G. Zhao, S.R. Sun, Z.L. Han, Q.P. Li, The existence of multiple positive solutions for boundary value problems of nonlinear fractional differential equations Commun. Nonlinear Sci. Numer. Simulat. 16 (2011) 2086-2097.

[29] D. Guo, V.Lakshmikantham, Nonlinear Problems in Abstract Cones. Boston and New York: Academic Press Inc, 1988.

[30] C.B. Zhai, L.L. Zhang, New fixed point theorems for mixed monotone operators and local existence Cuniqueness of positive solutions for nonlinear boundary value problems, J.Math.Anal. Appl. 382 (2011) 594-614.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 669 – 677

Generalized closed sets and some separation axioms on weak structure

A. M. Zahran * , A. K. Mousa † and A. Ghareeb ‡

Abstract

In this paper, we introduce and characterize the concepts of generalized closed (gw-closed, for short) sets in weak structures which introduced by Császár [3] and we give some properties of these concepts. The concept of gw-closed sets (in the sense of Al Omari and Noiri [1]) is a special case of gw-closed sets presented here. Finally, the concepts of $T_{\frac{1}{2}}$-, T_1-, normal, almost normal and weakly normal spaces are investigated by using the concepts of gw-closed, sgw-closed and mgw-closed sets in weak structures.

2000 AMS Classification: 54A05, 54C08.

Keywords: Weak structures, gw-closed sets, $T_{\frac{1}{2}}$-space, T_1-space, normal space, almost normal space, weakly normal spaces.

Received 16/09/2013 : Accepted 12/03/2014 Doi : 10.15672/HJMS.2015449748

1. Preliminaries

In 2002, Császár [2] introduced the concept of generalized topology and investigated some concepts such as continuity, generalized open sets. In 2005, Maki et al. [5] introduced the concept of minimal structure and investigated some of its properties. Finally, Császár [3] introduced the concept of weak structure (Let X be a non-empty set and $P(X)$ its power set. A class $w \subset P(X)$ is said to be a weak structure (WS, for short) on X if and only if $\phi \in w$). He defined a subset A is said to be w-open if $A \in w$ and its complement is called w-closed. Also, he defined two operations $i_w(A)$ and $c_w(A)$ in WS on X as the union of all w-open subsets of A and the intersection of all w-closed set containing A, respectively. Furthermore, he gave some properties of $c_w(A)$ and $i_w(A)$.

*Department of Mathematics, Faculty of Science, Al Azhar University, Assuit, Egypt. Email: amzahran60@aol.com

†Department of Mathematics, Faculty of Science, Al Azhar University, Assuit, Egypt. Email: a_k_musa@yahoo.com

‡Department of Mathematics, Faculty of Science, South Valley University, Qena, Egypt. Email: nasserfuzt@hotmail.com, a.ghareeb@sci.svu.edu.eg

The purpose of this paper is to introduce and study the concepts of generalized closed sets in weak structures and we give some characterizations and properties of these concepts. The concept of gw-closed sets (in the sense of Al Omari and Noiri [1]) is a special case of gw-closed sets in a weak structure. Finally, the concepts of $T_{\frac{1}{2}}$-, T_1-, normal, almost normal and weakly normal spaces are investigated by using the concepts of gw-closed, sgw-closed and mgw-closed sets in weak structures. It is shown that many results in previous papers [1, 6, 7] can be considered as special cases of our results.

1.1. Theorem. [3] *Let w be a WS on X and A, $B \subseteq X$. Then the following statements are true:*

(1) $A \subseteq c_w(A)$,
(2) *If* $A \subseteq B$, *then* $c_w(A) \subset c_w(B)$,
(3) *If* A *is* w*-closed, then* $A = c_w(A)$,
(4) $c_w(c_w(A)) = c_w(A)$,
(5) $A \supseteq i_w(A)$,
(6) *If* $A \subset B$, *then* $i_w(A) \subset i_w(B)$,
(7) $i_w(i_w(A)) = i_w(A)$,
(8) *If* A *is* w*-open, then* $A = i_w(A)$,
(9) $c_w(X - A) = X - i_w(A)$,
(10) $i_w(X - A) = X - c_w(A)$,
(11) $i_w(c_w(i_w(c_w(A)))) = i_w(c_w(A))$,
(12) $c_w(i_w(c_w(i_w(A)))) = c_w(i_w(A))$,
(13) $x \in i_w(A)$ *if only if there is a* w*-open set* U *such that* $x \in U \subset A$,
(14) $x \in c_w(A)$ *if and only if* $U \cap A \neq \phi$ *for each* w*-open set* U *containing* x.

1.2. Definition. [4] Let w be a WS on X and $A \subseteq X$. Then:

(1) $A \in r(w)$ (i.e., A is w-regular open subset) if $A = i_w(c_w(A))$,
(2) $A \in rc(w)$ (i.e., A is w-regular closed subset) if $A = c_w(i_w((A))$.

1.3. Definition. Let w be a WS on X and $A \subset X$. A point $x \in X$ is said to be w-boundary point of a subset A if and only if $x \in c_w(A) \bigcap c_w(X - A)$. By $Bd_w(A)$ we denote the set of all w-boundary points of A.

1.4. Theorem. *Let w be a WS on X and $A \subseteq X$. Then:*

(1) $Bd_w(A) = Bd_w(X - A)$,
(2) $Bd_w(A) = c_w(A) - i_w(A)$,
(3) *If* A *is* w*-open, then* $A \cap Bd_w(A) = \phi$,
(4) *If* A *is* w*-closed, then* $Bd_w(A) \subset A$.

Proof. It follows from Definition 1.3 and Theorem 1.1. □

1.5. Remark. One may notice that the converses of (3) and (4) in Theorem 1.4 are not true as shown by the following example.

1.6. Example. Let $X = \{a, b, c\}$ and $w = \{\phi, \{a\}, \{b\}, \{c\}\}$. One may notice that:

(1) The subset $A = \{a, c\}$ satisfy $A \cap Bd_w(A) = \phi$, but A is not w-open,
(2) The subset $A = \{c\}$ satisfy $Bd_w(A) \subset A$, but A is not w-closed.

2. Generalized w-Closed and Generalized w-Open Sets

2.1. Definition. Let w be a WS on X. We define the concepts of generalized closed and generalized open sets in weak structure as follows:

(1) A subset A is said to be generalized w-closed (gw-closed, for short) if $c_w(A) \subset U$, whenever $A \subset U$ and U is w-open.

(2) The complement of a generalized w-closed set is said to be generalized w-open (gw-open, for short).

The family of all gw-closed (resp. gw-open) sets in a weak structure X will be denoted by $gwC(X)$ (resp. $gwO(X)$)

2.2. Theorem. *Let w be a WS on X. A subset A is gw-open if and only if $i_w(A) \supseteq F$, whenever $A \supseteq F$ and F is w-closed.*

Proof. It follows from Theorem 1.1 and the fact the complement of w-open set is w-closed. □

2.3. Remark. By the following two examples, we show that union and intersection of two gw-closed sets is not gw-closed.

2.4. Example. Let $X = \{a, b, c\}$ and $w = \{\emptyset, \{a\}\}$. If $A = \{a, b\}$ and $B = \{a, c\}$, then A and B are gw-closed sets but $A \cap B = \{a\}$ is not gw-closed set.

2.5. Example. Let $X = \{a, b, c, d\}$ and $w = \{\emptyset, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \{a, b\}\}$. Then $A = \{a\}$ and $B = \{c, d\}$ are gw-closed sets in X, but their union $A \cup B = \{a, c, d\}$ is not gw-closed.

2.6. Theorem. *Let w be a WS on X. If $\{A_i : i \in I\}$ is a family of subsets of X, then $c_w(\bigcup A_i) \supseteq \bigcup c_w(A_i)$.*

Proof. It is clear. □

2.7. Definition. Let w be a WS on X. A family $\{A_i : i \in I\}$ is said to be w-locally finite if $c_w(\bigcup A_i) = \bigcup c_w(A_i)$.

2.8. Theorem. *Let w be a WS on X. The arbitrary union of gw-closed sets $A_i, i \in I$ in X is a gw-closed set if the family $\{A_i : i \in I\}$ is w-locally finite.*

Proof. Let w be a WS on X, let $\{A_i : i \in I\}$ be a family of gw-closed sets in X and U be a w-open set such that $\bigcup A_i \subset U$. Then $A_i \subset U$ for each $i \in I$ and hence $c_w(A_i) \subset U$ which implies $\bigcup c_w(A_i) \subseteq U$. Since the family $\{A_i : i \in I\}$ is w-locally finite, then $c_w(\bigcup A_i) = \bigcup c_w(A_i) \subseteq U$. Therefore $\bigcup A_i$ is gw-closed. □

2.9. Theorem. *Let w be a WS on X. The arbitrary intersection of gw-open sets $A_i, i \in I$ in X is a gw-open set if the family $\{A_i : i \in I\}$ is w-locally finite.*

Proof. It follows from Theorem 1.1 and Theorem 2.27 and the fact the complement of a gw-open set is a gw-closed. □

2.10. Theorem. *Let w be a WS on X. If A is a w-closed set, then A is gw-closed.*

Proof. Let A be a w-closed set and U be a w-open set in X such that $A \subset U$. Then $c_w(A) = A \subset U$ and hence A is gw-closed. □

2.11. Corollary. *Let w be a WS on X. If A is a w-open set, then A is gw-open.*

2.12. Remark. By the following example, we show that the converse of Theorem 2.10 need not be true in general.

2.13. Example. In Example 2.5, if $A = \{d\}$, then A is gw-closed and not w-closed.

2.14. Theorem. *Let w be a WS on X. If A is a gw-closed set in X, then $c_w(A) - A$ contains no non empty w-closed.*

Proof. Suppose that F is a non empty w-closed subset of $c_w(A)-A$. Now $F \subset c_w(A)-A$. Then $F \subset c_w(A) \bigcap X - A$ and hence $F \subset c_w(A)$ and $F \subset X - A$. Since $X - F$ is w-open and A is gw-closed, then $c_w(A) \subset X - F$ and hence $F \subset X - c_w(A)$. Thus $F \subset c_w(A) \cap X - c_w(A) = \phi$ and hence $F = \phi$. Therefore $c_w(A) - A$ does not contain non empty w-closed. □

2.15. Remark. In general topology, Levine [6] proved that the above theorem is true for "if and only if". But in the weak structures the converse of the above theorem need not be true in general as shown by the following example.

2.16. Example. Let $X = \{a, b, c\}$ and $w = \{\phi, \{b\}, \{c\}\}$. One may notice that if $A = \{b\}$, then $c_w(A) - A = \{a, b\} - \{b\} = \{a\}$ does not contain any non empty w-closed, but A is not a gw-closed set in X, since A is an w-open set contains itself and $c_w(A) = \{a, b\} \nsubseteq A$.

2.17. Corollary. *Let w be a WS on X and $A \subseteq X$ is a gw-closed set. If $c_w(A) - A$ is w-closed, then $c_w(A) = A$.*

Proof. Let $c_w(A)-A$ be w-closed and A be a gw-closed set in X. Then by Theorem 2.14, $c_w(A) - A$ contains no non empty w-closed set. Since $c_w(A) - A$ is a w-closed subset of itself, $c_w(A) - A = \phi$ and hence $c_w(A) = A$. □

2.18. Remark. If A is a gw-closed set in a WS on X and $c_w(A) = A$, then $c_w(A) - A$ is need not be w-closed as shown by the following example.

2.19. Example. Let $X = \{a, b, c\}$, $w = \{\phi, \{a\}, \{c\}, \{a, b\}\}$ and $A = \{b\}$. One may notice that $c_w(A) = A$ and hence $c_w(A) - A = \phi$, which is not w-closed.

2.20. Theorem. *Let w be a WS on X. Then $A \subseteq X$ is a gw-closed if $c_w(\{x\}) \bigcap A \neq \phi$ for each $x \in c_w(A)$.*

Proof. Let $c_w(\{x\}) \bigcap A \neq \phi$ for each $x \in c_w(A)$ and U be any w-open set with $A \subseteq U$. Let $x \in c_w(A)$. Then $c_w(\{x\}) \bigcap A \neq \phi$ and hence there exists $y \in c_w(\{x\}) \bigcap A$, so $y \in A \subseteq U$. Thus $\{x\} \bigcap U \neq \phi$ and hence $x \in U$. Therefore $c_w(A) \subseteq U$, which implies A is gw-closed. □

2.21. Remark. Al Omari and Noiri [1, Theorem 2.9] proved that the converse of the above theorem is true. The following example shows that the converse needn't be true generally.

2.22. Example. Let $X = \{a, b, c\}, w = \{\phi, \{a\}, \{b\}\}$. one may notice that $A = \{a, b\}$ is gw-closed and $c_w(\{c\}) = \{c\}$. So $A \bigcap c_w(\{c\}) = \phi$.

2.23. Theorem. *Let w be a WS on X. If A is a gw-closed set in X, then $c_w(A) - A$ is gw-open.*

Proof. Let A is a gw-closed set in X and F be a w-closed subset such that $F \subset c_w(A)-A$. Then by Theorem 2.14 we have $F = \phi$ and hence $F \subset i_w(c_w(A) - A))$. So by Theorem 2.2, we have $c_w(A) - A$ is gw-open. □

2.24. Remark. In topological space, Levine [6] proved that the above theorem is true for "if and only if". But in the weak structures the converse of the above theorem need not be true in general as shown by the following example.

2.25. Example. Let $X = \{a, b, c\}$, $w = \{\phi, \{a\}, \{c\}, \{a, b\}\}$ and $A = \{a\}$. One may notice that $c_w(A) - A = \{a, b\} - \{a\} = \{b\}$ which is gw-open, but A is not a gw-closed set, since A is a generalized w-open set contain itself, but $c_w(A) = \{a, b\} \not\subset A$

2.26. Theorem. *Let w be a WS on X and A be a gw-closed set with $A \subset B \subset c_w(A)$, then B is gw-closed.*

Proof. Let H be a w-open set in X such that $B \subset H$, then $A \subset H$. Since A is gw-closed, then $c_w(A) \subset H$ and hence $c_w(B) \subset c_w(A) \subset H$. Thus B is gw-closed. □

2.27. Theorem. *Let w be a WS on X and A be a gw-closed set with $A \subset B \subset c_w(A)$, then $c_w(B) - B$ contains no non empty w-closed.*

Proof. It follows from Theorems 2.14 and 2.26. □

2.28. Remark. Let w be a WS on X and A be a gw-open set with $i_w(A) \subset B \subset A$, then B is gw-open.

2.29. Remark. Let w be a WS on X. Then each subset of X is gw-closed if each w-open set is w-closed.

2.30. Remark. In topological space, Levine [6] proved that the above theorem is true for "if and only if". But in the weak structures the converse of the above theorem need not be true in general as shown by the following example.

2.31. Example. Let $X = \{a, b, c\}$, $w = \{\phi, \{a\}, \{b\}, \{c\}, \{a, c\}, \{b, c\}, X\}$. One may notice that every subset of X is gw-closed, but $A = \{c\}$ is w-open set in X and it is not w-closed.

2.32. Remark. Let w be a WS on X. Then each subset of X is gw-closed if and only if $c_w(A) \subseteq A$ for each w-open set A in X.

2.33. Theorem. *Let w be a WS on X. If A is a gw-open set in X, then $U = X$ whenever U is w-open and $i_w(A) \cup (X - A) \subset U$.*

Proof. Let U be a w-open set in X and $i_w(A) \cup (X - A) \subset U$ for any gw-open set A. Then $X - U \subset [X - i_w(A)] \cap A$ and hence $X - U \subset (c_w(X - A)) - (X - A)$. Since $X - A$ is a gw-closed, then by Theorem 2.14, we have $X - U = \phi$ and hence $U = X$. □

3. Separation Axioms on Weak Structures

3.1. Definition. Let w be a WS on X. We define the concepts of strongly generalized closed and strongly generalized open sets in weak structure as follows:

(1) A subset A is said to be strongly generalized w-closed (sgw-closed, for short) if $c_w(A) \subseteq U$, whenever $A \subseteq U$ and U is gw-open.
(2) A subset A is said to be mildly w-closed (mgw-closed, for short) if $c_w(i_w((A)) \subseteq U$, whenever $A \subseteq U$ and U is gw-open.
(3) The complement of a sgw-closed (resp. mgw-closed) set is said to be sgw-open (resp. mgw-open).

3.2. Definition. A weak structure w on X is said to be $w - T_{\frac{1}{2}}$ if each gw-closed set A of X, $c_w(A) = A$.

3.3. Remark. (1) In a topological space X, X is $T_{\frac{1}{2}}$ [6] if and only if each singleton is either closed or open. By the following examples we show that "if X is a weak structure and each singleton is w-open or $c_w(A) = A$, then X need not be $w - T_{\frac{1}{2}}$".
(2) We think that in a weak structure X, if X is $w - T_{\frac{1}{2}}$, then there exists a singleton $x \in X$ such x is neither w-closed nor $\{x\} \neq i_w\{x\}$.

3.4. Example. Let $X = \{a, b, c\}$, $w = \{\phi, \{b\}, \{c\}, \{b, c\}\}$. One may notice that each singleton is w-open or w-closed. But there exists $A = \{a, b\}$ which is gw-closed and $c_w(A) = X \neq A$. So X is not $w - T_{\frac{1}{2}}$.

3.5. Theorem. *Let w be a WS on X. If $i_w\{x\}$ is an w-open set and each singleton is either w-closed or $\{x\} = i_w\{x\}$, then X is $w - T_{\frac{1}{2}}$.*

Proof. Let A be a gw-closed subset of X and $x \in c_w(A)$.

Case 1. If $\{x\}$ is w-closed and $x \notin A$, then $x \in (c_w(A) - A)$ and hence $\{x\} \subseteq X - A$, which implies $A \subseteq X - \{x\}$. Since A is a gw-closed set and $X - \{x\}$ is an w-open set, then $c_w(A) \subseteq X - \{x\}$ and hence $\{x\} \subseteq X - c_w(A)$. Therefore $\{x\} \subseteq c_w(A) \bigcap X - c_w(A) = \phi$, which is a contradiction. Thus $x \in A$ and hence $c_w(A) = A$.

Case 2. If $\{x\} = i_w\{x\}$ and $x \in c_w(A)$, then for each w-open set V with $x \in V$, we have $V \bigcap A \neq \phi$. Since $i_w\{x\}$ is an w-open set and $\{x\} = i_w\{x\}$, then $\{x\} \bigcap A \neq \phi$ and hence $x \in A$.Thus $c_w(A) = A$. Therefore in the two cases we have $c_w(A) = A$ and hence X is $w - T_{\frac{1}{2}}$. □

3.6. Definition. A weak structure w on X is said to be $w - T_1$ if for any points $x, y \in X$ with $x \neq y$, there exist two w-open sets U and V such that $x \in U$, $y \notin U$, $x \notin V$ and $y \in V$.

3.7. Theorem. *A weak structure w on X is $w - T_1$ if every singleton in X is w-closed.*

Proof. It is clear. □

3.8. Remark. In a topological space one may notice that:

(1) The above theorem is true for if and only if,
(2) If X is T_1, then each g-closed set in X is closed.

By the following example we show that the converse of the above theorem (the second part of item 1 above) need not be true and the item 2 above need not be true too in an WS on X in general.

3.9. Example. Let $X = \{a, b, c\}$, $w = \{\phi, \{a\}, \{b\}, \{c\}\}$. One may notice that:

(1) w is $w - T_1$, but the singleton $\{b\}$ is not w-closed.
(2) w is $w - T_1$ and the singleton $\{b\}$ is gw-closed, but is not w-closed.

3.10. Definition. A weak structure w on X is said to be:

(1) w-normal if for each two w-closed sets F and H with $F \bigcap H = \phi$, there exist two w-open sets U and V such that $F \subseteq U$, $H \subseteq V$ and $U \bigcap V = \phi$.
(2) Almost w-normal if for each w-closed set F and $H \in rc(w)$ with $F \bigcap H = \phi$, there exist two w-open sets U and V such that $F \subseteq U$, $H \subseteq V$ and $U \bigcap V = \phi$.
(3) Weakly w-normal if for each $F, H \in rc(w)$ with $F \bigcap H = \phi$, there exist two w-open sets U and V such that $F \subseteq U$, $H \subseteq V$ and $U \bigcap V = \phi$.

3.11. Theorem. *Let w be a WS on X. Consider the following statements:*

(1) *X is w-normal;*
(2) *For each w-closed set F and w-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq U$;*
(3) *For each w-closed set F and each gw-closed set H with $F \bigcap H = \phi$, there exist two w-open sets U and V such that $F \subseteq U, H \subseteq V$ and $U \bigcap V = \phi$;*
(4) *For each w-closed set F and gw-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq U$.*

Then the implications (1) ⇒ (2) *and* (3) ⇒ (4) ⇒ (2) *are hold.*

Proof. It is clear. □

3.12. Theorem. *Let w be a WS on X. If $c_w(A)$ is w-closed for each w-open or gw-closed, then the statements in Theorem 3.11 are equivalent.*

Proof. From Theorem 3.11 we need to prove (2) ⇒ (1) and (1) ⇒ (3) only.

(2) ⇒ (1): Let A and B be two disjoint w-closed subsets of X. Then $X - B$ is an w-open set containing A. Thus by (2) there exists an w-open set U such that $A \subseteq U \subseteq c_w(U) \subseteq X - B$ and hence $A \subseteq U$ and $B \subseteq X - c_w(U)$. Since $c_w(U)$ is w-closed for each w-open set U, then $X - c_w(U) = V$ is w-open and $U \bigcap V = \phi$. Hence X is w-normal.

(1) ⇒ (3). Let F be an w-closed set and H be a gw-closed set with $F \bigcap H = \phi$. Then $H \subseteq X - F$ which is w-open. Since H is gw-closed and $H \subseteq X - F$, then $c_w(H) \subseteq X - F$. Since H is gw-closed, then $c_w(H)$ is w-closed. By (1) there exist two w-open sets U and V such that $c_w(H) \subseteq U$, $F \subseteq V$ and $U \bigcap V = \phi$. Hence $H \subseteq U$, $F \subseteq V$ and $U \bigcap V = \phi$. □

3.13. Theorem. *Let w be a WS on X. Consider the following statements:*

(1) *X is almost w-normal;*
(2) *For each w-closed set F and $U \in r(w)$ with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$;*
(3) *For each w-closed set F and mgw-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$;*
(4) *For each w-closed set F and sgw-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$;*
(5) *For each w-closed set F and w-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$.*

Then the implications (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (5) *are hold.*

Proof. (1) ⇒ (2): Obvious.

(2) ⇒ (3): Let F be an w-closed set and U be a mgw-open with $F \subseteq U$. Then $F \subseteq i_w(C_w(U)) \in r(w)$. By (2) there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(i_w(c_w(U)))) = i_w(c_w(U))$.

(3) ⇒ (4) ⇒ (5): Obvious. □

3.14. Theorem. *Let w be a WS on X. If $c_w(A)$ is w-closed for each w-open A, then the statements in Theorem 3.13 are equivalent.*

Proof. From Theorem 3.13 we need to prove that (5) ⇒ (1) only.

(5) ⇒ (1): Let F be an w-closed set and $H \in rc(w)$ with $F \bigcap H = \phi$. Then $F \subseteq X - H = i_w(c_w(X - H))$. By (5) there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(i_w(c_w(X-H)))) = i_w(c_w(X-H))$ and hence $F \subseteq V, H = c_w(i_w(H)) \subseteq X - c_w(V)$. Since V is an w-open, then $c_w(V)$ is w-closed and hence $X - c_w(V) = W$ which is w-open contains H. Thus $V \bigcap W = \phi$. Therefore X is almost w-normal. □

3.15. Theorem. *Let w be a WS on X. Consider the following statements:*

(1) *X is almost w-normal;*
(2) *For each w-open set U and $F \in rc(w)$ with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq U$;*
(3) *For each mgw-closed set F and w-open U with $F \subseteq U$, there exist w-open sets V such that $c_w(i_w(F)) \subseteq V \subseteq c_w(V) \subseteq U$;*
(4) *For each gw-closed set F and w-open U with $F \subseteq U$, there exist w-open sets V such that $c_w(i_w(F)) \subseteq V \subseteq c_w(V) \subseteq U$.*

Then the implications (1) ⇒ (2) ⇒ (3) *and* (2) ⇒ (4) *are hold.*

Proof. (1) ⇒ (2): Obvious.

(2) ⇒ (3): Let F be a mgw-closed set and U be a w-open with $F \subseteq U$. Then $c_w(i_w(F)) \subseteq U$. Since $c_w(i_w(F)) \in rc(w)$, then by (2) there exists an w-open set V such that $c_w(i_w(F)) \subseteq V \subseteq c_w(V) \subseteq U$.

(2) ⇒ (4): Let F be a gw-closed set and U be a w-open with $F \subseteq U$. Then $c_w(F) \subseteq U$ and hence $c_w(i_w(F)) \subseteq U$. Since $c_w(i_w(F)) \in rc(w)$, then by (2) there exists an w-open set V such that $c_w(i_w(F)) \subseteq V \subseteq c_w(V) \subseteq U$. □

3.16. Theorem. *Let w be a WS on X. If $c_w(A)$ is w-closed for each w-open set A or $A \in r(w)$, then the statements in Theorem 3.15 are equivalent.*

Proof. From Theorem 3.15 we need to prove that (3) ⇒ (1) and (4) ⇒ (1) only.

(3) ⇒ (1): Let F be an w-closed set and $H \in rc(w)$ with $F \bigcap H = \phi$. Then $H \subseteq X - F$. Since $H \in rc(w)$, then H is mgw-closed. By (3) there exist w-open sets V such that $H = c_w(i_w(H)) \subseteq V \subseteq c_w(V) \subseteq X - F$ and hence $H \subseteq V$ and $F \subseteq X - c_w(V) = W$ which is w-open. Thus there exist two w-open sets V and W such that $H \subseteq V, F \subseteq W$ and $V \bigcap W = \phi$. Therefore X is almost w-normal.

(4) ⇒ (1): Let F be an w-closed set and $H \in rc(w)$ with $F \bigcap H = \phi$. Then $H \subseteq X - F$. Since $H \in rc(w)$, then $c_w(i_w(H)) \subseteq X - F$. Since $H \in rc(w)$, then $i_w(H)$ is an w-open and hence $c_w(i_w(H))$ is w-closed which is gw-closed. By (4) there exist w-open sets V such that $c_w(i_w(c_w(i_w(H))) \subseteq V \subseteq c_w(V) \subseteq X - F$ and hence $H \subseteq V$ and $F \subseteq X - c_w(V) = W$ which is w-open. Thus there exist two w-open sets V and W such that $H \subseteq V, F \subseteq W$ and $V \bigcap W = \phi$. Therefore X is almost w-normal. □

3.17. Theorem. *Let w be a WS on X. Consider the following statements:*

(1) *X is weakly w-normal,*
(2) *For each $F \in rc(w)$ and $U \in r(w)$ with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$,*
(3) *For each $F \in rc(w)$ and mgw-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$,*
(4) *For each $F \in rc(w)$ and sgw-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$,*
(5) *For each $F \in rc(w)$ and w-open U with $F \subseteq U$, there exist w-open sets V such that $F \subseteq V \subseteq c_w(V) \subseteq i_w(c_w(U))$.*

Then the implications (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (5) *are hold.*

Proof. It is similar to that of Theorem 3.13. □

3.18. Theorem. *Let w be a WS on X. If $c_w(A)$ is w-closed for each w-open A, then the statements in Theorem 3.17 are equivalent.*

Proof. It is similar to that of Theorem 3.14. □

References

[1] A. Al-Omari and T. Noiri, *A unified theory of generalized closed sets in weak structure*, Acta Math. Hungar. **135** (2012), no. 1-2, 174–183.
[2] A. Császár, *Generalized topology, generalized continuity*, Acta Math. Hungar. **96** (2002), 351–357.
[3] ———, *Weak structure*, Acta Math. Hungar. **131** (2011), no. 1-2, 193–195.
[4] E. Ekici, *On weak struture due to Császár*, Acta Math. Hungar. **134** (2012), no. 4, 565–570.

[5] H. Maki J. Umehara and T. Noiri, *Every topological space is pre* $T_{\frac{1}{2}}$, Mem. Fac. Sci. Kochi. Univ. Ser. A Math. **17** (1996), 53–66.
[6] N. Levine, *Generalized closed sets in topology*, Rend. Circ. Mat. Palermo **2** (1970), 89–96.
[7] J. K. Park and J. H. Park, *Mildly generalized closed sets, almost normal and mildly normal space*, Chaos, Solitoen and Fractals **20** (2004), 1103–1111.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 679 – 687

Using of fractional factorial design (r^{k-p}) in data envelopment analysis to selection of outputs and inputs

Hulya Bayrak*, Oday Jarjies† and Kubra Durukan‡§

Abstract

Data envelopment analysis (DEA) is a linear programming based technique for measuring the relative performance of organisational units where the presence of multiple inputs and outputs makes comparisons difficult. We used, Morita and Avkiran propose after it has been developed an input-output selection method that uses fractional factorial design, which is a statistical approach to find an optimal combination. Energy efficiency and greenhouse gas emissions are closely linked in the last two decades. We demonstrate the proposed method using data that increase energy efficiency and heating gas emissions in the European Union (EU) countries.

2000 AMS Classification: 47N30, 62H30

Keywords: Data Envelopment Analysis (DEA), Decision-Making Unit (DMU), Fractional Factorial Design, Mahalanobis Fistance (MD).

Received 06/08/2013 : Accepted 30/07/2014 Doi : 10.15672/HJMS.2014247475

*Gazi University, Science Faculty, Statistics Department, Ankara, TURKEY,
Email: hbayrak@gazi.edu.tr
†Department Operations Research and Technology Intelligence Mosul University, IRAQ
Email: oday_alubade@yahoo.com
‡Kirikkale University, Faculty of Arts and Sciences, Department of Statistics, 71100, Kampus, Yahsihan, Kirikkale - TURKEY,
Email: kubraba@yahoo.com
§Corresponding Author.

1. Introduction

Data envelopment analysis (DEA), introduced by Charnes, Cooper and Rhodes (CCR) [1], is a mathematical programming method for measuring the relative efficiency of decision-making units (DMUs) with multiple inputs and outputs. Most models DEA has the best performance and efficiency to determine the degree of expertise and decision-making units (DMUs). Differentiating efficient DMUs is an interesting research area. The original DEA method evaluates each DMU against a set of efficient DMUs and cannot identify which efficient DMU is a better option with respect to the inefficient DMU. This is because all efficient DMUs have an efficiency score of one. Authors have proposed methods for ranking the best performers, for instance using super-efficiency DEA model.

In this paper, in order to rank DMUs, we use the evaluation contexts that are obtained by partitioning the set of DMUs into several levels of efficiency, and rank all DMUs with two criteria: the high and low performers. The influence of all DMUs, both efficient and inefficient, in ranking is this method's preference.

1.1. Data Envelopment Analysis. Consider n decision making units ($DMUj, j = 1, \dots, n$) in which each DMU consumes input levels $x_{ij}(i = 1, ..., m)$ to produce output levels $y_{rj}(j = 1, ..., s)$. Suppose that $x_j = (x_{1j}, ..., x_{mj})^T$ and $y_j = (y_{1j}, ..., y_{sj})^T$ are the vectors of inputs and outputs values respectively, the relative efficiency score of the DMU_O, $O \in \{1, ..., n\}$ is obtained from the following model which is called input-oriented CCR envelopment model [7, 8, 9]

$$(1.1) \qquad \begin{aligned} &\theta_O^* = \min \theta \\ s.t. &\sum_j \lambda_j x_{ij} \le \theta x_{io}, i = 1, ..., m \\ &\sum_j \lambda_j y_{ij} x_{ij} \ge y_{ro}, r = 1, ..., s \\ &\lambda_j \ge 0, j = 1, ..., n \end{aligned}$$

This model is an input oriented constant returns to scale (CRS) model. The efficiency of DMUo is determined from efficiency score θ_O^* and its slack values. If and only if $\theta_O^* = 1$ there is no slack, DMUo is said to be efficient. If and only if $\theta_O^* = 1$ there are non-zero slacks, DMUo is inefficient and we can called it a weak-efficient. The weak-efficient DMUs and efficient DMUs comprise the efficient frontier [6].

Morita and Haba in a previous study, select the output of the of preference between the two groups based on public information and previous experience has nothing to do with data where they are exploiting the experience of planning two-level orthogonal and optimal variables can be found statistically. On the other hand, Ediridsinghe and Zhang proposed DEA generalized approach to determine the input and output by maximizing the correlation coefficient between the DEA and the result of external performance indicator. Morita and Avkiran propose an input output selection method that uses diagonal layout experiments and demonstrate the proposed method using financial statement data from NIKKEI 500 index. They utilize a two-step heuristic algorithm that combines random sampling and local search to find an optimal combination of inputs and outputs [5, 4].

In this paper, we show the method of selection of inputs and outputs based on an analysis using the Mahalanobis distance of difference between the two group of data. We use 3-level orthogonal layout experiment to find a suitable combination of inputs and outputs, where trials are independent of each other.

2. 3^{k-p} Fractional Factorial Designs and Selecting Input and Output Variables:

The whole point of looking at this structure is because sometimes we want to only conduct a fractional factorial. We sometimes can't afford 27 runs. Often we can only afford a fraction of the design. So, let's construct a 3^{3-1} design which is a $1/3$ fraction of a 3^3 design. In this case, $N = 3^{3-1} = 3^2 = 9$, the total number of runs. This is a small, compact design [2].

We again start out with a 3^3 design which has 27 treatment combinations and assign them to 3 blocks. What we want to do in this part, going beyond the 3^2 design, is to describe the ANOVA for this 3^3 design. Then we also want to look at the connection between confounding in blocks and 3^{k-p} fractional factorials, See Appendix 1.

2.1. 3-level Full Factorial Designs and Other Factorials. The 3-level design is written as a 3^k factorial design. It means that k factors are considered, each at 3-levels. These are (usually) referred to as low, intermediate and high levels. These levels are numerically expressed as $0, 1$ and 2. One could have considered the digits $-1, 0$ and $+1$, but this may be confusing with respect to the 2-level designs since 0 is reserved for center points. Therefore, we will use the $0, 1, 2$ scheme. The reason that the 3-level designs were proposed is to model possible curvature in the response function and to handle the case of nominal factors at 3-levels. A third level for a continuous factor facilitates investigation of a quadratic relationship between the response and each of the factors [2].

Unfortunately, the 3-level design is prohibitive in terms of the number of runs, and thus in terms of cost and effort. For example a 2-level design with center points is much less expensive while it still is a very good (and simple) way to establish the presence or absence of curvature. Table 1 shows us the difference between full factorial designs and other factorials.

Table 1. 3-level designs

Factors	3	4	**5**	6	7
Full	27	81	**243**	729	2187
1/3	9	27	**81**	243	729
1/9	3	9	**27**	81	243
1/27	NA	NA	**9**	27	81

3. Mahalanobis Distance

In statistics, Mahalanobis distance (MD) is a distance measure introduced by P.C. Mahalanobis in 1936. It is based on correlations between variables by which different patterns can be identified and analyzed. It gauges similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant. In other words, it is a multivariate effect size [3].

Formally, the MD of a multivariate vector $x = (x_1, x_2, ..., x_N)^T$ from a group of values with mean vector $\mu = (\mu_1, \mu_2, ..., \mu_N)^T$ and covariance matrix Σ is defined as [4, 12]:

$$D_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)} \tag{3.1}$$

Equation (3.1) is rewritten for the sample following as

$$\hat{D}_M(x) = \sqrt{\left(x-\overline{x}\right)^T S^{-1}\left(x-\overline{x}\right)}$$

where the mean vector and covariance matrix of the sample are given as $\overline{x}$ and S respectively.

MD is widely used in cluster analysis and classification techniques. It is closely related to Hotelling's T-square distribution used for multivariate statistical testing and Fisher's Linear Discriminant Analysis that is used for supervised classification [12].

In order to use the MD to classify a test point as belonging to one of N classes, one first estimates the covariance matrix of each class, usually based on samples known to belong to each class. Then, given a test sample, one computes the MD to each class, and classifies the test point as belonging to that class for which the MD is minimal [3].

MD and leverage are often used to detect outliers, especially in the development of linear regression models. A point that has a greater the MD from the rest of the sample population of points is said to have higher leverage since it has a greater influence on the slope or coefficients of the regression equation. MD is also used to determine multivariate outliers. Regression techniques can be used to determine if a specific case within a sample population is an outlier via the combination of two or more variable scores. A point can be a multivariate outlier even if it is not a univariate outlier on any variable [7, 8].

3.1. MD Threshold Selection. The MD threshold is another important element of prognostics analysis. An MD threshold value which is either too large or too small leads to false negatives or false positives, respectively. In this study, we consider the distance of one-dimensional variables, where MD coincides with the Welch statistics [5]. The Welch statistics is given as

$$\hat{d} = \frac{\overline{x}_h - \overline{x}_l}{\sqrt{\frac{S_h^2}{n_h} + \frac{S_l^2}{n_l}}} \tag{3.2}$$

where $\overline{x}_h, S_h^2$ and n_h are the sample mean, sample variance and sample size of high group, respectively. Also $\overline{x}_l$, S_l^2 and n_l are the sample mean, sample variance and sample size of low group, respectively.

For example, in run No. 5, x_1, x_{11}, x_{12} variables are selected as an input, variables $x_2, x_3, x_4, x_5, x_6, x_7$ are selected as an output; and variables x_8, x_9, x_{10} are not selected as an input or an output. Based on the fractional factorial design in Appendix 1, we calculate the efficiency scores and MD between the two groups using selected inputs and outputs. Where "1" means that the variable is selected as an input, "2" means that the variable is selected as an output, and "3" means that the variable is not selected.

The ANOVA table for the fractional factorial design appears in Table 2. The sum of squares and the degrees of freedom are given as

$$S_T = \sum_{i=1}^{27}\left(\hat{d}_i - \overset{\overline{\wedge}}{d}\right)^2, \quad df_T = 26 \tag{3.3}$$

$$S_i = 3\left[\overset{\overline{\wedge}}{d}^2(x_i=1) + \overset{\overline{\wedge}}{d}^2(x_i=2) + \overset{\overline{\wedge}}{d}^2(x_i=3)\right] - 27\overset{\overline{\wedge}}{d}^2, \quad df_i = 2, \quad i = 1, ..., 12 \tag{3.4}$$

$$S_E = S_T - (S_1 + S_2 + ... + S_{12}) \tag{3.5}$$

Table 2. ANOVA table for fractional factorial design of 3^{12-9}

Variables	SS	Df	MS	F Statistics
X_1	S_1	2	$V_1 = S_1/2$	V_1/V_E
X_2	S_2	2	$V_2 = S_2/2$	V_2/V_E
X_3	S_3	2	$V_3 = S_3/2$	V_3/V_E
X_4	S_4	2	$V_4 = S_4/2$	V_4/V_E
X_5	S_5	2	$V_5 = S_5/2$	V_5/V_E
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
X_{12}	S_{12}	2	$V_{12} = S_{12}/2$	V_{12}/V_E
Error	S_E	2	$V_E = S_E/2$	
Total	S_T	26		

where $\bar{\hat{d}}\,(x_i = 1)$ is the mean of the Mahalanobis distances observed when $(x_i = 1)$. The null hypothesis that the candidate has no effect as an input or output is tested by using the F statistics

$$F = \frac{S_i/df_i}{S_E/df_E} \tag{3.6}$$

and hypotesis tests is as following:
H0: The variable candidate has no effect on output and input.
H1: The variable candidate has effect on output and input.

This results in the optimal combination of input and output variables. The following is a summary procedure for the selection of variables.

Step 1. Choose a list of data envelope (DEA), which contains the input and output variables are possible.
Step 2. The use of external standards to distinguish between the performance of the two groups. For example, the high and low performance.
Step 3. To create a table perpendicular to try to set the input and output variables that are not determined.
Step 4. Calculate MD between the two groups using Welch statistics.
Step 5. Determine the optimal mix of input and output variables based on the results Analysis of variance.
Step 6. Determine the optimal variables are statistically significant either input or output using sum of MD.
Step 7. We use DEA model (1) with the data that have been selected from the output and input.

4. A Case Study Using Greenhouse Gas Emissions Intensity of Energy Consumption Data

We used the greenhouse gas intensity of energy consumption that is the ratio between energy-related greenhouse gas emissions (carbon dioxide, methane and nitrous oxide) and gross inland energy consumption for EU countries.

There are key factors leading to greenhouse emissions: Electricity production, Transportation, Industry, Commercial and Residential, Agriculture and Land Use and Forestry [10, 11].

The table in the Appendix 2 shows a part of the data set all variables have large rangs. In Step 1 the following twelve variables are collected to evaluate the managerial performance, that is, the standard deviation is greater than the mean.

A. Total emissions.
B. Total net emissions.
C. Energy.
D. Energy industries.
E. Manufacturing industries and construction.
F. Transport.
G. Road transportation.
H. Other sectors.
I. Industrial processes.
J. Solvent and other product use.
K. Agriculture.
L. Waste.

In Step 2, we construct two groups, high-performers and low-performers, the table in the Appendix 3 shows the mean and standard deviation for each variable. When we select the variables to capture the difference between high-performers and low-performers, we choose a variable with a large difference between these two groups. MD between the 15 high cuntery and 15 low cuntery for each variable is also shown in the Appendix 3, where we find that (A),(B),(C),(E), (I) and (L) have a large d and may be intuitively selected as inputs or outputs.

In Step 3, we assign 12 factors into a 3-level orthogonal layout, where at least 27 runs are required. That is, we utilize the fractional factorial design 3^{12-9}. Table 3 shows the selected variable combinations for efficiency score calculation. The MD for each experiment is calculated in Step 4, which is also shown in the last column of Table 3.

Table 3. Selected inputs and outputs and MD

Runs	A	B	C	D	E	F	G	H	I	J	K	L	Selected Input	Selected Output	Not Selected	$\hat{d}$
1	1	1	1	1	1	1	1	1	1	1	1	1	A,B,C,D,E,F,G,H,I,J,K,L	None	None	2.44
2	1	1	1	1	2	2	2	2	2	2	2	2	A,B,C,D	E,F,G,H,I,J,K,L	None	1.19
3	1	1	1	1	3	3	3	3	3	3	3	3	A,B,C,D	None	E,F,G,H,I,J,K,L	1.4
4	1	2	2	2	1	1	1	2	2	2	3	3	A,E,F,G	B,C,D,H,I,J	K,L	0
5	1	2	2	2	2	2	2	3	3	3	1	1	A,K,L	B,C,D,E,F,G	H,I,J	-0.13
6	1	2	2	2	3	3	3	1	1	1	2	2	A,H,I,J	B,C,D,K,L	E,F,G	-0.35
7	1	3	3	3	1	1	1	3	3	3	2	2	A,E,F,G	K,L	B,C,D,H,I,J	1.18
8	1	3	3	3	2	2	2	1	1	1	3	3	A,H,I,J	E,F,G	B,C,D,K,L	0.55
9	1	3	3	3	3	3	3	2	2	2	1	1	A,K,L	H,I,J	B,C,D,E,F,G	0.98
10	2	1	2	3	1	2	3	1	2	3	1	2	B,E,H,K	A,C,F,I,L	D,G,J	-0.26
11	2	1	2	3	2	3	1	2	3	1	2	3	B,G,J	A,C,E,H,K	D,F,I,L	-0.15
12	2	1	2	3	3	1	2	3	1	2	3	1	B,F,I,L	A,C,G,J	D,E,H,K	-0.46
13	2	2	3	1	1	2	3	2	3	1	3	1	D,E,J,L	A,B,F,H	C,G,I,K	-1.11
14	2	2	3	1	2	3	1	3	1	2	1	2	D,G,I,K	A,B,E,J,L	C,F,H	-0.87
15	2	2	3	1	3	1	2	1	2	3	2	3	D,F,H	A,B,G,I,K	C,E,J,L	-0.7
16	2	3	1	2	1	2	3	3	1	2	2	3	C,E,I	A,D,F,J,K	B,G,H,L	0.03
17	2	3	1	2	2	3	1	1	2	3	3	1	C,G,H,L	A,D,E,I	B,F,J,K	-0.24
18	2	3	1	2	3	1	2	2	3	1	1	2	C,F,J,K	A,D,G,H,L	B,E,I	-0.21
19	3	1	3	2	1	3	2	1	3	2	1	3	B,E,H,K	D,G,J	A,C,F,I,L	0.59
20	3	1	3	2	2	1	3	2	1	3	2	1	B,F,I,L	D,E,H,K	A,C,G,J	0.55
21	3	1	3	2	3	2	1	3	2	1	3	2	B,G,J	D,F,I,L	A,C,E,H,K	0.67
22	3	2	1	3	1	3	2	2	1	3	3	2	C,E,I	B,G,H,L	A,D,F,J,K	0.02
23	3	2	1	3	2	1	3	3	2	1	1	3	C,F,J,K	B,E,I	A,D,G,H,L	-0.28
24	3	2	1	3	3	2	1	1	3	2	2	1	C,G,H,L	B,F,J,K	A,D,E,I	-0.05
25	3	3	2	1	1	3	2	3	2	1	2	1	D,E,J,L	C,G,I,K	A,B,F,H	-0.8
26	3	3	2	1	2	1	3	1	3	2	3	2	D,F,H	C,E,J,L	A,B,G,I,K	-0.08
27	3	3	2	1	3	2	1	2	1	3	1	3	D,G,I,K	C,F,H	A,B,E,J,L	-0.65

Table 4 shows the analysis of variance for the data in Table 3, where we have pooled the negligible variables into the residual (Step 5). The level of significance is shown as the p value, where we find four variables (A,B,C, F, H and K) significant at the 5% level and their p values are very low, we leave them in the analysis for illustrative purposes.

Table 4. Table of ANOVA

Variables	Sum of Squares	Degrees of Freedom	Mean Squares	F Statistics	p value
A	4.32	2	2.16	11.37**	0.0837
B	3.78	2	1.39	7.32**	0.1265
C	1.58	2	0.79	4.16**	0.2109
F	1.90	2	0.95	5.00**	0.1790
H	1.20	2	0.60	3.16**	0.2662
K	1.40	2	0.70	3.68**	0.2340
Error	2.68	14	0.19		
Total	16.86	26			

Step 6, the final step in our procedure, generates Table 5 which shows the sum of MD for each variable at each level in Table 3. For example, when variable A is selected as an input, the sum of MD is 7.26, and when variable A is selected as an output, $\hat{d}$ is -3.97, it should be selected as an input. Maxima are indicated in bold font in Table 5. Thus we select four input (A) Total emissions, (B) Total net emissions, (C) Energy and (F) Transport and two outputs, namely, (H) Other sectors and (K) Agriculture.

Step 7, we run the DEA model (1) using this inputs and outputs combination.

Table 5. The sum of MD

Variables	Selected as Input	Selected as Output
A	7.26	-3.97
B	5.97	-3.47
C	4.30	-2.88
F	2.44	0.24
H	1.90	2.62
K	1.61	1.90

Note, we got 80% of the major factors leading to emissions of greenhouse gases That was previously displayed.

5. Conclusion

It is possible to attempt more than fractional factorial design at level 3 for example Latin square design or partial design. The MD and ANOVA was used to distinguish between the two groups after selecting the input and output from ANOVA results note it was investigating maximum MD between the two groups we demonstrate the effectiveness of this new approach using a case study with any DEA can set inputs and outputs and measuring the efficiency of performance that can effectively distinguish between groups of high and low performance.

Situation as you know it can not always be perfect, but is close to ideal combination that have been obtained are a limited number of 27 trials experience It can experimentation on a larger number of factors and a larger number of experiments.

References

[1] Charnes, A., Cooper, W.W. and Rhodes, E. Measuring the efficiency of decision making units. European Journal of Operational Research Vol. 2, 429-441, 1978.

[2] Box, G.E., Hunter, J.S., Hunter, W.G. Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition. Wiley, 2005.

[3] Jahanshahloo, G.R. and Hosseinzadeh, L.F. Introduction to Data Envelopment Analysis, (forth coming) 2004.

[4] Taguchi, G., Wu, Y. and Chowdhury, S. Mahalanobis-Taguchi System. McGraw-Hill Professional, 2000.

[5] Morita, H. and Haba, Y. Variable selection in data envelopment analysis based on external information. Proceedings of the eighth Czech-Japan Seminar 2005 on Data Analysis and Decision Making under Uncertainty, 181–187, 2005.

[6] Morita, H. and Avrikan, N.K. Selecting Input and Output Variables in Data envelopment analysis by designing statictcal experiments. Journal of the Operations Research Society of Japan, 52(2), 163-173, 2009.

[7] Edirisinghe, N.C.P. and Zhang, X. Generalized DEA model of fundamental analysis and its application to portfolio optimization. Journal of Banking & Finance, 31, 3311–3335, 2007.

[8] Seiford, L.M. and Zhu, J. Context-dependent data envelopment analysis: measuring attractiveness and progress. Omega, 31, 397–408, 2003.

[9] Cooper, W.W, Seiford, L.M. and Tone, K. Data Envelopment Analysis. Kluwer Academic Publishers, 2000.

[10] http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes.

[11] http://www.eia.gov/countries/data.cfm.

[12] http://en.wikipedia.org/wiki/Mahalanobis_distance.

Appendix 1. Fractional factorial design for 3^{12-9} twelve factors at three levels (27 Runs)

Run	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	Selected Input	Selected Output	Not selected	D
1	1	1	1	1	1	1	1	1	1	1	1	1	$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	None	None	D_1
2	1	1	1	1	2	2	2	2	2	2	2	2	x_1, x_2, x_3, x_4	$x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	None	D_2
3	1	1	1	1	3	3	3	3	3	3	3	3	x_1, x_2, x_3, x_4	None	$x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	D_3
4	1	2	2	2	1	1	1	2	2	2	3	3	x_1, x_5, x_6, x_7	$x_2, x_3, x_4, x_8, x_9, x_{10}$	x_{11}, x_{12}	D_4
5	1	2	2	2	2	2	2	3	3	3	1	1	x_1, x_{11}, x_{12}	$x_2, x_3, x_4, x_5, x_6, x_7$	x_8, x_9, x_{10}	D_5
6	1	2	2	2	3	3	3	1	1	1	2	2	x_1, x_8, x_9, x_{10}	$x_2, x_3, x_4, x_{11}, x_{12}$	x_5, x_6, x_7	D_6
7	1	3	3	3	1	1	1	3	3	3	2	2	x_1, x_5, x_6, x_7	x_{11}, x_{12}	$x_2, x_3, x_4, x_8, x_9, x_{10}$	D_7
8	1	3	3	3	2	2	2	1	1	1	3	3	x_1, x_8, x_9, x_{10}	x_5, x_6, x_7	$x_2, x_3, x_4, x_{11}, x_{12}$	D_8
9	1	3	3	3	3	3	3	2	2	2	1	1	x_1, x_{11}, x_{12}	x_8, x_9, x_{10}	$x_2, x_3, x_4, x_5, x_6, x_7$	D_9
10	2	1	2	3	1	2	3	1	2	3	1	2	x_2, x_5, x_8, x_{11}	$x_1, x_3, x_6, x_9, x_{12}$	x_4, x_7, x_{10}	D_{10}
11	2	1	2	3	2	3	1	2	3	1	2	3	x_2, x_7, x_{10}	$x_1, x_3, x_5, x_8, x_{11}$	x_4, x_6, x_9, x_{12}	D_{11}
12	2	1	2	3	3	1	2	3	1	2	3	1	x_2, x_6, x_9, x_{12}	x_1, x_3, x_7, x_{10}	x_4, x_5, x_8, x_{11}	D_{12}
13	2	2	3	1	1	2	3	2	3	1	3	1	x_4, x_5, x_{10}, x_{12}	x_1, x_2, x_6, x_8	x_3, x_7, x_8, x_{11}	D_{13}
14	2	2	3	1	2	3	1	3	1	2	1	2	x_4, x_7, x_9, x_{11}	$x_1, x_2, x_5, x_{10}, x_{12}$	x_3, x_6, x_8	D_{14}
15	2	2	3	1	3	1	2	1	2	3	2	3	x_4, x_6, x_8	$x_1, x_2, x_7, x_9, x_{11}$	x_3, x_5, x_{10}, x_{12}	D_{15}
16	2	3	1	2	1	2	3	3	1	2	2	3	x_3, x_5, x_9	$x_1, x_4, x_6, x_{10}, x_{11}$	x_2, x_7, x_8, x_{12}	D_{16}
17	2	3	1	2	2	3	1	1	2	3	3	1	x_3, x_7, x_8, x_{12}	x_1, x_4, x_5, x_9	x_2, x_6, x_{10}, x_{11}	D_{17}
18	2	3	1	2	3	1	2	2	3	1	1	2	x_3, x_6, x_{10}, x_{11}	$x_1, x_4, x_7, x_8, x_{12}$	x_2, x_5, x_9	D_{18}
19	3	1	3	2	1	3	2	1	3	2	1	3	x_2, x_5, x_8, x_{11}	x_4, x_7, x_{10}	$x_1, x_3, x_6, x_9, x_{12}$	D_{19}
20	3	1	3	2	2	1	3	2	1	3	2	1	x_2, x_6, x_9, x_{12}	x_4, x_5, x_8, x_{11}	x_1, x_3, x_7, x_{10}	D_{20}
21	3	1	3	2	3	2	1	3	2	1	3	2	x_2, x_7, x_{10}	x_4, x_6, x_9, x_{12}	$x_1, x_3, x_5, x_8, x_{11}$	D_{21}
22	3	2	1	3	1	3	2	2	1	3	3	2	x_3, x_5, x_9	x_2, x_7, x_8, x_{12}	$x_1, x_4, x_6, x_{10}, x_{11}$	D_{22}
23	3	2	1	3	2	1	3	3	2	1	1	3	x_3, x_6, x_{10}, x_{11}	x_2, x_5, x_9	$x_1, x_4, x_7, x_8, x_{12}$	D_{23}
24	3	2	1	3	3	2	1	1	3	2	2	1	x_3, x_7, x_8, x_{12}	x_2, x_6, x_{10}, x_{11}	x_1, x_4, x_5, x_9	D_{24}
25	3	3	2	1	1	3	2	3	2	1	2	1	x_4, x_5, x_{10}, x_{12}	x_3, x_7, x_9, x_{11}	x_1, x_2, x_6, x_8	D_{25}
26	3	3	2	1	2	1	3	1	3	2	3	2	x_4, x_6, x_8	x_3, x_5, x_{10}, x_{12}	$x_1, x_2, x_7, x_9, x_{11}$	D_{26}
27	3	3	2	1	3	2	1	2	1	3	1	3	x_4, x_7, x_9, x_{11}	x_3, x_6, x_8	$x_1, x_2, x_5, x_{10}, x_{12}$	D_{27}

Appendix 2. 30 Curenty Data Extract

	Country	A	B	C	D	E	F	G	H	I	J	K	L
1	Belgium	132459	131417	707608	26434	23708	24257	23459	32606	13460	214	10042	1135
2	Bulgaria	61427	52796	46438	31464	3798	7949	7508	2013	3852	46	6406	4686
3	Czech Republic	139158	133639	115205	56251	23807	17448	16979	12340	12061	503	7777	3612
4	Denmark	61065	58895	48789	23915	4453	13248	12241	6600	1685	76	9520	995
5	Germany	936544	953827	782313	354506	115007	154730	146844	145928	72569	1944	67479	12239
6	Estonia	20517	16759	18185	14644	510	2260	2078	646	498	18	1344	478
7	Ireland	61314	60284	40510	13328	4549	11606	11061	10996	1934	72	17910	889
8	Greece	118287	115645	93213	52522	6764	22996	19184	9839	10542	316	9282	4934
9	Spain	355898	326944	269835	72418	63434	91423	83872	39252	2817	2938	40014	15094
10	France	522373	490148	370022	61564	68440	132154	125273	103199	37509	1224	93876	19741
11	Italy	501318	444787	415727	133255	61375	118849	110018	90453	31963	1658	33741	18229
12	Cyprus	10838	10673	7505	3880	716	2313	2313	580	807	1	670	1857
13	Latvia	12077	−5070	8401	2261	1087	3222	2963	1729	639	42	2330	666
14	Lithuania	20810	9095	12848	5446	1120	4565	4125	1431	2249	93	4458	1161
15	Luxembourg	12075	11780	10647	1271	1401	6288	6273	1643	660	14	690	64
16	Hungary	67679	64307	49070	16668	3903	11870	11598	14270	6386	269	8267	3687
17	Malta	3035	2973	2657	1893	46	577	529	140	100	1	78	199
18	Netherlands	210053	213054	177819	66613	27326	34988	34242	45820	10432	170	16624	5008
19	Austria	84594	80983	64328	14293	15618	22452	21890	11401	10680	327	7453	1806
20	Poland	400865	357985	327117	173536	30969	48766	47151	61954	29769	779	34624	8576
21	Portugal	70599	60719	49654	14586	9487	18936	18255	5242	5756	228	7515	7446
22	Romania	121355	95545	86038	33353	18577	15133	14129	10192	12732	125	16777	5683
23	Slovenia	19522	11031	15980	6219	1900	5272	5228	2228	971	30	1963	577
24	Slovakia	45982	39893	32008	9507	9316	6654	6547	4545	8522	164	3065	2222
25	Finland	74553	52474	60649	30547	9905	13570	12001	5083	5766	73	5882	2186
26	Sweden	66232	32177	49359	13091	10122	20744	19122	4239	6841	311	7873	1848
27	United Kingdom	590247	586493	501113	192184	67546	118455	111735	108315	26731	1	45908	16495
28	Iceland	4542	5276	1866	5	213	900	844	556	1810	6	646	214
29	Norway	53896	20951	40726	14899	3612	15142	10097	3592	7479	170	4273	1248
30	Switzerland	54247	53367	44017	4190	5985	16422	16092	17050	3689	215	5688	612
	Range	933509	958897	780447	354501	114961	154153	146315	145788	72469	2937	93798	19677
	Mean	161119	149628.2	128321.6	48148.1	19823.13	32106.3	30121.7	25252.6	11870.3	400.93	15739.17	4786.03
	Standard Deviation	220938	219088.3	181771.5	75728.93	27823.22	43507.51	41002.2	23603.62	15541.64	678.39	21641.52	5786.77

Appendix 3. High and Low Curenty Data Extract

	Country	A	B	C	D	E	F	G	H	I	J	K	L
1	Belgium	132459	131417	107608	26434	23708	24257	23459	32602	13460	214	10042	1135
2	Czech Republic	139158	133639	115205	56251	23807	17448	16979	12340	12061	503	7777	3612
3	Germany	936544	953827	782313	354506	115007	154730	146844	145928	72569	1944	67479	12239
4	Greece	118287	115645	93213	52222	6764	22996	19184	9839	10542	316	9282	4934
5	Spain	355898	326944	269835	72418	63434	91423	83872	39252	28017	2938	40014	15094
6	France	522373	490148	370022	61564	68440	132154	125273	103199	37509	1224	93876	19741
7	Italy	501318	444787	415727	133255	61375	118849	110018	94153	31963	1658	33741	18229
8	Hungary	67679	64307	49070	16668	3903	11870	11598	14270	6386	269	8267	3687
9	Netherlands	210053	213054	177819	66613	27326	34988	34242	45820	10432	170	16624	5008
10	Austria	84594	80983	64328	14293	15618	22452	21890	11401	10680	327	7453	1806
11	Poland	400865	357985	327117	173536	30969	48766	47151	61954	29769	779	34624	8576
12	Portugal	70599	60719	49654	14586	9487	18936	18255	5242	5756	228	7515	7446
13	Romania	121355	95545	86038	33353	18577	15133	14129	10192	12732	125	16777	5683
14	Finland	74556	52474	60649	30547	9905	13570	12001	5083	5766	73	5882	2186
15	United Kingdom	590247	586493	501113	192184	67546	118455	111735	108315	26731	0	45908	16495
High 15	Range	868865	901353	733243	340213	111104	142860	135246	140845	66813	2938	87994	18606
	Mean	288399	273864.5	231314.1	86562	36391.07	56401.8	53108.67	46639.33	20958.2	717.87	27017.4	8391.4
	Standard Deviation	256553.6	256629	213030.7	92924.08	31703.96	51160.87	48155	45665.55	17771.85	854.18	26026.4	6336.72
1	Bulgaria	61427	52796	46438	31464	3798	7949	7508	2013	3852	46	6406	4686
2	Denmark	61065	58895	48789	23915	4453	13248	12241	6600	1685	76	9520	995
3	Estonia	20517	16759	18185	14644	510	2260	2078	646	498	18	1344	472
4	Ireland	61314	60284	40510	13328	4549	11606	11061	10996	1934	72	17910	889
5	Cyprus	10838	10673	7505	3880	716	2313	2313	580	807	0	670	1857
6	Latvia	12077	−5070	8401	2261	1087	3222	2963	1729	639	42	2330	666
7	Lithuania	20810	9095	12848	5446	1120	4565	4125	1431	2249	93	4458	1161
8	Luxembourg	12075	11780	10647	1271	1401	6288	6273	1643	660	14	690	64
9	Malta	3035	2973	2657	1893	46	577	529	140	100	1	78	199
10	Slovenia	19522	11031	15980	6219	1900	5272	5228	2228	971	30	1963	577
11	Slovakia	45982	39893	32008	9507	9316	6654	6547	4545	8522	164	3065	2222
12	Swedeen	66232	32177	49359	13091	10122	20744	19122	4239	6841	311	7873	1848
13	Iceland	4542	5276	1866	5	213	900	844	556	1810	6	646	214
14	Norway	53896	20951	40726	14899	3612	15142	10097	3592	7479	170	4273	1248
15	Switzerland	54247	53367	44017	4190	5985	16422	16092	17050	3689	215	5688	612
Low 15	Range	63197	65354	47493	31459	10076	20167	18593	16910	8422	311	17832	4622
	Mean	33838.6	25392	25329.07	9734.2	3255.2	7810.8	7134.73	3865.87	2782.4	83.87	4460.93	1180.67
	Standard Deviation	24066.08	22253.43	18152, 81	8974.83	3195.32	6221.92	5590.46	4631.91	2742.38	91.88	4692.22	1164.93

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 689 – 705

Efficient exponential ratio estimator for estimating the population mean in simple random sampling

Ekpenyong, Emmanuel John*† and Enang, Ekaette Inyang‡

Abstract

This paper proposes, with justification, two exponential ratio estimators of population mean in simple random sampling without replacement. Their biases and mean squared error are derived and compared with existing related ratio estimators. Analytical and numerical results show that at optimal conditions, the proposed ratio estimators are always more efficient than the regression estimator and some existing estimators under review.

2000 AMS Classification: 62D05

Keywords: Exponential ratio estimators, mean squared error, optimal conditions, efficiency, simple random sampling.

Received 06/03/2014 : Accepted 31/05/2014 Doi : 10.15672/HJMS.2014197470

1. Introduction

In Sample Surveys, auxiliary information are always used to improve the precision of estimates of population parameters. This can be done at either estimation or selection stage or both stages. The commonly used estimators, which make use of auxiliary variables, include ratio estimator, regression estimator, product estimator and difference estimator. The classical ratio estimator is preferred when there is a high positive correlation between the variable of interest, Y and the auxiliary variable, X with the regression line passing through the origin. The classical product estimator, on the other hand is mostly preferred when there is a high negative correlation between Y and X while the linear regression estimator is most preferred when there is a high positive correlation between the two variables and the regression line of the study variable on the auxiliary

*Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria,
E-mail: ekpesstat@yahoo.com
†Corresponding Author.
‡Department of Mathematics/Statistics&Computer Science, University of Calabar, Calabar, Cross River State, Nigeria,
E-mail: ekkaass@yahoo.com

variable has intercept on Y axis. Ratio estimation has gained relevance in Estimation theory because of its improved precision in estimating the population parameters. It has been widely applied in Agriculture to estimate the mean yield of crops in a certain area and in Forestry, to estimate with high precision, the mean number of trees or crops in a forest or plantation. Other areas of relevance include Economics and Population studies to estimate the ratio of income to family size.

According to [13], regression estimator, in spite of its lesser practicability, seems to be holding a unique position due to its sound theoretical basis. The classical ratio and product estimators even though considered to be more useful in many practical situation have efficiencies which does not exceed that of the linear regression. As a result of this limitation, most authors have carried out several researches towards the modification of the existing ratio, product or classes of ratio and product estimators of the population mean in simple random sampling without replacement to improve efficiency. Among authors, who have carried out researches in this direction are [9], [10], [11], [25], [14], [15] , [5], [2], [3], [1], [20], [21], [22], [23], [15], [16], [4], [19] and [28].

So far, only the estimators proposed by [17], which is a modification of those of [9] and [10] is more efficient than the linear regression estimator.. This paper therefore proposes ratio estimators using an exponential ratio estimator, whose efficiencies would be better than regression estimator, [5] and compared with other ratio estimators including [17]. Authors like [6], [7],[13] and [18] extended related works of ratio estimators to stratified sampling.

This work reviews some related existing estimators, proposes new improved estimators and derive their properties. Their efficiencies are used to compare with other existing estimators and empirical results used to validate every theoretical claim.

2. Review of some related existing Estimators

Consider a finite population $\Pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$ of size N. Let Y and X be the study and auxiliary variables with population means $\bar{Y}$ and $\bar{X}$ respectively. It is assumed that information on the population mean $\bar{X}$ of the auxiliary variable is known and $Y_i, X_i \geq 0$ (since the survey variables are generally non-negative). Let a sample of size n be drawn by simple random sampling without replacement (SRSWOR) from the population Π and the sample means $\bar{y}$ and $\bar{x}$ of the study and auxiliary variables obtained respectively. Given the above population, a summary of some related existing estimators with their Mean Squared Errors (MSE's) are given below:

Table 1: Existing related estimators with their MSEs

S/N	Estimators	MSE
1	$\bar{y}$, unbiased sample mean	$\bar{Y}^2\lambda C_y^2$
2	$\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X}$, Classical Ratio	$\bar{Y}^2\lambda[C_y^2 - 2\rho C_y C_x + C_x^2]$
3	$\bar{y}_{\Re} = \bar{y}\exp\left[\frac{(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}\right]$ Bahl and Tuteja [1]	$\bar{Y}^2\lambda[C_y^2 + \frac{C_x^2}{4}(1-4k)]$
4	$\bar{y}_{\text{GS}} = \left[\omega_1^*\bar{y} + \omega_2^*\left(\bar{X}-\bar{x}\right)\right]\left(\frac{\eta\bar{X}+\delta}{\eta\bar{x}+\delta}\right)$ Gupta and Shabbir [5]	$\bar{Y}^2[1-\nu_1]$
5	$\bar{y}_{\text{GS}} = \psi_1^*\bar{y}\left(\frac{\eta\bar{X}+\delta}{\eta\bar{x}+\delta}\right) + \psi_2^*\left(\bar{X}-\bar{x}\right)\left(\frac{\eta\bar{X}+\delta}{\eta\bar{x}+\delta}\right)^2$ Singh and Solanki [17]	$\bar{Y}^2[1-\nu_2]$
6	$\bar{y}_{\text{reg}} = \bar{y} + b\left(\bar{X}-\bar{x}\right)$, Regression Estimator	$\bar{Y}^2\lambda C_y^2(1-\rho^2)$
7	$t_{(\alpha,\zeta)} = \bar{y}\{2 - \left(\frac{\bar{x}}{\bar{X}}\right)^\alpha \exp\left[\frac{\zeta\left(\bar{x}-\bar{X}\right)}{\left(\bar{X}+\bar{x}\right)}\right]\}$ Solanki et al [25]	$\bar{Y}^2\lambda\{C_y^2 + \frac{(2\alpha+\zeta)}{4}C_x^2[(2\alpha+\zeta)+4k]\}$

where
$C_x = \frac{S_x}{\bar{X}}$ be the coefficient of variation of the auxiliary variable,
$C_y = \frac{S_y}{\bar{Y}}$ be the coefficient of variation of the study variable,
$\rho = \frac{S_{xy}}{S_x S_y}$ be the correlation coefficient between the auxiliary and study variables
$k = \frac{\rho C_y}{C_x}$ and $f = \frac{n}{N}$, the sampling fraction; where

$$S_x{}^2 = (N-1)^{-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2,$$

population variance of the auxiliary variable;

$$S_y{}^2 = (N-1)^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2,$$

population variance of the study variable;

$$S_{xy} = (N-1)^{-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)\left(y_i - \bar{Y}\right),$$

population covariance between the auxiliary and study variables;

$$\bar{X} = N^{-1}\sum_{i=1}^{N}x_i,\ \text{population mean of the auxiliary variable}$$

$$\bar{Y} = N^{-1}\sum_{i=1}^{N}y_i,\ \text{population mean of the study variable}$$

$$\bar{x} = n^{-1}\sum_{i=1}^{n}x_i,\ \text{sample mean of the auxiliary variable,}$$

$$\bar{y} = n^{-1}\sum_{i=1}^{n}y_i,\ \text{sample mean of the study variable,}$$

$$\alpha_1 = \left\{1 + \lambda\left[C_y^2 + \tau C_x^2\left(3\tau - 4k\right)\right]\right\}, \alpha_2 = \lambda C_x^2, \alpha_3 = \lambda C_x^2\left(k - 2\tau\right),$$
$$\alpha_4 = \left[1 - \lambda\tau C_x^2\left(k - \tau\right)\right], \alpha_5 = \lambda\tau C_x^2$$
$$\tau = \frac{\eta\bar{X}}{\left(\eta\bar{X} + \delta\right)}.$$
$$A = \left\{1 + \lambda\left[C_x^2 + \tau C_x^2\left(3\tau - 4k\right)\right]\right\}, B = \lambda C_x^2, C = \lambda C_x^2\left(3\tau - k\right),$$
$$D = \left[1 + \lambda\tau C_x^2\left(\tau - k\right)\right], E = 2\lambda\tau C_x^2$$
$$\omega^* = \frac{\left(\alpha_2\alpha_4 + \alpha_3\alpha_5\right)}{\left(\alpha_1\alpha_2 - \alpha_3^2\right)}, \omega^* = \frac{R\left(\alpha_1\alpha_5 + \alpha_3\alpha_4\right)}{\left(\alpha_1\alpha_2 - \alpha_3^2\right)},$$
$$R = \frac{\bar{Y}}{\bar{X}}, \nu_1 = \frac{\left(\alpha_2\alpha_4^2 + 2\alpha_3\alpha_4\alpha_5 + \alpha_1\alpha_5^2\right)}{\left(\alpha_1\alpha_2 - \alpha_3^2\right)}$$
$$\psi_1^* = \frac{(\text{BD} - \text{CE})}{(\text{AB} - C^2)}, \psi_2 = \frac{(\text{AE} - \text{CD})}{(\text{AB} - C^2)}, \nu_2 = \frac{\left(BD^2 - 2\text{CDE} + AE^2\right)}{(\text{AB} - C^2)}$$

ω_1^*, ω_2^*, ψ_1^* and ψ_2^* are optimum values of ω_1, ω_2, ψ_1 and ψ_2 respectively, η $(\eta \neq 0)$, α, δ and ζ are suitably chosen constants or functions of the known parameters such as standard deviation S_x, moment ratios $\beta_1(x)$, $\beta_2(x)$, Coefficient of Variation, C_x, and Correlation Coefficient $\rho_{Y,X}$ between the variables Y and X, and so on.

[17] made corrections on the Mean Squared Error(MSE) of the class of estimators proposed by [5] to obtain the correct expression of the MSE. The corrected version would be used in this study. They went further to compare the efficiency of the estimators of [5] with those proposed by [9], [10], [11] and found that a class of estimators proposed by [5] was more efficient than those of [9], [10], [11]. [17] proceeded to propose a new class of modified estimators from that of [5]. These estimators were more efficient than those of [5], [9], [10], [15] and the regression estimator. In this paper, two alternative ratio estimators which are more efficient than the linear regression estimators are proposed with justification.

3. Proposed Estimator I

The first ratio estimator is proposed as

(3.1) $$\bar{y}_{\text{pr1}} = \theta_1\bar{y} + \theta_2\left(\bar{X} - \bar{x}\right)\exp\left[\left(\bar{X} - \bar{x}\right)/\left(\bar{X} + \bar{x}\right)\right]$$

θ_1 and θ_2 are suitably chosen scalars, such that $\theta_1 > 0$ and $-\infty < \theta_2 < \infty$.

3.1. The bias and Mean Squared Error of the proposed estimator. The proposed estimator in terms of e's, is expressed as

(3.2) $$\bar{x} = \bar{X}\left(1 + e_x\right)\bar{y} = \bar{Y}\left(1 + e_y\right)$$

where $e_x = \bar{x} - \bar{X}/\bar{X} \quad e_y = \bar{y} - \bar{Y})/\bar{Y}$.

(3.3) $$E[e_x] = E[e_y] = 0, E[e_x]^2 = \frac{1-f}{n}{C_x}^2; E[e_y]^2 = \frac{(1-f}{n}{C_y}^2;$$
$$E[e_xe_y] = \frac{(1-f)}{n}\rho\ C_xC_y = \frac{(1-f)}{n}k{C_x}^2.$$

(3.4) $$\bar{y}_{\text{pr1}} = \bar{Y}\left[\theta_1 + \theta_1e_y - \theta_2\frac{\bar{X}}{\bar{Y}}\left[1 - \frac{e_x}{2}\left(1 + \frac{e_x}{2}\right)^{-1} + \frac{e_x^2}{2}\left(1 + \frac{e_x}{2}\right)^{-2} + \ldots\right]\right].$$

It is assumed that $|e_x| < 1$; $|e_y| < 1$ so that $\left(1 + \frac{e_x}{2}\right)^{-1}$ and $\left(1 + \frac{e_x}{2}\right)^{-2}$ can be expanded.

Expanding equation (3.4) by Taylor series approximation and neglecting terms of e's having powers greater than two, we have:

$\bar{y}_{\text{pr1}} = \bar{Y}\left[\theta_1 + \theta_1 e_y - \theta_2 m e_x [1 - (e_x/2)\left(1 - e_x/2 + e_x^2/4\right) + e_x^2/8]\right]$

where $m = \bar{X}/\bar{Y}$, leading to

$$\bar{y}_{\text{pr1}} - \bar{Y} = \bar{Y}\left\{(\theta_1 - 1) + \theta_1 e_y - \theta_2 m e_x + \theta_2 m \frac{e_x^2}{2}\right\}. \tag{3.5}$$

Therefore, the Bias of the estimator is given as

$$B(\bar{y})_{pr1} = E[\bar{y}_{\text{pr1}} - \bar{Y}] = \bar{Y}[(\theta_1 - 1) + \theta_2 \text{m}\lambda \frac{{C_x}^2}{2}] \tag{3.6}$$

The MSE of $\bar{y}_{\text{pr1}}$ to first degree approximation is obtained by squaring equation (3.5) and ignoring powers of 'e' greater than two and taking the expectation of the square as follows:

$\left(\bar{y}_{\text{pr1}} - \bar{Y}\right)^2 = \bar{Y}^2\left[(\theta_1 - 1)^2 + \theta_2\left(\theta_2 - 1\right) m e_x^2 + \theta_1^2 e_y^2 - 2\theta_1\theta_2 m e_y e_x + \theta_2^2 m^2 e_x^2\right]$

$= \bar{Y}^2\left[\theta_1^2 - 2\theta_1 + 1 + \theta_1\theta_2 m e_x^2 - \theta_2 m e_x^2 + \theta_1^2 e_y^2 - 2\theta_1\theta_2 m e_y e_x + \theta_2^2 m^2 e_x^2\right]$

$= \bar{Y}^2\left[1 + \theta_1^2\left(1 + e_y^2\right) - 2\theta_1 - 2\theta_1\theta_2 m\left(e_y e_x - \frac{e_x^2}{2}\right) - 2\theta_2 m\frac{e_x^2}{2} + \theta_2^2 m^2 e_x^2\right].$

$$\begin{aligned}\text{MSE}\left(\bar{y}_{\text{pr1}}\right) &= E\left(\bar{y}_{\text{pr1}} - \bar{Y}\right)^2 = \bar{Y}^2[1 + \theta_1^2\left(1 + \lambda C_y^2\right) - 2\theta_1 - \\ &- 2\theta_1\theta_2 \text{m}\lambda C_x^2\left(k - \frac{1}{2}\right) - 2\theta_2 \text{m}\lambda\frac{C_x^2}{2} + \theta_2^2 m^2 \lambda C_x^2] \\ &= \bar{Y}^2[1 + \theta_1^2\gamma_1 - 2\theta_1 - 2\theta_1\theta_2 m\gamma_2 - 2\theta_2 m\gamma_3 + \theta_2^2 m^2\gamma_4]\end{aligned} \tag{3.7}$$

where $\gamma_1 = 1 + \lambda C_y^2, \gamma_2 = C_x^2\lambda\left(k - \frac{1}{2}\right), \gamma_3 = \frac{\lambda C_x^2}{2}, \gamma_4 = \lambda C_x^2$

3.2. Optimal conditions for MSE of proposed estimator I. To obtain the optimum values of θ_1 and θ_2 that would minimize the MSE of the estimator, the partial derivative of (3.7) is taken with respect to θ_1 and θ_2 respectively and equated to zero as shown below:

$$\frac{\partial \text{MSE}\left(\bar{y}_{\text{pr1}}\right)}{\partial\theta_1} = 2\theta_1\gamma_1 - 2\theta_2 m\gamma_2 - 2 = 0 \Rightarrow \theta_1\gamma_1 - \theta_2 \text{m}\gamma_2 = 1 \tag{3.8}$$

$$\frac{\partial \text{MSE}\left(\bar{y}_{\text{pr1}}\right)}{\partial\theta_2} = -2\theta_1 m\gamma_2 - 2m\gamma_3 + 2\theta_2 m^2\gamma_4 = 0 \Rightarrow -\theta_1 m\gamma_2 + \theta_2 m^2\gamma_4 = m\gamma_3 \tag{3.9}$$

Solving equations (3.8) and (3.9) simultaneously gives the optimal values of θ_1 and θ_2 as

$$\theta_1^* = (\gamma_4 + \gamma_2\gamma_3)/(\gamma_1\gamma_4 - \gamma_2^2) \tag{3.10}$$

$$\theta_2^* = R(\gamma_2 + \gamma_1\gamma_3)/(\gamma_1\gamma_4 - \gamma_2^2) \tag{3.11}$$

where $R = \bar{Y}/\bar{X}$

Substituting equations (3.10) and (3.11) in (3.7) gives the minimum MSE as:

$$\text{MSE}_{\min}\left(\bar{y}_{\text{pr1}}\right) = \bar{Y}^2\left\{1 - [(\gamma_4 + 2\gamma_2\gamma_3 + \gamma_1\gamma_3^2)/(\gamma_1\gamma_4 - \gamma_2^2)]\right\} \tag{3.12}$$

which leads to

$$\text{MSE}_{\min}\left(\bar{y}_{\text{pr1}}\right) = \bar{Y}^2[1 - q_1] \tag{3.13}$$

where $q_1 = (\gamma_4 + 2\gamma_2\gamma_3 + \gamma_1\gamma_3^2)/(\gamma_1\gamma_4 - \gamma_2^2)$. These results can be summarized in theorem I below:

3.1. Theorem. *If $\theta_1 \to \theta_1^*$ and $\theta_2 \to \theta_2^*$ such that $\theta_1^* > 0$ and $-\infty < \theta_2^* < \infty$ the proposed estimator will have a Mean Squared Error,*
$MSE(\bar{y}_{pr1}) \geq \bar{Y}^2 \left\{1 - [(\gamma_4 + 2\gamma_2\gamma_3 + \gamma_1\gamma_3^2)/(\gamma_1\gamma_4 - \gamma_2^2)]\right\}$,
with strict equality holding if $\theta_1 = \theta_1^*$ and$\theta_2 = \theta_2^*$.

3.3. Some special cases of proposed estimator I.

Case I: When $\theta_1 = 1$. The proposed estimator becomes

$$\bar{y}_{\text{pr11}} = \bar{y} + \theta_2 \left(\bar{X} - \bar{x}\right) \exp\left[(\bar{X} - \bar{x})/\left(\bar{X} + \bar{x}\right)\right], \tag{3.14}$$

which is obtained by setting $\theta_1 = 1$ in (3.7). The optimum value of θ_2 that would make the MSE a minimum is:

$$\theta_2^{'} = \frac{\gamma_2 + \gamma_3}{m\gamma_4} = B \tag{3.15}$$

where B is the regression coefficient. Substitution of equation (3.15) into (3.7) with $\theta_1 = 1$, gives the minimum MSE as

$$\text{MSE}_{\min}\left(\bar{y}_{\text{pr11}}\right) = \lambda \bar{Y}^2 C_y^2 (1 - \rho^2) \tag{3.16}$$

Remark I: It should be noted here that equation (3.16) gives the same expression as the Variance of the linear regression estimator

$$\bar{y}_{\text{reg}} = \bar{y} + b(\bar{X} - \bar{x}) \tag{3.17}$$

where b is the sample regression coefficient. Therefore, when $\theta_1 = 1$ and θ_2 is optimal, the proposed estimator I has the same efficiency as the simple linear regression estimator.

Case II: When $\theta_1 = 1$ and $\theta_2 = 1$. The proposed estimator reduces to

$$\bar{y}_{\text{pr12}} = \bar{y} + \left(\bar{X} - \bar{x}\right) \exp[(\bar{X} - \bar{x})/(\bar{X} + \bar{x})] \tag{3.18}$$

with MSE given as

$$\text{MSE}\left(\bar{y}_{\text{pr12}}\right) = \lambda \bar{Y}^2 [C_y^2 - mC_x^2 \left(2k - m\right)] \tag{3.19}$$

Case III: When $\theta_1 = 1, \theta_2 = 0$ The proposed estimator reduces to unbiased sample mean estimator $\bar{y}$, with Variance given as:

$$V\left(\bar{y}_{\text{pr13}}\right) = \lambda \bar{Y}^2 C_y^2 \tag{3.20}$$

These cases are specific members of the family of the proposed estimator I obtained by varying the values of θ_1 and θ_2. Table 2 gives a summary of some members of this proposed family of estimators.

Table 2: Some members of the family of proposed estimator I and their MSE's.

S/N	Estimator	θ_1	θ_2	MSE
1	$\bar{y} + \theta_2(\bar{X} - \bar{x}) \exp\left[\frac{(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}\right]$	1	$\frac{\gamma_2+\gamma_3}{m\gamma_4} = b$	$\lambda \bar{Y}^2 C_y^2 (1 - \rho^2)$
2	$\bar{y} + \left(\bar{X} - \bar{x}\right) \exp[\frac{(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}]$	1	1	$\bar{Y}^2 [C_y^2 - mC_x^2 \left(2k - m\right)]$
3	$\bar{y}$	1	0	$\lambda \bar{Y}^2 C_y^2$
4	$\theta_1^* \bar{y} + \theta_2^* (\bar{X} - \bar{x}) \exp\left[\frac{(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}\right]$	θ_1^*	θ_2^*	$\bar{Y}^2 \left\{1 - [\frac{\gamma_4+\gamma_1\gamma_3^2+2\gamma_2\gamma_3}{\gamma_1\gamma_4-\gamma_2^2}]\right\}$

4. Proposed estimator II

The second proposed estimator takes the form

(4.1) $$\bar{y}_{\text{pr2}} = \varphi_1 \bar{y} + \varphi_2 \left(\bar{X} - \bar{x}\right) \exp[2(\bar{X} - \bar{x})/(\bar{X} + \bar{x})]$$

Where φ_1 and φ_2 are suitable scalars and $\varphi_1 > 0, -\infty < \varphi_2 < \infty$. Expressing (4.1) in terms of e's gives

(4.2) $$\bar{y}_{\text{pr2}} = \bar{Y}\left\{\varphi_1 + \varphi_1 e_y - \varphi_2 m[1 - e_1\Big(1 + \frac{e_x}{2}\Big)^{-1} + \frac{e_x^2}{2}\Big(1 + \frac{e_x}{2}\Big)^{-2} + \ldots]\right\}.$$

The first degree approximation of equation (4.2) is obtained as: $\bar{y}_{\text{pr2}} = \bar{Y}[\varphi_1 + \varphi_1 e_y - \varphi_2 m e_x \left(1 - e_x + e_x^2\right)]$

$$= \bar{Y}[\varphi_1 + \varphi_1 e_y - \varphi_2 m e_x + \varphi_2 m e_x^2]$$

(4.3) $$\bar{y}_{\text{pr2}} - \bar{Y} = \bar{Y}[(\varphi_1 - 1) + \varphi_1 e_y - \varphi_2 m e_x + \varphi_2 m e_x^2]$$

The Bias of $\bar{y}_{\text{pr2}}$ is obtained from equation (4.3) as:

(4.4) $$B\left(\bar{y}_{\text{pr2}}\right) = E\left(\bar{y}_{\text{pr2}} - \bar{Y}\right) = \bar{Y}\left[(\varphi_1 - 1) + \varphi_2 \mathrm{m} \lambda C_x^2\right].$$

Squaring equation (4.3) and ignoring powers of 'e' greater than two, we have:

(4.5) $$\begin{aligned}\left(\bar{y}_{\text{pr2}} - \bar{Y}\right)^2 &= \bar{Y}^2[(\varphi_1 - 1)^2 + 2\varphi_2\left(\varphi_1 - 1\right) m e_x^2 \\ &+ \varphi_2 e_y^2 - 2\varphi_1\varphi_2 m e_y e_x + \varphi_2^2 m^2 e_x^2] \\ &= \bar{Y}^2[1 + \varphi_1^2\left(1 + e_y^2\right) - 2\varphi_1 - 2\varphi_1\varphi_2 m\left(e_y e_x - e_x^2\right) - 2\varphi_2 m e_x^2 + \varphi_2^2 m^2 e_x^2]\end{aligned}$$

Its MSE is obtained by taking the expectation of equation (4.5) as shown below:

(4.6) $$\begin{aligned}\mathrm{MSE}\left(\bar{y}_{\text{pr2}}\right) &= E(\bar{y}_{\text{pr2}} - \bar{Y})^2 = \bar{Y}^2\left[1 + \varphi_1^2\left(1 + \lambda C_y^2\right)\right. \\ &-2\varphi_1 - 2\varphi_1\varphi_2 \mathrm{m}\lambda C_x^2\left(k - 1\right) - \varphi_2 \mathrm{m}\lambda C_x^2 + \varphi_2^2 m^2 \lambda C_x^2\big] \\ &= \bar{Y}^2[\varphi_1^2\gamma_1 - 2\varphi_1 - 2\varphi_1\varphi_2 m\gamma_5 - 2\varphi_2 m\gamma_4 + \varphi_2^2 m^2\gamma_4]\end{aligned}$$

where $\gamma_5 = \lambda C_x^2(k-1)$

4.1. Optimality conditions for estimator II. To investigate the optimal conditions for estimator II, let

$$\frac{\partial \mathrm{MSE}(\bar{y}_{\text{pr2}})}{\partial \varphi_1} = \frac{\partial \mathrm{MSE}(\bar{y}_{\text{pr2}})}{\partial \varphi_2} = 0$$

so that,

(4.7) $$\varphi_1\gamma_1 - \varphi_2 m\gamma_5 = 1$$

(4.8) $$-\varphi_1 m\gamma_5 + \varphi_2 m^2\gamma_4 = m\gamma_4.$$

Solving equations (4.7) and (4.8) simultaneously give the optimal values of φ_1 and φ_2 as:

(4.9) $$\varphi_1^\star = (\gamma_4 + \gamma_4\gamma_5)/(\gamma_1\gamma_4 - \gamma_5^2)$$

(4.10) $$\varphi_2^\star = R(\gamma_5 + \gamma_1\gamma_4)/\left(\gamma_1\gamma_4 - \gamma_5^2\right).$$

Substituting equations (4.9) and (4.10) in (4.6) yields the minimum MSE of the estimator as:

(4.11) $$\mathrm{MSE}\left(\bar{y}_{\text{pr2}}\right) = \bar{Y}^2\{1 - \left[(\gamma_4 + 2\gamma_4\gamma_5 + \gamma_1\gamma_4^2)/(\gamma_1\gamma_4 - \gamma_5^2)\right]\} = \bar{Y}^2[1 - q_2]$$

where,

$$q_2 = (\gamma_4 + 2\gamma_4\gamma_5 + \gamma_1\gamma_4^2)/(\gamma_1\gamma_4 - \gamma_5^2)$$

These results are summarized in the following theorem.

4.1. Theorem. *If $\varphi_1 \to \varphi_1^\star$ and $\varphi_2 \to \varphi_2^\star$ such that $\varphi_1^\star > 0$ and $-\infty < \varphi_2^\star < \infty$, the proposed estimator will have a Mean Squared Error of*
$MSE(\bar{y}_{pr2}) \geq \bar{Y}^2 \left\{1 - [(\gamma_4 + 2\gamma_4\gamma_5 + \gamma_1\gamma_4^2)/(\gamma_1\gamma_4 - \gamma_5^2)]\right\}$,
with strict equality holding if $\varphi_1 = \varphi_1^\star$ and $\varphi_2 = \varphi_2^\star$.

4.2. Some special cases of proposed estimator II. Some special cases of $\bar{y}_{\mathrm{pr2}}$ with varying values of φ_1 and φ_2 and MSEs are given in Table 3.

Table 3: Some types of estimator II and their MSEs

S/N	Estimator	φ_1	φ_2	MSE
1	$\bar{y}+\varphi_2(\bar{X}-\bar{x})\exp\left[\frac{2(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}\right]$	1	$\frac{\gamma_4+\gamma_5}{m\gamma_4}=b$	$\lambda\bar{Y}^2C_y^2(1-\rho^2)$
2	$\bar{y}+\left(\bar{X}-\bar{x}\right)\exp[\frac{2(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}]$	1	1	$\bar{Y}^2[C_y^2-mC_x^2\left(2k-m\right)]$
3	$\bar{y}$	1	0	$\lambda\bar{Y}^2C_y^2$
4	$\varphi_1^\star\bar{y}+\varphi_2^\star(\bar{X}-\bar{x})\exp\left[\frac{2(\bar{X}-\bar{x})}{(\bar{X}+\bar{x})}\right]$	$\varphi_1^\star$	$\varphi_2^\star$	$\bar{Y}^2\left\{1-[\frac{\gamma_4+\gamma_1\gamma_4^2+2\gamma_4\gamma_5}{\gamma_1\gamma_4-\gamma_5^2}]\right\}$

5. Efficiency Comparison

In this section, the MSE of some existing ratio estimators are compared with the optimal MSE of the proposed estimators.

5.1. Unbiased simple random sample mean, $\bar{y}$. The Variance of the simple random mean expressed in terms of $\gamma' s$ is:

(5.1) $$V\left(\bar{y}\right)=\bar{Y}^2(\gamma_1-1)$$

Therefore, for the proposed estimator I to be more efficient than the simple sample random mean, $\bar{y}$, $V\left(\bar{y}\right)-\text{MSE}\left(\bar{y}_{\text{pr1}}\right)>0$

$$\Rightarrow\bar{Y}^2\left[\gamma_1+q_1-2\right]>0$$

(5.2) $$\Rightarrow\left[\gamma_1+q_1-2\right]>0.$$

Also for $\bar{y}_{\text{pr1}}$ to be more efficient than $\bar{y}$

$$V\left(\bar{y}\right)-\text{MSE}\left(\bar{y}_{\text{pr2}}\right)>0$$

$$\Rightarrow\bar{Y}^2\left[\gamma_1+q_2-2\right]>0$$

(5.3) $$\Rightarrow\left[\gamma_1+q_2-2\right]>0.$$

If equations (5.2) and (5.3) hold, then the proposed estimators would be more efficient than the simple random sample mean.

5.2. Classical ratio estimator, $\bar{y}_R$. The MSE of $\bar{y}_R$ expressed in terms of $\gamma' s$ is given by: For estimator I,

(5.4) $$\text{MSE}\left(\bar{y}_R\right)=\bar{Y}^2[\gamma_1-2\gamma_2-1]$$

And for estimator II

(5.5) $$\text{MSE}\left(\bar{y}_R\right)=\bar{Y}^2[\gamma_1-2\gamma_4-\gamma_5-1]$$

Therefore, for the proposed estimator $\bar{y}_{\text{pr1}}$ to be more efficient than the classical ratio estimator,

$$\text{MSE}\left(\bar{y}_R\right)-\text{MSE}\left(\bar{y}_{\text{pr1}}\right)>0$$

$$\Rightarrow\bar{Y}^2\left[\gamma_1-2\gamma_2-2+q_1\right]>0$$

(5.6) $$\Rightarrow\left[\left(\gamma_1+q_1\right)-2\left(\gamma_2+1\right)\right]>0$$

Similarly, for $\bar{y}_{\text{pr2}}$ to be more efficient than $\bar{y}_R$

$$\text{MSE}\left(\bar{y}_R\right)-\text{MSE}\left(\bar{y}_{\text{pr2}}\right)>0$$

$$\Rightarrow\left[\gamma_1-2\gamma_4-\gamma_5-2+q_2\right]>0$$

(5.7) $$\Rightarrow\left[\left(\gamma_1+q_2\right)-2\left(\gamma_4+1\right)-\gamma_5\right]>0$$

Therefore, for the proposed estimators to be more efficient than the classical ratio estimator, equations (5.6) and (5.7) must hold.

5.3. Regression Estimator, $\bar{y}_{\mathbf{reg}}$. The Variance of the regression estimator expressed in terms of $\gamma's$ is given as: For estimator I

(5.8) $$V\left(\bar{y}_{\text{reg}}\right) = \bar{Y}^2\{\gamma_1 - [(\gamma_2+\gamma_3)^2/\gamma_4] - 1\}$$

and for estimator II

(5.9) $$V\left(\bar{y}_{\text{reg}}\right) = \bar{Y}^2[\gamma_1 - [(\gamma_4+\gamma_5)^2/\gamma_4] - 1]$$

Therefore, for the proposed estimators to be more efficient than the regression estimator,

$$V\left(\bar{y}_{\text{reg}}\right) - \text{MSE}\left(\bar{y}_{\text{pr1}}\right) > 0$$

$$\Rightarrow \bar{Y}^2\left[\gamma_1 - [(\gamma_2+\gamma_3)^2/\gamma_4] - 1 - (1-q_1)\right] > 0$$

$$\Rightarrow \bar{Y}^2\left[\gamma_1 - [(\gamma_2+\gamma_3)^2/\gamma_4] - 2 + q_1\right] > 0$$

(5.10) $$\Rightarrow [\gamma_4(\gamma_1-1) - \gamma_2(\gamma_2+\gamma_3)]^2/\gamma_4\left(\gamma_1\gamma_4 - \gamma_2^2\right) > 0$$

(5.10) holds if $[\ \gamma_4\left(\gamma_1\gamma_4 - \gamma_2^2\right) > 0]$. Therefore,

$$\gamma_4\left(\gamma_1\gamma_4 - \gamma_2^2\right) > 0$$

$$\Rightarrow \gamma_1\gamma_4^2 - \gamma_2^2\gamma_4 > 0$$

$$\Rightarrow \lambda^2 C_x^4\left(1+\lambda C_y^2\right) - \lambda^2 C_x^4\left(k-\tfrac{1}{2}\right)^2 > 0$$

$$\Rightarrow 1+\lambda C_y^2 > \left(k-\frac{1}{2}\right)^2$$

$$\Rightarrow \text{Var}\left(\bar{y}\right) + \bar{Y}^2 > \frac{1}{C_x^2}[\text{MSE}\left(\bar{y}_{\Re}\right) + \bar{Y}^2\lambda C_y^2(\rho^2-1)]$$

(5.11) $$\Rightarrow \text{Var}\left(\bar{y}\right) + \bar{Y}^2 > \frac{1}{C_x^2}[\text{MSE}\left(\bar{y}_{\Re}\right) - \bar{Y}^2\lambda C_y^2(1-\rho^2)]$$

Clearly, from equation (5.11), $\text{MSE}\left(\bar{y}_{\Re}\right)$, the Mean Square Error of [1] is smaller than $\text{Var}\left(\bar{y}\right)$, the Variance of the simple random sample mean. Also, the second term in the bracket on the right hand side of equation (5.11) is the Variance of regression estimator, which is smaller than $V(\bar{y})$. Therefore, the expression on the left hand side of equation (5.11) is always greater than that of the right hand side. Hence, equation (5.11) holds. It follows therefore that $[\gamma_4\left(\gamma_1\gamma_4 - \gamma_2^2\right) > 0\ ]$ and the numerator of (5.10) is a square, which implies that (5.10) holds. Hence, the proposed estimator I is always more efficient than classical regression estimator.
Also,

$$V\left(\bar{y}_{\text{reg}}\right) - \text{MSE}\left(\bar{y}_{\text{pr2}}\right) > 0$$

$$\Rightarrow \bar{Y}^2\left[\gamma_1 - [(\gamma_4+\gamma_5)^2/\gamma_4] - 1 - (1-q_2)\right] > 0$$

$$\Rightarrow \bar{Y}^2\left[\gamma_1 - [(\gamma_4+\gamma_5)^2/\gamma_4] - 2 + q_2\right] > 0$$

(5.12) $$\Rightarrow [\gamma_4(\gamma_1-1) - \gamma_5(\gamma_5+\gamma_4)]^2/\gamma_4\left(\gamma_1\gamma_4 - \gamma_5^2\right) > 0$$

Similarly, for (5.12) to be satisfied,

$$\gamma_4\left(\gamma_1\gamma_4 - \gamma_5^2\right) > 0$$

$$\Rightarrow \gamma_1\gamma_4^2 - \gamma_5^2 > 0$$

$$\Rightarrow 1+\lambda C_y^2 > (k-1)^2$$

$$\Rightarrow \bar{Y}^2 + \text{Var}\left(\bar{y}\right) > \frac{1}{C_x^2}[\text{MSE}\left(\bar{y}_R\right) + \bar{Y}^2\lambda C_y^2(\rho^2-1)]$$

(5.13) $$\Rightarrow \bar{Y}^2 + \text{Var}\left(\bar{y}\right) > \frac{1}{C_x^2}[\text{MSE}\left(\bar{y}_R\right) - \bar{Y}^2\lambda C_y^2(1-\rho^2)]$$

From (5.13), we observe that $\text{MSE}\left(\bar{y}_R\right)$, the Mean Square Error of the classical ratio estimator is always smaller than $\text{Var}\left(\bar{y}\right)$, the variance of simple random sample mean.

In addition, the second term in the bracket of the right hand side of (5.13) is the Variance of the classical regression estimator. Therefore the expression on the left hand side of equation (44) is greater than that of the right hand side. Hence, equation (44) holds and the numerator of (5.12) is positive, which implies that (43) always holds.
Remark II Since equations (5.10) and (5.12) are all greater than zero, then the proposed estimators are always more efficient than the regression estimator. Moreover, since the regression estimator is more efficient than the simple random sample mean, classical ratio estimator, estimators of [14], [9] and [10], and any other ratio estimators, it follows that the proposed estimators are more efficient than these estimators. The above remark is summarized in the following theorem.

5.1. Theorem. *If $\theta_1, \theta_2, \varphi_1\ \varphi_2$ attain or almost attain their optimal values in the proposed estimators, then the proposed estimators are always more efficient than the regression estimator.*

5.4. Gupta and Shabbir [5] estimator, $\bar{\mathbf{y}}_{\mathbf{GS}}$. The proposed estimators would be better than the Gupta and Shabbir's class of estimators if:

$$\mathrm{MSE}\left(\bar{y}_{\mathrm{GS}}\right) - \mathrm{MSE}\left(\bar{y}_{\mathrm{pr1}}\right) > 0$$

$$\Rightarrow \bar{Y}^2\left[(1-\nu_1) - (1-q_1)\right] > 0$$

$$\Rightarrow [q_1 - \nu_1] > 0 \tag{5.14}$$

$\mathrm{MSE}\left(\bar{y}_{\mathrm{GS}}\right) - \mathrm{MSE}\left(\bar{y}_{\mathrm{pr2}}\right) > 0$

$$\Rightarrow [q_2 - \nu_1] > 0 \tag{5.15}$$

5.5. Singh and Solanki [17] estimator, $\bar{\mathbf{y}}_{\mathbf{SS}}$. The proposed estimators would be more efficient than Singh and Solanki's class of estimators if:

$$\mathrm{MSE}\left(\bar{y}_{\mathrm{SS}}\right) - \mathrm{MSE}\left(\bar{y}_{\mathrm{pr1}}\right) > 0$$

$$\Rightarrow \bar{Y}^2\left[(1-\nu_2) - (1-q_1)\right] > 0$$

$$\Rightarrow [q_1 - \nu_2] > 0 \tag{5.16}$$

and

$$\mathrm{MSE}\left(\bar{y}_{\mathrm{SS}}\right) - \mathrm{MSE}\left(\bar{y}_{\mathrm{pr2}}\right) > 0$$

$$\Rightarrow \bar{Y}^2\left[(1-\nu_2) - (1-q_2)\right] > 0$$

$$\Rightarrow [q_2 - \nu_2] > 0 \tag{5.17}$$

6. Empirical Study

To investigate our theoretical results, as well as, test the optimality and efficiency performances of our proposed estimators over other existing ones considered in this study, we make use of data of the following populations.

Population I:

$N = 200, n = 50, \bar{Y} = 500, \bar{X} = 25, C_y = 15, C_x = 2, \rho = 0.90, \beta_2(x) = 50$

[Kadilar and Cingi [11]]

Population II:

$N = 106, n = 20, \bar{Y} = 2212.59, \bar{X} = 27421.70, C_y = 5.22, C_x = 2.10,$
$\rho = 0.86, \beta_2(x) = 34.57$

[Kadilar and Cingi,[9, 10]]

Population III:

$N = 104, n = 20, \bar{Y} = 625.37, \bar{X} = 13.93, C_y = 1.866, C_x = 1.653,$
$\rho = 0.865, \beta_2(x) = 17.516$

[Kadilar and Cingi [11]]

Population IV:

$N = 923, n = 180, \bar{Y} = 436.4345, \bar{X} = 11440.5, C_y = 1.7183, C_x = 1.8645,$
$\rho = 0.9543, \beta_2(x) = 18.7208$

[Koyuncu and Kadilar, [12]]

Table 4: Optimum values (θ_1^*, θ_2^*, MSEs and PREs of some Gupta and Shabbir[5] estimators and the proposed estimators.

Estimators	**Population I**				**Population II**				**Population III**				**Population IV**			
	θ_1^*	θ_2^*	MSE	PRE	θ_1^*	θ_2^*	MSE	PRE	θ_1^*	θ_2^*	MSE	PRE	θ_1^*	θ_2^*	MSE	PRE
$\bar{y}_{\text{GS1}}$	0.60	76.61	95396.05	884.471	0.74	0.09	1043368.08	518.642	0.96	2.41	13321.23	412.828	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS2}}$	0.59	76.48	95304.34	885.322	0.74	0.09	1043366.16	518.643	0.96	0.87	13316.65	412.970	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS3}}$	059	76.49	95308.27	885.285	0.74	0.09	1043366.29	518.643	0.96	0.98	13316.98	412.960	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS4}}$	0.60	76.71	95468.42	883.800	0.74	0.09	1043369.73	518.641	0.96	3.10	13323.21	412.766	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS5}}$	0.60	76.59	95386.75	884.557	0.74	0.09	1043367.79	518.642	0.96	2.19	13320.59	412.848	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS6}}$	0.59	76.48	95303.94	885.325	0.74	0.09	1043366.14	518.643	0.96	0.85	13316.58	412.972	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS7}}$	0.59	76.48	95304.34	885.322	0.74	0.09	1043366.03	518.643	0.96	0.83	13316.52	412.974	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS8}}$	0.60	76.73	95485.97	883.638	0.75	0.09	1049814.73	515.457	0.96	4.25	13326.36	412.669	0.999	-0.001	224.357	1121.014
$\bar{y}_{\text{GS9}}$	0.59	76.48	95303.94	885.325	0.74	0.09	1043366.03	518.643	0.96	0.81	13316.47	412.975	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{GS10}}$	0.59	76.47	95300.40	885.358	0.74	0.09	1043366.03	518.643	0.96	0.70	13316.13	412.986	0.999	-0.005	224.356	1121.019
$\bar{y}_{\text{reg}}$	-	-	160312.5	526.316	-	-	1409113.09	384.025	-	-	13846.05	397.180	-	-	224.625	1119.677
$\bar{y}_{\text{pr1}}$	0.58	83.08	72692.31	1160.7	0.71	0.13	713838.3	758.06	0.92	42.13	11051.73	497.60	0.99	6.23	212.519	1183.458
$\bar{y}_{\text{pr2}}$	0.56	4.23	45870.36	1839.4*	0.65	1.73	240765.7	2247.6*	0.87	0.98	6819.164	806.5*	0.99	0.35	183.147	1373.254*
$\bar{y}$	-	-	843750	100.000	-	-	5411348.28	100.000	-	-	54993.75	100.000	-	-	2515.074	100.000

Table 5: Optimum values (φ_1^*, φ_2^*, MSEs and PREs of some Singh and Solanki [17] estimators and the proposed estimators

Estimator	**Population I**				**Population II**				**Population III**				**Population IV**			
	φ_1^*	φ_2^*	MSE	PRE	φ_1^*	φ_2^*	MSE	PRE	φ_1^*	φ_2^*	MSE	PRE	φ_1^*	φ_2^*	MSE	PRE
$\bar{y}_{SS1}$	0.53	3.98	45246.7	1864.791	0.50	1.57	202185.29	2676.43	0.94	0.13	12986.83	423.458	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS2}$	0.53	3.98	44081.39	1914.073	0.50	1.57	202185.29	2676.82	0.94	0.09	13116.65	419.267	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS3}$	0.53	3.98	44131.01	1911.921	0.50	1.57	202155.53	2676.80	0.94	0.10	13107.25	419.567	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS4}$	0.53	3.98	46177.73	1827.179	0.50	1.57	202157.65	2676.09	0.94	0.15	12931.31	412.766	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS5}$	0.53	3.98	45127.48	1869.703	0.50	1.57	202210.82	2676.49	0.94	0.12	13005.06	425.276	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS6}$	0.53	3.98	44076.42	1914.289	0.50	1.57	202180.85	2676.83	0.94	0.09	13118.60	422.864	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS7}$	0.53	3.98	44081.39	1914.073	0.50	1.57	202155.27	2676.82	0.94	0.09	13116.65	419.204	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS8}$	0.53	3.98	46405.23	1818.222	0.54	1.52	202155.53	1817.35	0.94	0.17	12844.01	419.267	0.99	-0.03	224.510	1120.250
$\bar{y}_{SS9}$	0.53	3.98	44076.42	1914.289	0.50	1.57	297759.98	2676.85*	0.94	0.09	13121.61	428.167	1.00	-0.12	224.527	1120.166
$\bar{y}_{SS10}$	0.53	3.98	44031.68	1916.234*	0.50	1.57	202153.61	2676.85*	0.94	0.09	13131.21	419.108	1.00	-0.12	224.527	1120.166
$\bar{y}_{reg}$	-	-	160312.5	526.316	-	-	1409113.542	384.025	-	-	13846.05	397.180	-	-	224.625	1119.677
$\bar{y}_{pr1}$	0.58	83.08	72692.31	1160.7	0.71	0.13	713838.3	758.06	0.92	42.13	11051.73	497.60	0.99	6.23	212.519	1183.458
$\bar{y}_{pr2}$	0.56	4.23	45870.36	1839.4	0.65	1.73	240765.7	2247.6	0.87	0.98	6819.164	806.5*	0.99	0.35	183.147	1373.254*
$\bar{y}$	-	-	843750	100.000	-	-	5411348.28	100.000	-	-	54993.75	100.000	-	-	2515.074	100.000

***indicates the largest PRE**

7. Discussion

The ratio-type class of estimators considered in Tables (IV) and (V) was adapted from the work of [17], where he made corrections on the MSE of the general class of [5] estimators. It is observed from Table (IV) that the proposed estimator (I) fares better at optimum condition than the unbiased sample mean, regression estimator and [5] class of estimators in all the four populations. This is evident on the larger Percent Relative Efficiencies (PREs) and the smaller Mean Squared Errors of the proposed estimator (I) than those of sample mean, regression and estimators of [5]. On the other hand, the proposed estimator (II) becomes more efficient than the simple random sample mean, regression estimator, the class of estimators of [5] and proposed estimator (I) in the four populations. This is evident on the fact that the proposed estimator (II) has the largest PRE in the four populations considered in this study. This therefore, shows that the proposed estimators are more efficient than any other proposed estimators that have less efficiency than the regression estimator and estimators of [5]. Table (V) clearly shows that [17] and our proposed estimators fare better than the class of estimators of [5], regression estimator and simple random sample mean in all the populations considered in this study. A clear difference is also observed between the class of estimators of [17] and the proposed estimators. In populations (I) and (II), estimator of [17], ($\bar{y}_{\mathrm{SS10}}$) fares better than the proposed estimators. Also, [17] estimator ($\bar{y}_{\mathrm{SS9}}$) is equally efficient with ($\bar{y}_{\mathrm{SS10}}$) and more efficient than the proposed estimators. On the other hand, the proposed estimators (I) and (II) fares better than [17] estimators in populations (III) and (IV), but the proposed estimator (II) is most efficient in the populations (III) and (IV). This indicates that the proposed estimators using exponential estimator may fare in some populations better than [17] class of estimators, while [17] may be more efficient than the proposed estimators in some other populations. On the whole, the proposed estimators have shown significant efficiencies in the four populations considered in this study. It can also be deduced that the proposed estimators always fare better than the usual regression estimator and [5].

8. Conclusion

From the above result and discussion, It can be concluded that the two proposed estimators at optimal condition are each more efficient that the general regression estimator which have always been preferred because of its minimum MSE. The two proposed estimators are also more efficient than most of the exiting ratio estimators, thus providing better alternative estimators in practical situations.

References

[1] Bahl, S. and Tuteja, R. K. *Ratio and product type exponential estimator*, Journal of Information and Optimization Sciences **12** (1), 159-164, 1991.

[2] Choudhury, S. and Singh, B. K. *An efficient class of dual to product-cum- dual to ratio estimators of finite population mean in sample surveys*, Global Journal of Science Frontier Research Mathematics and Decision Sciences **12** (3) Version 1.0 25-33, 2012.

[3] Choudhury, S. and Singh, B. K. *A class of chain ratio-cum-dual to ratio type estimator with two auxiliary characters under double sampling in sample surveys*, Statistics in Transition-New Series **13** (3), 519-536, 2012.

[4] Diana, G., Giordan, M. Perri P. F. *An improved class of estimators for the population mean*, Statistical methods and Applications **20** 123-140, 2011.

[5] Gupta, S. and Shabbir, J. *On the improvement in estimating the population mean in simple random sampling*, Journal of Applied Statistics **35** (5), 559-566, 2008.

[6] Haq, A. and Shabbir, J. *Improved family of ratio estimators in simple and stratified random Sampling*, Communications in Statistics - Theory and Methods **42** (5), 782-799, 2013.

[7] Haq, A, Shabbir, J. and Gupta, S. *Improved exponential ratio type estimators in stratified sampling*, Pakistan Journal of Statistics **29** (1), 13-31, 2013.

[8] Kadilar, C. and Cingi, H. *Ratio estimators in simple random sampling*, Applied Mathematics and Computation **151**, 893 - 902, 2004.

[9] Kadilar C. and Cingi H. *Improvement in estimating the population mean in simple random sampling*, Applied Mathematics Letters, **19** (1), 75-79, 2006.

[10] Kadilar, C. & Cingi, H. *An improvement in estimating the population mean by using the correlation coefficient*, Hacettepe Journal of Mathematics and Statistics **35** (1), 103-109, 2006.

[11] Kadilar, C. & Cingi, H. *New Ratio Estimators Using Correlation Coefficient.* http://interstat.statjournals.net/YEAR/2006/articles/0603004.pdf, 2006.

[12] Koyuncu, N. and *Kadilar C. Efficient estimators for the population mean*, Hacettepe Journal of Mathematics and Statistics, **38**(2), 217-225, 2009.

[13] Shabbir, J and Gupta, S. *On estimating finite population mean in simple and stratified sampling*, Communications in Statistics - Theory and Methods, **40**, 199-212, 2011.

[14] Singh, H. P. and Tailor, R. *Use of known correlation coefficient in estimating the finite population mean*, Statistics in Transition, **6** (4), 555-560, 2003.

[15] Singh, H.P. and Tailor, R. *Estimation of finite population mean with known coefficient of variation of an auxiliary character*, Statistica Anno **LXV** (3), 301-313, 2005.

[16] Singh, H. P. and Tailor, R. *Estimation of finite population mean using known correlation coefficient between auxiliary characters*, Statistica, Anno **LXV**(4), 407-418, 2005.

[17] Singh, H. P & Solanki, R. S. *An alternative procedure for estimating the population mean in simple random sampling*, Pakistan Journal of Statistics and Operations Research **VIII** (2), 213-232, 2012.

[18] Singh, H. P & Solanki, R. S. *Efficient ratio and product estimators in stratified random sampling*, Communications in Statistics - Theory and Methods, **42**,1008-1023, 2013.

[19] Singh, H. P & Solanki, R. S. *An Efficient class of estimators for the population mean using auxiliary information*, Communications in Statistics -Theory and Methods, **42** (5), 145-163, 2013.

[20] Subramani, J. & Kumarapandiyan, G. *Estimation of population mean using coefficient of variation and median of an auxiliary variable*, International Journal of Probability and Statistics **1** (4), 111 -118, 2012.

[21] Subramani, J. & Kumarapandiyan, G. *Modified ratio estimators using known median and coefficient of kurtosis.* American Journal of Mathematics and Statistics **2** (4), 95 -100, 2012.

[22] Subramani, J. & Kumarapandiyan, G. *Estimation of population mean using known median and coefficient of skweness* American Journal of Mathematics and Statistics**2** (5), 101-107, 2012.

[23] Subramani, J. & Kumarapandiyan, G. A class of modified ratio estimators using deciles of an auxiliary variable. *International Journal of Probability and Statistics.* 2(6):101 - 107, 2012.

[24] Singh, H. P and Solanki, R. S. *Improved estimation of population mean in simple random sampling using information on auxiliary attribute*, Applied Mathematics and Computation **218** (15), 7798-7812, 2012.

[25] Solanki, R. S., Singh, H. P.,and Rathour, A. *An alternative estimator for estimating finite population mean using auxiliary information in sample surveys*,ISRN Probability and Statistics **2012**, 1-14, 2012.

[26] Solanki, R. S., Singh, H. P., & Pal, S. K.*Improved estimation of finite population mean in sample surveys*,Journal of Advanced Computing **2** (2), 70-78, 2013.

[27] Upadhyaya, L. N.and Singh, H. P. *Use of transformed auxiliary variable in estimating the finite population mean*, Biometrical Journal **41** (5), 627-636, 1999.

[28] Yadav, S. K. & Kadilar, C. *Efficient family of exponential estimators for the population mean*, Hacettepe Journal of Mathematics and Statistics **42** (6), 671-677, 2013.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 707 – 713

Composite quantile regression for linear errors-in-variables models

Rong Jiang *

Abstract

Composite quantile regression can be more efficient and sometimes arbitrarily more efficient than least squares for non-normal random errors, and almost as efficient for normal random errors. Therefore, we extend composite quantile regression method to linear errors-in-variables models, and prove the asymptotic normality of the proposed estimators. Simulation results and a real dataset are also given to illustrate our the proposed methods.

2000 AMS Classification: primary 60G08, 62G10; secondary 62G20.

Keywords: Errors-in-variables; composite quantile regression.

Received 25/05/2014 : Accepted 17/07/2014 Doi : 10.15672/HJMS.2014237474

1. Introduction

Consider a linear errors-in-variables model as follows:

$$\begin{cases} Y = x^T\beta_0 + \varepsilon, \\ X = x + u, \end{cases} \tag{1.1}$$

where x is a p-dimensional vector of unobserved latent covariates which is measured in an error-prone way, X is the observed surrogate of x, β_0 is a p-dimensional unknown parameter vector, Y are responses vector, $(\varepsilon, u^T)^T$ is a p+1-dimensional spherical error vector, and they are independent with a common error distribution that is spherically symmetric. Spherically symmetric implies that ε and each component u have the same distribution, which ensures model identifiability. We restrict ourselves to structural models where x are independently and identically distributed random variables. If x stem from non-stochastic designs, the model is said to have a functional relationship, see Fuller (1987) for details. Model (1.1) belongs to a kind of model called the errors-in-variables model or measurement error model which was proposed by Deaton (1985) to correct for the effects of sampling error and is somewhat more practical than the ordinary regression model. Fuller (1987) gave a systematic survey on this research topic and present many

*Department of Applied Mathematics, Donghua University, Shanghai, 201620, China
Email: jrtrying@dhu.edu.cn

applications of measurement error data. Other references can see Cui (1997a), He and Liang (2000), Huang and Wang (2001), Ma and Tsiatis (2006), Schennach (2007), Liang and Li (2009), Wei and Carroll (2009), Hu and Cui (2009), Jiang et al. (2012a) and so on.

The composite quantile regression (CQR) was first proposed by Zou and Yuan (2008) for estimating the regression coefficients in the classical linear regression model. Zou and Yuan (2008) showed that the relative efficiency of the CQR estimator compared with the least squares estimator is greater than 70% regardless of the error distribution. Furthermore, the CQR estimator could be more efficient and sometimes arbitrarily more efficient than the least squares estimator. Other references about CQR method can see Kai, Li and Zou (2010), Kai, Li and Zou (2011), Tang et al. (2012a), Tang et al. (2012b), Guo et al. (2012) and Jiang et al. (2012b, 2012c, 2013, 2014a, 2014b). These nice theoretical properties of CQR in linear regression motivate us to consider linear errors-in-variables models based on CQR method so as to make the method of CQR more effective and convenient.

This paper is organized as follows. The main results are given in Section 2. Some simulations and a real data application are conducted in Section 3 to illustrate our methodology. Final remarks are given in Section 4. All the conditions and technical proofs are collected in the Appendix.

2. Methodology and main results

If the true covariates x are observed, the parameters β in model (1.1) can be estimated through (Zou and Yuan, 2008)

$$(\tilde{b}_1, ..., \tilde{b}_K, \tilde{\beta}) = argmin_{b_1,...,b_K,\beta} \sum_{k=1}^{K} \sum_{i=1}^{n} \rho_{\tau_k} \left(Y_i - b_k - x_i^T \beta \right),$$

where $\rho_{\tau_k}(r) = \tau_k r - rI(r < 0), k = 1, 2, \ldots, K$, be K check loss functions with $0 < \tau_1 < \tau_2 < \cdots \tau_K < 1$. Typically, we use the equally spaced quantiles: $\tau_k = \frac{k}{K+1}$ for $k = 1, 2, \ldots, K$. $\tilde{b}_k$ is estimator of b_{τ_k}, where $P(\varepsilon \le b_{\tau_k} | x_i) = \tau_k$, b_{τ_k} is the τ_k quantile of ε.

Taking into account the measurement error in X, we consider estimating β as follows

$$(\hat{a}_1, ..., \hat{a}_K, \hat{\beta}) = argmin_{a_1,...,a_K,\beta} \sum_{k=1}^{K} \sum_{i=1}^{n} \rho_{\tau_k} \left(\frac{Y_i - a_k - X_i^T \beta}{\sqrt{1 + \|\beta\|^2}} \right), \tag{2.1}$$

where $\hat{a}_k$ is the estimator of $b_{\tau_k} \sqrt{1 + \|\beta_0\|^2}$, $k = 1, \ldots, K$. The measurement error correction factor $\frac{1}{\sqrt{1+\|\beta\|^2}}$ is widely used in linear models with additive errors (see Ma and Yin, 2011) . The main intuition is the following. In the usual regression, one minimizes the vertical standardized distance $d\{(Y - a_k - X^T\beta)/s.d.(Y - a_k - X^T\beta)\}$ where d stands for a suitable distance measure and s.d. is the standard deviation, because only the vertical Y direction has errors. However, in the measurement error situation, errors also occur along the horizontal X direction, hence a distance containing both vertical and horizontal components should be favored. In fact, the minimization of the same standardized distance with X replaced by x automatically corrects for this. If we denote the variance of ε as Σ_ε and the variance-covariance matrix of u as Σ_u, we have

$$\frac{(Y - a_k - X^T\beta)}{s.d.(Y - a_k - X^T\beta)} = \frac{(Y - a_k - X^T\beta)}{\sqrt{\Sigma_\varepsilon + \beta^T \Sigma_u \beta}}$$

which is proportional to $(Y - a_k - X^T\beta)/\sqrt{1+\|\beta\|^2}$ under the spherical symmetry assumption. The following theorem gives the the asymptotic normality for the composite quantile regression estimator $\hat{\beta}$.

Theorem 1 Assuming Conditions A1-A2 in the Appendix are satisfied, then

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{L} N\left(0, \left(\sum_{k=1}^{K} f(b_{\tau_k})\right)^{-2} (1+\|\beta_0\|^2)\Sigma_x^{-1} S \Sigma_x^{-1}\right),$$

where $\xrightarrow{L}$ stands for convergence in distribution, $\Sigma_x = E(xx^T)$ and $S = \sum_{k,k'=1}^{K} \min(\tau_k, \tau_{k'})(1 - \max(\tau_k, \tau_{k'}))\Sigma_x + Cov\left[\sum_{k=1}^{K} \psi_{\tau_k}\left(\frac{\varepsilon - u^T\beta_0}{\sqrt{1+\|\beta_0\|^2}} - b_{\tau_k}\right)\left(u + \frac{(\varepsilon - u^T\beta_0)\beta_0}{1+\|\beta_0\|^2}\right)\right]$.

Remark 1: In practice, there is constant term in model (1.1), then model (1.1) can be write as

$$\begin{cases} Y = \alpha_0 + x^T\beta_0 + \varepsilon, \\ X = x + u. \end{cases}$$

The parameter α_0 and β_0 can be estimated as follows (Cui, 1997b)

$$(\hat{a}_1^*, ..., \hat{a}_K^*, \hat{\beta}^*) = argmin_{a_1,...,a_K,\beta} \sum_{k=1}^{K}\sum_{i=1}^{n} \rho_{\tau_k}\left(\frac{Y_i - \bar{Y} - a_k - (X_i - \bar{X})^T\beta}{\sqrt{1+\|\beta\|^2}}\right),$$
$$\hat{\alpha} = \bar{Y} - \bar{X}^T\hat{\beta}^*,$$

where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

3. Numerical studies

In this section, we conduct simulation studies to assess the finite sample performance of the proposed procedures and illustrate the proposed methodology on AIDS clinical trials. Furthermore, we compare CQR method with least square (LS) method proposed by Fuller (1987), t-type (TT) method proposed by Hu and Cui (2009) and quantile regression method with $\tau = 0.5$ ($\text{QR}_{0.5}$) proposed by He and Liang (2000).

3.1. Simulation example. We conduct a small simulation study with n = 100 and the data are generated from model (1.1), where the random error variables are taken to be 0.5*N(0,1), 0.2*t(3) and 0.05*C(0,1) distribution. The covariate vector $x = (x_1, x_2, \ldots, x_p)$ are generated from standard normal distribution N(0,1) and p=1,2,5 are considered. We focus on $K = 5$, $K = 9$ and $K = 19$ for composite quantile regression, respectively. The mean squared errors (MSE) and their standard deviations (STD) over 1000 simulations are summarized in Table 1, where MSE=$\|\hat{\beta} - \beta_0\|^2$. It can be seen from Table 1 that the CQR estimators have better performance than LS for heavy-tailed error distributions (t(3) and C(0,1)), but is less efficient when the error is normal distribution N(0,1). Moreover, the results show that our method is more efficient than TT and $\text{QR}_{0.5}$ in most cases. The three CQR estimators perform very similarly. Further, one can see that the three CQR estimators are close to the true value.

3.2. Real data example. In this section, we present an analysis of an AIDS clinical trial group (ACTG 315) study. One of the purposes of this study is to investigate the relationship between virologic and immunologic responses in AIDS clinical trials. In general, it is believed that the virologic response RNA (measured by viral load) and immunologic response (measured by CD4+ cell counts) are negatively correlated during treatment. Our preliminary investigations suggested that viral load depends linearly on CD4+ cell count. We therefore model the relationship between viral load and CD4+ cell counts by model (1.1). Let Y_i be the viral load and let x_i be the CD4+ cell count

Table 1 Simulation results for simulation example.

p	Method	N(0,1) MSE	N(0,1) STD	t(3) MSE	t(3) STD	C(0,1) MSE	C(0,1) STD
p=1	LS	0.0056	0.0077	0.0037	0.0063	1.8657	8.2551
	TT	0.0072	0.0082	0.0025	0.0035	0.0214	0.1407
	$QR_{0.5}$	0.0086	0.0095	0.0030	0.0039	0.0209	0.1422
	CQR_5	0.0064	0.0081	0.0022	0.0033	0.0219	0.1399
	CQR_9	0.0064	0.0084	0.0023	0.0034	0.0217	0.1407
	CQR_{19}	0.0063	0.0083	0.0022	0.0033	0.0315	0.1686
p=2	LS	0.0192	0.0204	0.0198	0.1749	3.2342	10.1476
	TT	0.0257	0.0284	0.0065	0.0074	0.6477	5.7348
	$QR_{0.5}$	0.0295	0.0330	0.0072	0.0083	0.6295	5.6614
	CQR_5	0.0216	0.0225	0.0061	0.0064	0.1525	0.5327
	CQR_9	0.0209	0.0216	0.0062	0.0066	0.1512	0.5065
	CQR_{19}	0.0206	0.0216	0.0061	0.0066	0.1586	0.5313
p=5	LS	0.1081	0.0724	0.0796	0.1203	4.7832	8.6034
	TT	0.1498	0.1010	0.0426	0.0309	0.8189	1.9204
	$QR_{0.5}$	0.1613	0.1100	0.0446	0.0326	0.7865	1.9108
	CQR_5	0.1203	0.0775	0.0397	0.0286	0.6306	0.9306
	CQR_9	0.1170	0.0765	0.0390	0.0283	0.6560	0.9254
	CQR_{19}	0.1156	0.0772	0.0396	0.0286	0.7147	0.9592

for subject i. To reduce the marked skewness of CD4+ cell counts, and make treatment times equal space, we take log-transformations of both variables. The x_i are measured with error (Liang et al., 2003). The model we used is

$$Y = x^T\beta_0 + \varepsilon, \quad X = x + u,$$

where X is the observed CD4+ cell counts. The performances of CQR method with different K are very similar (see Table 1), and considering computing time, K=5 is a good choice in practice. Therefore, K=5 for CQR method is considered in this example. The parameter estimator by using our proposed method is -0.0717. Moreover, the standard deviation of the parameter is 0.0078 and the 90% confidence interval is [-0.0824,-0.0576] by using the random weighting method (see Jiang et al., 2012a).

4. Conclusion

In this work, we have focused on the CQR method for linear errors-in-variables models and proven its nice theoretical properties. Moreover, the proposed approaches are demonstrated by simulation examples and a real data application.

Appendix

To prove main results in this paper, the following technical conditions are imposed.

A1. Assume (ε, u^T) is spherically symmetric with finite first moment, and the distribution functions F of ε are absolutely continuous, with continuous densities f uniformly bounded away from 0 and ∞ at the points $b_{\tau_k}, k = 1, \ldots, K$ and $E\varepsilon^2 < \infty$.

A2. $E(x) = 0$ and $\Sigma_x = E(xx^T)$ is positive definite.

Remark 2: Conditions A1-A2 are standard conditions, see He and Liang (2000).

Now we proceed to prove the theorems.

Proof of Theorem 1. Denote

$$
\begin{aligned}
f_{1ik}(b_k,\beta) =&\rho_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta}{\sqrt{1+\|\beta\|^2}}-b_{\tau_k}-\frac{x_i^T(\beta-\beta_0)}{\sqrt{1+\|\beta\|^2}}-(b_k-b_{\tau_k})\right)-\rho_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\\
&-\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\left(\frac{\varepsilon_i-u_i^T\beta}{\sqrt{1+\|\beta\|^2}}-\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-\frac{x_i^T(\beta-\beta_0)}{\sqrt{1+\|\beta\|^2}}-(b_k-b_{\tau_k})\right),\\
f_{2ik}(b_k,\beta) =&\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\left(\frac{\varepsilon_i-u_i^T\beta}{\sqrt{1+\|\beta\|^2}}-\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-\frac{x_i^T(\beta-\beta_0)}{\sqrt{1+\|\beta\|^2}}-(b_k-b_{\tau_k})\right).
\end{aligned}
$$

$(\hat a_1,...,\hat a_K,\hat\beta)$ is the minimizer of the following criterion:

$$
\begin{aligned}
&\sum_{k=1}^K\sum_{i=1}^n\left[\rho_{\tau_k}\left(\frac{Y_i-a_k-X_i^T\beta}{\sqrt{1+\|\beta\|^2}}\right)-\rho_{\tau_k}\left(\frac{Y_i-a_{\tau_k}-X_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}\right)\right]\\
=&\sum_{k=1}^K\sum_{i=1}^n f_{1ik}(b_k,\beta)+\sum_{k=1}^K\sum_{i=1}^n f_{2ik}(b_k,\beta)\\
\equiv&Q_n(b_1,...,b_K,\beta)
\end{aligned}
$$

Therefore, by applying the identity in Knight (1998)

$$
\rho_\tau(x-y)-\rho_\tau(x)=-y\psi_\tau(x)+\int_0^y\{I(x\le z)-I(x\le 0)\}dz.
$$

We have

$$
\begin{aligned}
EQ_n(b_1,...,b_K,\beta)=&\sum_{k=1}^K\sum_{i=1}^n E\left[\rho_{\tau_k}\left(\frac{Y_i-a_k-X_i^T\beta}{\sqrt{1+\|\beta\|^2}}\right)-\rho_{\tau_k}\left(\frac{Y_i-a_{\tau_k}-X_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}\right)\right]\\
=&\sum_{k=1}^K\sum_{i=1}^n E\left[\rho_{\tau_k}\left(\varepsilon_i-b_{\tau_k}-\frac{x_i^T(\beta-\beta_0)}{\sqrt{1+\|\beta\|^2}}-(b_k-b_{\tau_k})\right)-\rho_{\tau_k}\left(\varepsilon_i-b_{\tau_k}\right)\right]\\
=&\sum_{k=1}^K\sum_{i=1}^n E\left[\int_0^{\frac{x_i^T(\beta-\beta_0)}{\sqrt{1+\|\beta\|^2}}+(b_k-b_{\tau_k})}\{F(\varepsilon_i\le b_{\tau_k}+z|x_i)-F(\varepsilon_i\le z|x_i)\}\right]dz\\
\to&\frac{1}{2}\sum_{k=1}^K f(b_{\tau_k})(\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0))\left[\begin{array}{cc}1&0\\0&\frac{\Sigma_x}{1+\|\beta_0\|^2}\end{array}\right](\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0)^T)^T.
\end{aligned}
$$

Next, we study $f_{2ik}(b_k,\beta)$,

$$
\begin{aligned}
f_{2ik}(b_k,\beta)=&-\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\frac{1}{\sqrt{1+\|\beta_0\|^2}}\left(x_i+u_i+\frac{(\varepsilon_i-u_i^T\beta_0)\beta_0}{1+\|\beta_0\|^2}\right)^T(\beta-\beta_0)\\
&-\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)(b_k-b_{\tau_k})+R.
\end{aligned}
$$

Similar to the proof of Theorem 3 in Cui (1997a), we can obtain

$$
\begin{aligned}
&\sum_{i=1}^n[f_{1ik}(b_k,\beta)-Ef_{1ik}(b_k,\beta)]=o_p\left(\left\|\left(\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0)\right)\right\|\right)\\
&\sum_{i=1}^n[R-ER]=o_p\left(\left\|\left(\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0)\right)\right\|\right).
\end{aligned}
$$

Thus it follows that

$$
\begin{aligned}
&Q_n(b_1,...,b_K,\beta) \to Q_0(b_1,...,b_K,\beta)\\
=&-\frac{1}{\sqrt{n}}\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\frac{1}{\sqrt{1+\|\beta_0\|^2}}\left(x_i+u_i+\frac{(\varepsilon_i-u_i^T\beta_0)\beta_0}{1+\|\beta_0\|^2}\right)^T\sqrt{n}(\beta-\beta_0)\\
&-\frac{1}{\sqrt{n}}\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\sqrt{n}(b_k-b_{\tau_k})\\
&+\frac{1}{2}\sum_{k=1}^{K}f(b_{\tau_k})(\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0))\begin{bmatrix}1 & 0\\ 0 & \frac{\Sigma_x}{1+\|\beta_0\|^2}\end{bmatrix}(\sqrt{n}(b_k-b_{\tau_k}),\sqrt{n}(\beta-\beta_0)^T)^T.
\end{aligned}
$$

The convexity of the limiting objective function, $Q_0(b_1,...,b_K,\beta)$, assures the uniqueness of the minimizer and, consequently, that

$$
\sqrt{n}(\hat{\beta}-\beta_0)=\frac{\Sigma_x^{-1}\sqrt{1+\|\beta_0\|^2}}{\sqrt{n}\sum_{k=1}^{K}f(b_{\tau_k})}\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\left(\frac{\varepsilon_i-u_i^T\beta_0}{\sqrt{1+\|\beta_0\|^2}}-b_{\tau_k}\right)\left(x_i+u_i+\frac{(\varepsilon_i-u_i^T\beta_0)\beta_0}{1+\|\beta_0\|^2}\right)+o_p(1).
$$

The proof is completed.

References

[1] Cui, H. J. (1997a). Asymptotic normality of M-estimates in the EV model. Systems Science and Mathematical Sciences, 10, 225-236.

[2] Cui, H. J. (1997b). Asymptotic properties of generalized minimum L1-norm estimates in EV model. Science in China (Series A), in Chinese, 2, 119-131.

[3] Deaton, A. (1985). Panel data from a time series of cross-sections. Journal of Econometrics, 30, 109-126.

[4] Fuller, W. A. (1987). Measurement Error Models. Wiley, New York.

[5] Guo, J., Tian, M. and Zhu, K. (2012). New Efficient and robust estimation in varying-coefficient models with heteroscedasticity. Statistica Sinica, 22, 1075-1101.

[6] He, X. and Liang, H. (2000). Quantile regression estimates for a class of linear and partially linear errors-in-variables models. Statistica Sinica, 10, 129-140.

[7] Huang, Y. J. and Wang, C. Y. (2001). Consistent functional methods for logistic regression with errors in covariates. Journal of the American Statistical Association, 96, 1469-1482.

[8] Hu, T. and Cui, H. J. (2009). t-type estimators for a class of linear errors-in-variables models. Statistica Sinica, 19, 1013-1036.

[9] Jiang, R., Yang, X. H. and Qian, W. M. (2012a). Random weighting M-estimation for linear errors-in-variables models. Journal of the Korean Statistical Society, 41, 505-514.

[10] Jiang, R., Zhou, Z. G., Qian, W. M. and Shao, W. Q. (2012b). Single-index composite quantile regression. Journal of the Korean Statistical Society, 3, 323-332.

[11] Jiang, R., Qian, W. M. and Zhou, Z. G. (2012c). Variable selection and coefficient estimation via composite quantile regression with randomly censored data. Statistics & Probability Letters, 2, 308-317.

[12] Jiang, R., Zhou, Z. G., Qian, W. M. and Chen, Y. (2013). Two step composite quantile regression for single-index models. Computational Statistics and Data Analysis, 64, 180-191.

[13] Jiang, R., Qian, W. M. and Li, J. R. (2014a). Testing in linear composite quantile regression models, Computational Statistics, In Print.

[14] Jiang, R. Qian, W. M. and Zhou, Z. G. (2014b). Test for single-index composite quantile regression. Hacettepe Journal of Mathematics and Statistics, Accepted.

[15] Kai, B., Li, R. and Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. Annals of Statistics, 39, 305-332.

[16] Kai, B., Li, R. and Zou, H. (2010). Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. Journal of the Royal Statistical Society, Series B, 72, 49-69.
[17] Knight, K. (1998). Limiting distributions for L_1 regression estimators under general conditions. Annals of Statistics, 26, 755-770.
[18] Liang H., Wu H. L. and Carroll R. J. (2003). The relationship between virologic and immunologic responses in AIDS clinical research using mixed-effect varying-coefficient semiparametric models with measurement error. Biostatistics, 4, 297-312.
[19] Liang, H. and Li, R. (2009). Variable selection for partially linear models with measurement errors. Journal of the American Statistical Association, 104, 234-248.
[20] Ma, Y. and Yin, G. (2011). Censored quantile regression with covariate measurement errors. Statistica Sinica, 21, 949-971.
[21] Ma, Y. and Tsiatis, A. A. (2006). Closed form semiparametric estimators for measurement error models. Statistica Sinica, 16, 183-193.
[22] Schennach, M. (2007). Instrumental variable estimation of nonlinear errors-in-variables models. Econometrica, 75, 201-239.
[23] Tang, L., Zhou, Z. and Wu, C. (2012a). Weighted composite quantile estimation and variable selection method for censored regression model. Statistics & Probability Letters, 3, 653-663.
[24] Tang, L., Zhou, Z. and Wu, C. (2012b). Efficient estimation and variable selection for infinite variance autoregressive models. Journal of Applied Mathematics and Computing, 40, 399-413.
[25] Wei, Y. and Carroll, R. J. (2009). Quantile regression with measurement error. Journal of the American Statistical Association, 104, 1129-1143.
[26] Zou, H. and Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. Annals of Statistics, 36, 1108-1126.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 715 – 733

Exact moments of generalized order statistics from type II exponentiated log-logistic distribution

Devendra Kumar *

Abstract

In this paper some new simple expressions for single and product moments of generalized order statistics from type II exponentiated log-logistic distribution have been obtained. The results for order statistics and record values are deduced from the relations derived and some ratio and inverse moments of generalized order statistics are also carried out. Further, a characterization result of this distribution by using the conditional expectation of generalized order statistics is discussed.

2000 AMS Classification: 62G30, 62E10.

Keywords: Exact moments, ratio and inverse moments, generalized order statistics, order statistics, upper record values, type II exponentiated log-logistic distribution and characterization.

Received 07/11/2013 : Accepted 29/06/2014 Doi : 10.15672/HJMS.2014217472

1. Introduction

A random variable X is said to have type II exponentiated log-logistic distribution if its probability density function (*pdf*) is given by

$$f(x) = \frac{\alpha\beta(x/\sigma)^{\beta-1}}{\sigma[1+(x/\sigma)^{\beta}]^{\alpha+1}}, \quad x \geq 0,\ \alpha,\ \sigma > 0,\ \beta > 1 \tag{1.1}$$

and the corresponding survival function is

$$\bar{F}(x) = \left(1 + \left\{\frac{x}{\sigma}\right\}^{\beta}\right)^{-\alpha}, \quad x \geq 0,\ \alpha,\ \sigma > 0,\ \beta > 1. \tag{1.2}$$

It is easy to see that

$$\alpha\beta\bar{F}(x) = \sigma[1+(x/\sigma)^{\beta}]xf(x). \tag{1.3}$$

Log-logistic distribution is considered as a special case of type II exponentiated log-logistic distribution when $\alpha = 1$. It is used in survival analysis as a parametric model where in the mortality rate first increases then decreases, for example in cancer diagnosis

*Department of Statistics, Amity Institute of Applied Sciences Amity University, Noida-201 303, India
Email: devendrastats@gmail.com

or any other type of treatment. It has also been used in hydrology to model stream flow and precipitation, and in economics to model the distribution of wealth or income.

Kamps [24] introduced the concept of generalized order statistics (*gos*) as follows: Let $X_1, X_2 \ldots$ be a sequence of independent and identically distributed (*iid*) random variables (*rv*) with absolutely continuous cumulative distribution function (*cdf*) $F(x)$ and *pdf*, $f(x)$, $x \in (\alpha, \beta)$. Let $n \in N$, $n \geq 2$, $k > 0$, $m \in \Re$, be the parameters such that

$$\gamma_r = k + (n-r)(m+1) > 0, \quad \text{for all} \quad r \in \{1, 2, \ldots, n-1\},$$

where $M_r = \sum_{j=r}^{n-1} mj$. Then $X(1,n,m,k),\ldots,X(n,n,m,k)$, $r = 1, 2, \ldots n$ are called *gos* if their joint *pdf* is given by

$$k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{i=1}^{n-1} [1 - F(x_i)]^m f(x_i) \right) [1 - F(x_n)]^{k-1} f(x_n) \tag{1.4}$$

on the cone $F^{-1}(0) \leq x_1 \leq x_2 \leq \ldots \leq x_n \leq F^{-1}(1)$.
The model of *gos* contains as special cases, order statistics, record values, sequential order statistics.
Choosing the parameters appropriately (Cramer, [18]), we get the variant of the *gos* given in Table 1.

Table 1: Variants of the generalized order statistics

	$\gamma_n = k$	γ_r	m_r
i) Sequential order statistics	α_n	$(n-r+1)\alpha_r$	$\gamma_r - \gamma_{r+1} - 1$
ii) Ordinary order statistics	1	$n-r+1$	0
ii) Record values	1	1	-1
iv) Progressively type II censored order statistics	$R_n + 1$	$n - r + 1 + \sum_{j=r}^{n} R_j$	R_r
v) Pfeifer's record values	β_n	β_r	$\beta_r - \beta_{r+1} - 1$

For simplicity we shall assume $m_1 = m_2 = \ldots = m_{n-1} = m$.
The *pdf* of the $r-$th *gos*, $X(r,n,m,k)$, $1 \leq r \leq n$, is

$$f_{X(r,n,m,k)}(x) = \frac{C_{r-1}}{(r-1)!} [\bar{F}(x)]^{\gamma_r - 1} f(x) g_m^{r-1}(F(x)) \tag{1.5}$$

and the joint *pdf* of $X(r,n,m,k)$ and $X(s,n,m,k)$, $1 \leq r < s \leq n$, is

$$f_{X(r,n,m,k),X(s,n,m,k)}(x,y) = \frac{C_{s-1}}{(r-1)!(s-r-1)!} [\bar{F}(x)]^m f(x) g_m^{r-1}(F(x))$$
$$\times [h_m(F(y)) - h_m(F(x))]^{s-r-1} [\bar{F}(y)]^{\gamma_s - 1} f(y), \quad x < y, \tag{1.6}$$

where

$$\bar{F}(x) = 1 - F(x), \quad C_{r-1} = \prod_{i=1}^{r} \gamma_i \,, \quad \gamma_i = k + (n-i)(m+1),$$

$$h_m(x) = \begin{cases} -\frac{1}{m+1}(1-x)^{m+1}, & m \neq -1 \\ -ln(1-x), & m = -1 \end{cases}$$

and

$$g_m(x) = h_m(x) - h_m(1), \;\; x \in [0,1).$$

Theory of record values and its distributional properties have been extensively studied in the literature, Ahsanullah [4], Balakrishnan *et al*. [12], Nevzorov [33], Glick [21] and Arnold *et al*. [8, 9]. Resnick [35] discussed the asymptotic theory of records. Sequential order statistics have been studies by Arnold and Balakrishnan [7], Kamps [24], Cramer and Kamps [19] and Schenk [37], among others.
Aggarawala and Balakrishnan [1] established recurrence relations for single and product moments of progressive type II right censored order statistics from exponential and truncated exponential distributions. Balasooriya and saw [14] develop reliability sampling plans for the two parameter exponential distribution under progressive censoring. Balakrishnan *et al*. [13] obtained bounds for the mean and variance of progressive type II censored order statistics. Ordinary via truncated distributions and censoring schemes and particularly progressive type II censored order statistics have been discuss by Kamps [24] and Balakrishnan and Aggarwala [10], among others.
Kamps [24] investigated recurrence relations for moments of *gos* based on non-identically distributed random variables, which contains order statistics and record values as special cases. Cramer and Kamps [20] derived relations for expectations of functions of *gos* within a class of distributions including a variety of identities for single and product moments of ordinary order statistics and record values as particular cases. Various developments on *gos* and related topics have been studied by Kamps and Gather [23], Ahsanullah [5], Pawlas and Szynal [34], Kamps and Cramer [22], Ahmad and Fawzy [2], Ahmad [3], Kumar [27, 28, 29] among others. Characterizations based on *gos* have been studied by some authors, Keseling [25] characterized some continuous distributions based on conditional distributions of *gos* . Bieniek and Szynal [15] characterized some distributions via linearity of regression of *gos*. Cramer *et al*. [17] gave a unifying approach on characterization via linear regression of ordered random variables. Khan *et al*. [26] characterized some continuous distributions through conditional expectation of functions of *gos*.

The aim of the present study is to give some explicit expressions and recurrence relations for single and product moments of *gos* from type II exponentiated log-logistic distribution. In Section 2, we give the explicit expressions and recurrence relations for single moments of type II exponentiated log-logistic distribution and some inverse moments of *gos* are also worked out. Then we show that results for order statistics and record values are deduced as special cases. In Section 3, we present the explicit expressions and recurrence relations for product moments of type II exponentiated log-logistic distribution and we show that results for order statistics and record values are deduced as special cases and ratio moments of *gos* are also established. Section 4, provides a characterization result on type II exponentiated log-logistic distribution based on conditional moment of *gos*. Two applications are performed in Section 5. Some concluding remarks are given in Section 6.

2. Relations for single Moments

In this Section, the explicit expressions, recurrence relations for single moments of *gos* and inverse moments of *gos* are considered. First we need the basic result to prove the main Theorem.

2.1. Lemma. *For type II exponentiated log-logistic distribution as given in (1.2) and any non-negative and finite integers a and b with $m \neq -1$*

$$J_j(a,0) = \alpha\sigma^j \sum_{p=0}^{\infty} \frac{(-1)^p (j/\beta)_{(p)}}{[\alpha(a+1)+p-(j/\beta)]}, \quad \beta > j \;\; and \;\; j = 0,1,2,\ldots, \tag{2.1}$$

where

$$(\alpha)_{(i)} = \begin{cases} \alpha(\alpha+1)\ldots(\alpha+i-1), & i>0 \\ 1, & i=0 \end{cases}.$$

and

$$J_j(a,b) = \int_0^{\infty} x^j [\bar{F}(x)]^a f(x) g_m^b(F(x)) dx. \tag{2.2}$$

Proof From (2.2), we have

$$J_j(a,0) = \int_0^{\infty} x^j [\bar{F}(x)]^a f(x) dx. \tag{2.3}$$

By making the substitution $z = [\bar{F}(x)]^{1/\alpha}$ in (2.3), we get

$$J_j(a,0) = \alpha\sigma^j \int_0^{\infty} (1-z)^{j/\beta} z^{\alpha(a+1)-(j/\beta)-1} dz.$$

$$= \alpha\sigma^j \sum_{p=0}^{\infty} (-1)^p (j/\beta)_{(p)} \int_0^1 z^{\alpha(a+1)-(j/\beta)+p-1} dz$$

and hence the result given in (2.1).

2.2. Lemma. *For type II exponentiated log-logistic distribution as given in (1.2) and any non-negative and finite integers a and b*

$$J_j(a,b) = \frac{1}{(m+1)^b} \sum_{u=0}^{b} (-1)^u \binom{b}{u} J_j(a+u(m+1),0) \tag{2.4}$$

$$= \frac{\alpha\sigma^j}{(m+1)^b} \sum_{p=0}^{\infty} \sum_{u=0}^{b} (-1)^{p+u} \binom{b}{u} \frac{(j/\beta)_{(p)}}{[\alpha\{a+(m+1)u+1\}+p-(j/\beta)]}, \quad m \neq -1 \tag{2.5}$$

$$= \alpha^{b+1} \sigma^j b! \sum_{p=0}^{\infty} \frac{(j/\beta)_{(p)}}{[\alpha(a+1)+p-(j/\beta)]^{b+1}}, \quad m = -1, \tag{2.6}$$

where $J_j(a,b)$ is as given in (2.2).
Proof: On expanding $g_m^b(F(x)) = \left[\frac{1}{m+1}(1-(F(x))^{m+1})\right]^b$ binomially in (2.2), we get when $m \neq -1$

$$J_j(a,b) = \frac{1}{(m+1)^b} \sum_{u=0}^{b} (-1)^u \binom{b}{u} \int_0^{\infty} x^j [F(x)]^{a+u(m+1)} f(x) dx$$

$$= \frac{1}{(m+1)^b} \sum_{u=0}^{b} (-1)^u \binom{b}{u} J_j(a+u(m+1),0).$$

Making use of Lemma 2.1, we establish the result given in (2.5)
and when $m = -1$ that
$J_j(a,b) = \frac{0}{0}$ as $\sum_{u=0}^{b} (-1)^u \binom{b}{u} = 0$.
Since (2.5) is of the form $\frac{0}{0}$ at $m = -1$, therefore, we have

$$J_j(a,b) = A\sum_{u=0}^{b}(-1)^u \begin{pmatrix} b \\ u \end{pmatrix} \frac{[\alpha\{a+u(m+1)+1\}+p-(j/\beta)]^{-1}}{(m+1)^b}, \tag{2.7}$$

where

$$A = \alpha\sigma^j \sum_{p=0}^{\infty}(-1)^p (j/\beta)_{(p)}.$$

Differentiating numerator and denominator of (2.7) b times with respect to m, we get

$$J_j(a,b) = A\alpha^b \sum_{u=0}^{b}(-1)^{u+b} \begin{pmatrix} b \\ u \end{pmatrix} \frac{u^b}{[\alpha\{a+u(m+1)+1\}+p-(j/\beta)]^{b+1}}.$$

On applying the L' Hospital rule, we have

$$lim_{m\to -1} J_j(a,b) = A\alpha^b \sum_{u=0}^{b}(-1)^{u+b} \begin{pmatrix} b \\ u \end{pmatrix} \frac{u^b}{[\alpha(a+1)+p-(j/\beta)]^{b+1}}. \tag{2.8}$$

But for all integers $n \geq 0$ and for all real numbers x, we have Ruiz [36]

$$\sum_{i=0}^{n}(-1)^i \begin{pmatrix} n \\ i \end{pmatrix} (x-i)^n = n!. \tag{2.9}$$

Therefore,

$$\sum_{u=0}^{b}(-1)^{u+b} \begin{pmatrix} b \\ u \end{pmatrix} u^b = b!. \tag{2.10}$$

Now on substituting (2.10) in (2.8), we have the result given in (2.6).

2.3. Theorem. *For type II exponentiated log-logistic distribution as given in (1.2) and* $1 \leq r \leq n$, $k = 1, 2, \ldots$ *and* $m \neq -1$

$$E[X^j(r,n,m,k)] = \frac{C_{r-1}}{(r-1)!} J_j(\gamma_r - 1, r-1) \tag{2.11}$$

$$= \frac{\alpha\sigma^j C_{r-1}}{(r-1)!(m+1)^{r-1}} \sum_{p=0}^{\infty}\sum_{u=0}^{r-1}(-1)^{p+u} \begin{pmatrix} r-1 \\ u \end{pmatrix}$$

$$\times \frac{(j/\beta)_{(}p)}{[\alpha\gamma_{r-u}+p-(j/\beta)]}, \quad \beta > j \quad and \quad j = 0,1,2,\ldots \tag{2.12}$$

where $J_j(\gamma_r - 1, r-1)$ is as defined in (2.2).
Proof. From (1.5) and (2.2), we have

$$E[X^j(r,n,m,k)] = \frac{C_{r-1}}{(r-1)!} J_j(\gamma_r - 1, r-1)$$

Making use of Lemma 2.2, we establish the result given in (2.12).
Identity 2.1. For $\gamma_r \geq 1$, $k \geq 1$, $1 \leq r \leq n$ and $m \neq -1$

$$\sum_{u=0}^{r-1}(-1)^u \begin{pmatrix} r-1 \\ u \end{pmatrix} \frac{1}{\gamma_{r-u}} = \frac{(r-1)!(m+1)^{r-1}}{\prod_{t=1}^{r}\gamma_t}. \tag{2.13}$$

Proof. (2.13) can be proved by setting $j = 0$ in (2.12).
Special Cases

i) Putting $m = 0$, $k = 1$ in (2.12), the explicit formula for the single moments of order statistics of the type II exponentiated log-logistic distribution can be obtained as

$$E[X_{r:n}^j] = C_{r:n} \sum_{p=0}^{\infty} \sum_{u=0}^{r-1} (-1)^{u+p} \begin{pmatrix} r-1 \\ u \end{pmatrix} \frac{\alpha \sigma^j (j/\beta)_{(p)}}{[\alpha(n-r+u+1)+p-(j/\beta)]},$$

where

$$C_{r:n} = \frac{n!}{(r-1)!(n-r)!},$$

ii) Setting $m = -1$ in (2.12), we deduce the explicit expression for the single moments of upper k record values for type II exponentiated log-logistic distribution in view of (2.11) and (2.6) in the form

$$E[X^j(r,n,-1,k)] = E[(Z_r^{(k)})^j] = (\alpha k)^r \sigma^j \sum_{p=0}^{\infty} \frac{(-1)^p (j/\beta)_{(p)}}{[\alpha k + p - (j/\beta)]^r}$$

and hence for upper records

$$E[(Z_r^{(1)})^j] = E[X_{U(r)}^j] = \alpha^r \sigma^j \sum_{p=0}^{\infty} \frac{(-1)^p (j/\beta)_{(p)}}{[\alpha + p - (j/\beta)]^r}.$$

Recurrence relations for single moments of *gos* from (1.5) can be obtained in the following theorem.

2.4. Theorem. *For the distribution given in (1.2) and* $2 \le r \le n$, $n \ge 2$ *and* $k = 1, 2, \ldots,$

$$\Big(1 - \frac{\sigma j}{\alpha \beta \gamma_r}\Big) E[X^j(r,n,m,k)] = E[X^j(r-1,n,m,k)] + \frac{j\sigma^{\beta+1}}{\alpha\beta\gamma_r} E[X^{j-\beta}(r,n,m,k)]. \tag{2.14}$$

Proof. From (1.5), we have

$$E[X^j(r,n,m,k)] = \frac{C_{r-1}}{(r-1)!} \int_0^{\infty} x^j [\bar{F}(x)]^{\gamma_r - 1} f(x) g_m^{r-1}(F(x)) dx. \tag{2.15}$$

Integrating by parts treating $[\bar{F}(x)]^{\gamma_r-1} f(x)$ for integration and rest of the integrand for differentiation, we get

$$E[X^j(r,n,m,k)] = E[X^j(r-1,n,m,k)] + \frac{jC_{r-1}}{\gamma_r (r-1)!} \int_0^{\infty} x^{j-1} [\bar{F}(x)]^{\gamma_r} g_m^{r-1}(F(x)) dx$$

the constant of integration vanishes since the integral considered in (2.15) is a definite integral. On using (1.3), we obtain

$$\begin{aligned} &E[X^j(r,n,m,k)] - E[X^j(r-1,n,m,k)] \\ &= \frac{\sigma j C_{r-1}}{\alpha\beta\gamma_r (r-1)!} \int_0^{\infty} x^j [\bar{F}(x)]^{\gamma_r-1} f(x) g_m^{r-1}(F(x)) dx \\ &+ \frac{\sigma^{\beta+1} j C_{r-1}}{\alpha\beta\gamma_r (r-1)!} \int_0^{\infty} x^{j-\beta} [\bar{F}(x)]^{\gamma_r-1} f(x) g_m^{r-1}(F(x)) dx \end{aligned}$$

and hence the result given in (2.14).

Remark 2.1: Setting $m = 0$, $k = 1$, in (2.14), we obtain a recurrence relation for single moments of order statistics for type II exponentiated log-logistic distribution in the form

$$\Big(1 - \frac{\sigma j}{\alpha\beta(n-r+1)}\Big) E[X_{r:n}^j] = E[X_{r-1:n}^j] + \frac{j\sigma^{\beta+1}}{\alpha\beta(n-r+1)} E[X_{r-1:n}^{j-\beta}].$$

Remark 2.2: Putting $m=-1$, in Theorem 2.4, we get a recurrence relation for single moments of upper k record values from type II exponentiated log-logistic distribution in the form

$$\Big(1-\frac{\sigma j}{\alpha\beta k}\Big)E[(X_{U(r)}^{(k)})^{j}]=E[(X_{U(r-1)}^{(k)})^{j}]+\frac{j\sigma^{\beta+1}}{\alpha\beta k}E[(X_{U(r)}^{(k)})^{j-\beta}].$$

Inverse moments of *gos* from type II exponentiated log-logistic distribution can be obtain by the following Theorem.

2.5. Theorem. *For type II exponentiated log-logistic distribution as given in (1.2) and* $1\le r\le n,\ k=1,2,\ldots,$

$$E[X^{j-\beta}(r,n,m,k)]=\sum_{p=0}^{\infty}\frac{\sigma^{j-\beta}(-1)^{p}\Gamma\Big(\frac{j}{\beta}\Big)}{p!\Gamma\Big(\frac{j}{\beta}-p\Big)\prod_{i=1}^{r}\Big(1+\frac{p+1-(j/\beta)}{\alpha\gamma_i}\Big)},\quad \beta>j. \tag{2.16}$$

Proof. From (1.5), we have

$$E[X^{j-\beta}(r,n,m,k)]=\frac{C_{r-1}}{(r-1)!(m+1)^{r-1}}\sum_{u=0}^{r-1}(-1)^{u}\binom{r-1}{u}$$
$$\times\int_{0}^{\infty}x^{j-\beta}[\bar F(x)]^{\gamma_{r-u}-1}f(x)dx. \tag{2.17}$$

Now letting $t=[\bar F(x)]^{1/\alpha}$ in (2.17), we get

$$E[X^{j-\beta}(r,n,m,k)]=\frac{\sigma^{j-\beta}C_{r-1}}{(r-1)!(m+1)^{r}}\sum_{u=0}^{r-1}\sum_{p=0}^{\infty}(-1)^{u+p}\binom{r-1}{u}\frac{\Gamma\Big(\frac{j}{\beta}\Big)}{p!\Gamma\Big(\frac{j}{\beta}-p\Big)}$$
$$\times B\Big(\frac{k}{m+1}+n-r+u+\frac{p+1-(j/\beta)}{\alpha(m+1)},1\Big).$$

Since

$$\sum_{a=0}^{b}(-1)^{a}\binom{b}{a}B(a+k,c)=B(k,c+b) \tag{2.18}$$

where $B(a,b)$ is the complete beta function.
Therefore,

$$E[X^{j-\beta}(r,n,m,k)]=\frac{\sigma^{j-\beta}C_{r-1}}{(m+1)^{r}}\sum_{p=0}^{\infty}(-1)^{p}\frac{\Gamma\Big(\frac{j}{\beta}\Big)}{p!\Gamma\Big(\frac{j}{\beta}-p\Big)}$$
$$\times\frac{\Gamma\Big(\frac{\alpha\{k+(n-r)(m+1)\}+p+1-(j/\beta)}{\alpha(m+1)}\Big)}{\Gamma\Big(\frac{\alpha\{k+n(m+1)\}+p+1-(j/\beta)}{\alpha(m+1)}\Big)} \tag{2.19}$$

and hence the result given in (2.16).
Special Cases
iii) Putting $m=0,\ k=1$ in (2.19), we get inverse moments of order statistics from type II exponentiated log-logistic distribution as;

$$E[X_{r:n}^{j-\beta}]=\frac{\sigma^{j-\beta}n!}{(n-r)!}\sum_{p=0}^{\infty}\frac{(-1)^{p}\Gamma\Big(\frac{j}{\beta}\Big)\Gamma[\alpha(n-r+1)+p+1-(j/\beta)]}{p!\Gamma\Big(\frac{j}{\beta}-p\Big)\Gamma[\alpha(n+1)+p+1-(j/\beta)]}.$$

iv) Putting $m=-1$ in (2.16), to get inverse moments of k record values from type II exponentiated log-logistic distribution as;

$$E[X_{U(r)}^{j-\beta}]=\sigma^{j-\beta}\sum_{p=0}^{\infty}\frac{(-1)^p\Gamma\Big(\frac{j}{\beta}\Big)}{p!\Gamma\Big(\frac{j}{\beta}-p\Big)\Big(1+\frac{p+1-(j/\beta)}{\alpha k}\Big)^r}.$$

Recurrence relations for inverse moments of *gos* from (1.2) can be obtained in the following theorem.

2.6. Theorem. *For type II exponentiated log-logistic distribution and for* $2\le r\le n$, $n\ge 2$ $k=1,2,\dots,$

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta\gamma_r}\Big)E[X^{j-\beta}(r,n,m,k)]=E[X^{j-\beta}(r-1,n,m,k)]$$
$$+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta\gamma_r}E[X^{j-2\beta}(r,n,m,k)],\quad \beta>j. \tag{2.20}$$

Proof. The proof is easy.

Remark 2.3: Setting $m=0$, $k=1$ in (2.20), we obtain a recurrence relation for inverse moments of order statistics for type II exponentiated log-logistic distribution in the form

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta(n-r+1)}\Big)E[X_{r:n}^{j-\beta}]=E[X_{r-1:n}^{j-\beta}]+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta(n-r+1)}E[X_{r:n}^{j-2\beta}].$$

Remark 2.4: Putting $m=-1$, in Theorem 2.6, we get a recurrence relation for inverse moments of upper k record values from type II exponentiated log-logistic distribution in the form

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta k}\Big)E[(X_{U(r)}^{(k)})^{j-\beta}]=E[(X_{U(r-1)}^{(k)})^{j-\beta}]+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta k}E[(X_{U(r)}^{(k)})^{j-2\beta}].$$

3. Relations for product moments

In this Section, the explicit expressions and recurrence relations for single moments of *gos* and ratio moments of *gos* are considered. First we need the following Lemmas to prove the main result.

3.1. Lemma. *For type II exponentiated log-logistic distribution as given in (1.2) and any non-negative integers a, b, c with* $m\ne -1$

$$J_{i,j}(a,0,c)=\alpha^2\sigma^{i+j}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(-1)^{p+q}(j/\beta)_{(p)}(j/\beta)_{(q)}}{[\alpha(c+1)+p-(j/\beta)]}$$
$$\times\frac{1}{[\alpha(a+c+2)+p+q-\{(i+j)/\beta\}]}, \tag{3.1}$$

where

$$J_{i,j}(a,b,c)=\int_0^\infty\int_x^\infty x^iy^j[\bar{F}(x)]^af(x)[h_m(F(y))-h_m(F(x))]^b[\bar{F}(y)]^cf(y)dydx. \tag{3.2}$$

Proof: From (3.2), we have

$$J_{i,j}(a,0,c)=\int_0^\infty x^i[\bar{F}(x)]^af(x)G(x)dx, \tag{3.3}$$

where

$$G(x)=\int_x^\infty y^j[\bar{F}(y)]^cf(y)dy. \tag{3.4}$$

By setting $z = [\bar{F}(y)]^{1/\alpha}$ in (3.4), we find that

$$G(x) = \alpha\sigma^j \sum_{p=0}^{\infty} (-1)^p \frac{(j/\beta)_p [\bar{F}(x)]^{c+1+\{p-(j/\beta)\}/\alpha}}{[\alpha(c+1)+p-(j/\beta)]}.$$

On substituting the above expression of $G(x)$ in (3.3), we get

$$J_{i,j}(a,0,c) = \alpha\sigma^j \sum_{p=0}^{\infty} \frac{(-1)^p (j/\beta)_p}{[\alpha(c+1)+p-(j/\beta)]} \times \int_0^{\infty} x^i [\bar{F}(x)]^{a+c+1+\{p-(j/\beta)\}/\alpha} f(x) dx. \tag{3.5}$$

Again by setting $t = [\bar{F}(x)]^{1/\alpha}$ in (3.5) and simplifying the resulting expression, we derive the relation given in (3.1).

3.2. Lemma. *For the distribution as given in (1.2) and any non-negative integers a, b, c*

$$J_{i,j}(a,b,c) = \frac{1}{(m+1)^b} \sum_{v=0}^{b} (-1)^v \binom{b}{v} J_{i,j}(a+(b-v)(m+1), 0, c+v(m+1)) \tag{3.6}$$

$$= \frac{\alpha^2 \sigma^{i+j}}{(m+1)^b} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \sum_{v=0}^{b} (-1)^{p+q+v} \binom{b}{v} \frac{(j/\beta)_p}{[\alpha\{c+(m+1)v+1\}+p-(j/\beta)]} \times \frac{(i/\beta)_q}{[\alpha\{a+c+(m+1)b+2\}+p+q-\{(i+j)/\beta\}]}, \quad m \neq -1 \tag{3.7}$$

$$= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-1)^{p+q} \alpha^{b+2} \sigma^{i+j} b! \, (j/\beta)_{(p)} (i/\beta)_{(q)}}{[\alpha(c+1)+p-(j/\beta)]^{b+1} [\alpha(a+c+2)+p+q-\{(i+j)/\beta\}]}, \quad m = -1 \tag{3.8}$$

where $J_{i,j}(a,b,c)$ is as given in (3.2).

Proof: When $m \neq -1$, we have

$$[h_m(F(y)) - h_m(F(x))]^b = \frac{1}{(m+1)^b} [(\bar{F}(x))^{m+1} - (\bar{F}(y))^{m+1}]^b$$

$$= \frac{1}{(m+1)^b} \sum_{v=0}^{b} (-1)^v \binom{b}{v} [\bar{F}(y)]^{v(m+1)} [\bar{F}(x)]^{(b-v)(m+1)}.$$

Now substituting for $[h_m(F(y)) - h_m(F(x))]^b$ in equation (3.2), we get

$$J_{i,j}(a,b,c) = \frac{1}{(m+1)^b} \sum_{v=0}^{b} (-1)^v \binom{b}{v} J_{i,j}(a+(b-v)(m+1), 0, c+v(m+1)).$$

Making use of the Lemma 3.1, we derive the relation given in (3.7).
When $m = -1$, we have
$J_{i,j}(a,b,c) = \frac{0}{0}$ as $\sum_{v=0}^{b} (-1)^v \binom{b}{v} = 0$.
On applying L' Hospital rule, (3.8) can be proved on the lines of (2.6).

3.3. Theorem. *For type II exponentiated log-logistic distribution as given in (1.2) and* $1 \leq r < s \leq n$, $k = 1, 2, \ldots$ *and* $m \neq -1$

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] = \frac{C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{r-1}} \sum_{u=0}^{r-1} (-1)^u$$

$$\times \begin{pmatrix} r-1 \\ u \end{pmatrix} J_{i,j}(m+u(m+1), s-r-1, \gamma_s - 1) \tag{3.9}$$

$$= \frac{\alpha^2 \sigma^{i+j} C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{s-2}} \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{u=0}^{r-1}\sum_{v=0}^{s-r-1} (-1)^{p+q+u+v} \begin{pmatrix} r-1 \\ u \end{pmatrix}$$

$$\times \begin{pmatrix} s-r-1 \\ v \end{pmatrix} \frac{(j/\beta)_{(p)}\ (i/\beta)_{(q)}}{[\alpha\gamma_{s-v} + p - (j/\beta)][\alpha\gamma_{r-u} + p + q - \{(i+j)/\beta\}]},$$

$$\beta > max(i,j) \ \ and \ \ i,j = 0,1,2,\ldots. \tag{3.10}$$

Proof: From (1.6), we have

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] = \frac{C_{s-1}}{(r-1)!(s-r-1)!} \int_0^\infty \int_x^\infty x^i y^j [\bar{F}(x)]^m f(x)$$

$$\times g_m^{r-1}(F(x))[h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s - 1} f(y)dydx. \tag{3.11}$$

On expanding $g_m^{r-1}(F(x))$ binomially in (3.11), we get

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] = \frac{C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{r-1}}$$

$$\times \sum_{u=0}^{r-1} (-1)^u \begin{pmatrix} r-1 \\ u \end{pmatrix} J_{i,j}(m+u(m+1), s-r-1, \gamma_s - 1).$$

Making use of the Lemma 3.2, we derive the relation in (3.10).

Identity 3.1: For $\gamma_r, \gamma_s \geq 1$, $k \geq 1$, $1 \leq r < s \leq n$ and $m \neq -1$

$$\sum_{v=0}^{s-r-1} (-1)^v \begin{pmatrix} s-r-1 \\ v \end{pmatrix} \frac{1}{\gamma_{s-v}} = \frac{(s-r-1)!(m+1)^{s-r-1}}{\prod_{t=r+1}^{s} \gamma_t}. \tag{3.12}$$

Proof. At $i = j = 0$ in (3.10), we have

$$1 = \frac{C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{s-2}} \sum_{u=0}^{r-1}\sum_{v=0}^{s-r-1} (-1)^{u+v}$$

$$\times \begin{pmatrix} r-1 \\ u \end{pmatrix} \begin{pmatrix} s-r-1 \\ v \end{pmatrix} \frac{1}{\gamma_{s-v}\gamma_{r-u}}.$$

Now on using (2.13), we get the result given in (3.12).

At $r = 0$, (3.12) reduces to (2.13).

Special cases:

i) Putting $m = 0$, $k = 1$ in (3.10), the explicit formula for the product moments of order statistics of the type II exponentiated log-logistic distribution can be obtained as

$$E(X_{r:n}^i X_{s:n}^j) = \alpha^2 \sigma^{i+j} C_{r,s:n} \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{u=0}^{r-1}\sum_{v=0}^{s-r-1} (-1)^{p+q+u+v} \begin{pmatrix} n-s \\ u \end{pmatrix}$$

$$\times \begin{pmatrix} s-r-1 \\ v \end{pmatrix} \frac{(j/\beta)_{(p)}}{[\alpha(n-s+1+v) + p - (j/\beta)]}$$

$$\times \frac{(i/\beta)_{(q)}}{[\alpha(n-r+1+u) + p + q - \{(i+j)/\beta\}]}.$$

where,

$$C_{r,s:n} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}.$$

ii)Putting $m=-1$ in (3.10), we deduce the explicit expression for the product moments of upper k record values for the type II exponentiated log-logistic distribution in view of (3.9) and (3.8) in the form

$$E[(X_{U(r)}^{(k)})^i (X_{U(s)}^{(k)})^j)] = (\alpha k)^s \sigma^{i+j} \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(-1)^{p+q}(j/\beta)_{(p)}}{[\alpha k + p - (j/\beta)]^{s-r}}$$
$$\times \frac{(i/\beta)_{(q)}}{[\alpha k + p + q - \{(i+j)/\beta\}]^r}$$

and hence for upper records

$$E(X_{U(r)}^i X_{U(s)}^j) = \alpha^s \sigma^{i+j} \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(-1)^{p+q}(j/\beta)_{(p)}(i/\beta)_{(q)}}{[\alpha + p - (j/\beta)]^{s-r}[\alpha + p + q - \{(i+j)/\beta\}]^r}.$$

Remark 3.1 At $j=0$ in (3.10), we have

$$E[X^i(r,n,m,k) = \frac{\alpha\sigma^i C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{s-2}} \sum_{q=0}^{\infty}\sum_{u=0}^{r-1}\sum_{v=0}^{s-r-1}(-1)^{q+u+v}$$
$$\times \binom{r-1}{u}\binom{s-r-1}{v} \frac{(i/\beta)_q}{\gamma_{s-v}[\alpha\gamma_{r-u} + q - (i/\beta)]}. \tag{3.12}$$

Making use of (3.12) in (3.13) and simplifying the resulting expression, we get

$$E[X^i(r,n,m,k) = \frac{\alpha\sigma^i C_{r-1}}{(r-1)!(m+1)^{s-1}} \sum_{q=0}^{\infty}\sum_{u=0}^{r-1}(-1)^{q+u}$$
$$\times \binom{r-1}{u} \frac{(i/\beta)_q}{[\alpha\gamma_{r-u} + q - (i/\beta)]},$$

as obtained in (2.12).
Making use of (1.6), we can derive recurrence relations for product moments of *gos* from (1.2).

3.4. Theorem. *For the given type II exponentiated log-logistic distribution and $n \in N$, $m \in \Re$, $1 \le r < s \le n-1$*

$$\Big(1 - \frac{\sigma j}{\alpha\beta\gamma_s}\Big) E[X^i(r,n,m,k)X^j(s,n,m,k)] = E[X^i(r,n,m,k)X^j(s-1,n,m,k)]$$
$$+ \frac{j\sigma^{\beta+1}}{\alpha\beta\gamma_s} E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)]. \tag{3.14}$$

Proof: From (1.6), we have

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] = \frac{C_{s-1}}{(r-1)!(s-r-1)!}$$
$$\times \int_0^{\infty} x^i [\bar{F}(x)]^m f(x) g_m^{r-1}(F(x)) I(x) dx, \tag{3.15}$$

where

$$I(x) = \int_x^{\infty} y^j [\bar{F}(y)]^{\gamma_s - 1} [h_m(F(y)) - h_m(F(x))]^{s-r-1} f(y) dy.$$

Solving the integral in $I(x)$ by parts and substituting the resulting expression in (3.15), we get

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] - E[X^i(r,n,m,k)X^j(s-1,n,m,k)]$$
$$= \frac{jC_{s-1}}{\gamma_s(r-1)!(s-r-1)!}\int_0^\infty\int_x^\infty x^i y^j[\bar{F}(x)]^m f(x) g_m^{r-1}(F(x))$$
$$\times[h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s} dydx$$

the constant of integration vanishes since the integral in $I(x)$ is a definite integral. On using the relation (1.3), we obtain

$$E[X^i(r,n,m,k)X^j(s,n,m,k)] - E[X^i(r,n,m,k)X^j(s-1,n,m,k)]$$
$$= \frac{j\sigma C_{s-1}}{\alpha\beta\gamma_s(r-1)!(s-r-1)!}\int_0^\infty\int_x^\infty x^i y^j[\bar{F}(x)]^m f(x) g_m^{r-1}(F(x))$$
$$\times[h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s-1} f(y)dydx$$
$$+\frac{j\sigma^{\beta+1} C_{s-1}}{\alpha\beta\gamma_s(r-1)!(s-r-1)!}\int_0^\infty\int_x^\infty x^i y^{j-\beta}[\bar{F}(x)]^m f(x) g_m^{r-1}(F(x))$$
$$\times[h_m(F(y)) - h_m(F(x))]^{s-r-1}[\bar{F}(y)]^{\gamma_s-1} f(y)dydx$$

and hence the result given in (3.14).

Remark 3.2 Setting $m = 0,\ k = 1$ in (3.14), we obtain recurrence relations for product moments of order statistics of the type II exponentiated log-logistic distribution in the form

$$\Big(1 - \frac{\sigma j}{\alpha\beta(n-s+1)}\Big)E[X_{r,s:n}^{i,j}] = E[X_{r,s-1:n}^{i,j}] + \frac{j\sigma^{\beta+1}}{\alpha\beta(n-s+1)}E[X_{r,s:n}^{i,j-\beta}].$$

Remark 3.3 Putting $m = -1$, $k \geq 1$ in (3.5), we get the recurrence relations for product moments of upper k records of the type II exponentiated log-logistic distribution in the form

$$\Big(1 - \frac{\sigma j}{\alpha\beta k}\Big)E[(X_{U(r)}^{(k)})^i(X_{U(s)}^{(k)})^j] = E[(X_{U(r)}^{(k)})^i(X_{U(s-1)}^{(k)})^j]$$
$$+\frac{j\sigma^{\beta+1}}{\alpha\beta k}E[(X_{U(r)}^{(k)})^i(X_{U(s-1)}^{(k)})^{j-\beta}].$$

Ratio moments of *gos* from type II exponentiated log-logistic distribution can be obtain by the following Theorem.

3.5. Theorem. *For type II exponentiated log-logistic distribution as given in (1.2)*

$$E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)] = \sum_{p=0}^\infty\sum_{q=0}^\infty \frac{(-1)^{p+q}\sigma^{i+j-\beta}\Gamma\Big(\frac{j}{\beta}\Big)\Gamma\Big(\frac{i}{\beta}+1\Big)}{p!q!\Gamma\Big(\frac{j}{\beta}-p\Big)\Gamma\Big(\frac{i}{\beta}+1-p\Big)}$$
$$\times\frac{1}{\prod_{a=1}^r\Big(1+\frac{p+q-((i+j)/\beta)}{\alpha\gamma_a}\Big)\prod_{b=r+1}^s\Big(1+\frac{p+1-(j/\beta)}{\alpha\gamma_b}\Big)},\quad \beta > j. \tag{3.16}$$

Proof From (1.6), we have

$$E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)] = \frac{C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{s-2}}$$
$$\times\sum_{u=0}^{r-1}\sum_{v=0}^{s-r-1}(-1)^{u+v}\binom{r-1}{u}\binom{s-r-1}{v}$$
$$\times\int_0^\infty x^i[\bar{F}(x)]^{(s-r+u-v)(m+1)-1} f(x)J(x)dx, \tag{3.17}$$

where

$$J(x)=\int_x^\infty y^{j-\beta}[\bar{F}(y)]^{\gamma_{s-v}-1}f(y)dy. \tag{3.18}$$

By setting $z=[\bar{F}(y)]^{1/\alpha}$ in (3.18), we find that

$$J(x)=\sigma^{j-\beta}\sum_{p=0}^{\infty}\frac{(-1)^p\Gamma\left(\frac{j}{\beta}\right)[\bar{F(x)}]^{\gamma_{s-v}+\frac{p+1-(j/\beta)}{\alpha}}}{p!\Gamma\left(\frac{j}{\beta}-p\right)\left[\gamma_{s-v}+\frac{p+1-(j/\beta)}{\alpha}\right]}.$$

On substituting the above expression of $J(x)$ in (3.17), we get

$$E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)]=\frac{\sigma^{j-\beta}C_{s-1}}{(r-1)!(s-r-1)!(m+1)^{s-2}}\sum_{p=0}^{\infty}\sum_{u=0}^{r-1}$$

$$\times\sum_{v=0}^{s-r-1}(-1)^{u+v+p}\binom{r-1}{u}\binom{s-r-1}{v}\frac{\Gamma\left(\frac{j}{\beta}\right)}{p!\Gamma\left(\frac{j}{\beta}-p\right)}$$

$$\times\frac{1}{\left[\gamma_{s-v}+\frac{p+1-(j/\beta)}{\alpha}\right]}\int_0^\infty x^i[\bar{F}(x)]^{\gamma_{r-u}+\frac{p+1-(j/\beta)-1}{\alpha}-1}dx. \tag{3.19}$$

Again by setting $t=[\bar{F}(x)]^{1/\alpha}$ in (3.19), we get

$$E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)]=\frac{\sigma^{i+j-\beta}C_{s-1}}{(m+1)^s}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}(-1)^{p+q}$$

$$\times\frac{\Gamma\left(\frac{j}{\beta}\right)\Gamma\left(\frac{i}{\beta}+1\right)\Gamma\left[\frac{\alpha\{k+(n-r)(m+1)\}+p+q-\{(i+j)/\beta\}}{\alpha(m+1)}\right]}{p!q!\ \Gamma\left(\frac{j}{\beta}-p\right)\Gamma\left(\frac{i}{\beta}+1-q\right)\Gamma\left[\frac{\alpha\{k+n(m+1)\}+p+q-\{(i+j)/\beta\}}{\alpha(m+1)}\right]}$$

$$\times\frac{\Gamma\left[\frac{\alpha\{k+(n-s)(m+1)\}+p+1-(j/\beta)}{\alpha(m+1)}\right]}{\Gamma\left[\frac{\alpha\{k+(n-r)(m+1)\}+p+1-(j/\beta)}{\alpha(m+1)}\right]} \tag{3.20}$$

and hence the result given in (3.16).

Special cases

iii) Putting $m=0$, $k=1$ in (3.20), the explicit formula for the ratio moments of order statistics of the type II exponentiated log-logistic distribution can be obtained as

$$E[X^i_{r:n}X^{j-\beta}_{s:n}]=\frac{n!\sigma^{i+j-\beta}}{(n-s)!}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}(-1)^{p+q}\frac{\Gamma\left(\frac{j}{\beta}\right)\Gamma\left(\frac{i}{\beta}+1\right)}{p!q!\ \Gamma\left(\frac{j}{\beta}-p\right)\Gamma\left(\frac{i}{\beta}+1-q\right)}$$

$$\times\frac{\Gamma[\alpha(n-r+1)+p+q-\{(i+j)/\beta\}]\Gamma[\alpha(n-s+1)+p+1-(j/\beta)]}{\Gamma[\alpha(n+1)+p+q-\{(i+j)/\beta\}]\Gamma[\alpha(n-r+1)+p+1-(j/\beta)]}.$$

iv) Putting $m=-1$ in (3.16), the explicit expression for the ratio moments of upper k record values for the type II exponentiated log-logistic distribution can be obtained as

$$E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s)})^{j-\beta}]=\sigma^{i+j-\beta}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{(-1)^{p+q}}{p!q!}\frac{\Gamma\left(\frac{j}{\beta}\right)\Gamma\left(\frac{i}{\beta}+1\right)}{\Gamma\left(\frac{j}{\beta}-p\right)\Gamma\left(\frac{i}{\beta}+1-q\right)}$$

$$\times\frac{1}{\left(1+\frac{p+q-\{(i+j)/\beta\}}{\alpha k}\right)^r\left(1+\frac{p+1-(j/\beta)}{\alpha k}\right)^{s-r}}.$$

Making use of (1.6), we can derive recurrence relations for ratio moments of *gos* from (1.2).

3.6. Theorem. *For type II exponentiated log-logistic distribution*

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta\gamma_s}\Big)E[X^i(r,n,m,k)X^{j-\beta}(s,n,m,k)]$$
$$=E[X^i(r,n,m,k)X^{j-\beta}(s-1,n,m,k)]$$
$$+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta\gamma_s}E[X^i(r,n,m,k)X^{j-2\beta}(s,n,m,k)],\quad \beta>j. \tag{3.21}$$

Proof The proof is easy.

Remark 3.4 Setting $m=0$, $k=1$ in (3.21), we obtain a recurrence relation for Ratio moments of order statistics for type II exponentiated log-logistic distribution in the form

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta(n-s+1)}\Big)E[X^i_{r:n}X^{j-\beta}_{s:n}]=E[X^i_{r:n}X^{j-\beta}_{s-1:n}]+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta(n-s+1)}E[X^i_{r:n}X^{j-2\beta}_{s:n}].$$

Remark 3.5 Putting $m=-1$, in Theorem 3.6, we get a recurrence relation for ratio moments of upper k record values from type II exponentiated log-logistic distribution in the form

$$\Big(1-\frac{\sigma(j-\beta)}{\alpha\beta k}\Big)E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s)})^{j-\beta}]=E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s-1)})^{j-\beta}]$$
$$+\frac{(j-\beta)\sigma^{\beta+1}}{\alpha\beta k}E[(X^{(k)}_{U(r)})^i(X^{(k)}_{U(s)})^{j-2\beta}].$$

Remark 3.6 At $\gamma_r=n-r+1+\sum_{i=r}^{j}m_i$, $1\le r\le j\le n$, $m_i\in N$, $k=m_n+1$ in (3.16) the product moment of progressive type II censored order statistics of type II exponentiated log-logistic distribution can be obtained.

Remark 3.7 The result is more general in the sense that by simply adjusting $j-\beta$ in (3.16), we can get interesting results. For example if $j-\beta=-1$ then $E\Big[\frac{X(r,n,m,k)}{X(s,n,m,k)}\Big]^i$ gives the moments of quotient. For $j-\beta>0$, $E[X^i(r,n,m,k)\ X^{j-\beta}(s,n,m,k)]$ represent product moments, whereas for $j<\beta$, it is moment of the ratio of two generalized order statistics of different powers.

4. Characterization

This Section contains characterization of type II exponentiated log-logistic distribution by using the conditional expectation of *gos* .
Let $X(r,n,m,k)$, $r=1,2,\dots,n$ be *gos*, then from a continuous population with *cdf* $F(x)$ and *pdf* $f(x)$, then the conditional *pdf* of $X(s,n,m,k)$ given $X(r,n,m,k)=x$, $1\le r<s\le n$, in view of (1.5) and (1.6), is

$$f_{X(s,n,m,k)|X(r,n,m,k)}(y|x)=\frac{C_{s-1}}{(s-r-1)!C_{r-1}}$$
$$\times\frac{[h_m(F(y))-h_m(F(x))]^{s-r-1}[F(y)]^{\gamma_s-1}}{[\bar{F}(x)]^{\gamma_{r+1}}}f(y).\quad x<y \tag{4.1}$$

4.1. Theorem. *Let X be a non-negative random variable having an absolutely continuous distribution function $F(x)$ with $F(0)=0$ and $0<F(x)<1$ for all $x>0$, then*

$$E[X(s,n,m,k)|X(r,n,m,k)=x]=\sigma\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^{p}$$
$$\times\prod_{j=1}^{s-r}\Big(\frac{\gamma_{r+j}}{\gamma_{r+j}-p/\alpha}\Big) \tag{4.2}$$

if and only if

$$\bar{F}(x)=\Big(1+\Big\{\frac{x}{\sigma}\Big\}^{\beta}\Big)^{-\alpha},\quad x\geq 0,\ \alpha,\ \sigma>0,\ \beta>1.$$

Proof From (4.1), we have

$$E[X(s,n,m,k)|X(r,n,m,k)=x]=\frac{C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}$$
$$\times\int_{x}^{\infty}y\Big[1-\Big(\frac{\bar{F}(y)}{\bar{F}(x)}\Big)^{m+1}\Big]^{s-r-1}\Big(\frac{\bar{F}(y)}{\bar{F}(x)}\Big)^{\gamma_s-1}\frac{f(y)}{\bar{F}(x)}dy. \tag{4.3}$$

By setting $u=\frac{\bar{F}(y)}{\bar{F}(x)}=\Big(\frac{1+(x/\sigma)^{\beta}}{1+(y/\sigma)^{\beta}}\Big)^{\alpha}$ from (1.2) in (4.3), we obtain

$$E[X(s,n,m,k)|X(r,n,m,k)=x]=\frac{\sigma C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}$$
$$\times\int_{0}^{1}[\{1+(x/\sigma)^{\beta}\}u^{-1/\alpha}-1]^{1/\beta}u^{\gamma_s-1}(1-u^{m+1})^{s-r-1}du$$
$$=\frac{\sigma C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^{p}$$
$$\times\int_{0}^{1}u^{\gamma_s-(p/\alpha)-1}(1-u^{m+1})^{s-r-1}du \tag{4.4}$$

Again by setting $t=u^{m+1}$ in (4.4), we get

$$E[X(s,n,m,k)|X(r,n,m,k)=x]$$
$$=\frac{\sigma C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r}}\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^{p}$$
$$\times\int_{0}^{1}t^{\frac{k-(p/\alpha)}{m+1}+n-s-1}(1-t)^{s-r-1}dt$$
$$=\frac{\sigma C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r}}\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^{p}$$
$$\times\frac{\Gamma\Big(\frac{k-(p/\alpha)}{m+1}+n-s\Big)\Gamma(s-r)}{\Gamma\Big(\frac{k-(p/\alpha)}{m+1}+n-r\Big)}$$
$$=\frac{\sigma C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r}}\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^{p}$$
$$\times\frac{(m+1)^{s-r}\Gamma(s-r)}{\prod_{j=1}^{s-r}(\gamma_{r+j}-(p/\alpha))}$$

and hence the relation in (4.2).

To prove sufficient part, we have from (4.1) and (4.2)

$$\frac{C_{s-1}}{(s-r-1)!C_{r-1}(m+1)^{s-r-1}}\int_{x}^{\infty}y[(\bar{F}(x))^{m+1}-(\bar{F}(y))^{m+1}]^{s-r-1}$$

$$\times[\bar{F}(y)]^{\gamma_s-1}f(y)dy = [\bar{F}(x)]^{\gamma_{r+1}}H_r(x), \tag{4.8}$$

where

$$H_r(x) = \sigma\sum_{p=0}^{\infty}(1/\beta)_{(p)}[1+(x/\sigma)^{\beta}]^p\prod_{j=1}^{s-r}\Big(\frac{\gamma_{r+j}}{\gamma_{r+j}-p/\alpha}\Big).$$

Differentiating (4.5) both sides with respect to x and rearranging the terms, we get

$$-\frac{C_{s-1}[\bar{F}(x)]^m f(x)}{(s-r-2)!C_{r-1}(m+1)^{s-r-2}}\int_x^{\infty} y[(\bar{F}(x))^{m+1}-(\bar{F}(y))^{m+1}]^{s-r-2}$$
$$\times[\bar{F}(y)]^{\gamma_s-1}f(y)dy = H_r'(x)[\bar{F}(x)]^{\gamma_{r+1}} - \gamma_{r+1}H_r(x)[\bar{F}(x)]^{\gamma_{r+1}-1}f(x)$$

or

$$-\gamma_{r+1}H_{r+1}(x)[\bar{F}(x)]^{\gamma_{r+2}+m}f(x)$$
$$= H_r'(x)[\bar{F}(x)]^{\gamma_{r+1}} + \gamma_{r+1}H_r(x)[\bar{F}(x)]^{\gamma_{r+1}-1}f(x).$$

Therefore,

$$\frac{f(x)}{\bar{F}(x)} = -\frac{H_r'(x)}{\gamma_{r+1}[H_{r+1}(x)-H_r(x)]} = \frac{\alpha\beta(x/\sigma)^{\beta-1}}{\sigma[1+(x/\sigma)^{\beta}]}$$

which proves that

$$\bar{F}(x) = \Big(1+\Big\{\frac{x}{\sigma}\Big\}^{\beta}\Big)^{-\alpha}, \quad x\geq 0,\ \alpha,\ \sigma>0,\ \beta>1.$$

Remark For $m=0$, $k=1$ and $m=-1$, $k=1$, we obtain the characterization results of the type II exponentiated log-logistic distribution based on order statistics and record values respectively.

5. Applications

In this Section, we suggest some applications based on moments discussed in Section 2. Order statistics, record values and their moments are widely used in statistical inference [see for example Balakrishnan and Sandhu [11], Sultan and Moshref [38] and Mahmoud *et al.* [31], among several others].

i) Estimation: The moments of order statistics and record values given in Section 2 can be used to obtain the best linear unbiased estimate of the parameters of the type II exponentiated log-logistic distribution. Some works of this nature based on *gos* have been done by Ahsanullah and habibullah [6], Malinowska et al. [32] and Burkchat *et al.* [16].

ii) Characterization: The type II exponentiated log-logistic distribution given in (1.2) can be characterized by using recurrence of single moment of *gos* as follows:

Let $L(a,\ b)$ stand for the space of all integrable functions on $(a,\ b)$. A sequence $(f_n)\subset L(a,\ b)$ is called complete on $L(a,\ b)$ if for all functions $g\in L(a,\ b)$ the condition

$$\int_a^b g(x)f_n(x)dx = 0, \qquad n\in N,$$

implies $g(x)=0$ a.e. on $(a,\ b)$. We start with the following result of Lin [30].

Proposition 5.1 Let n_0 be any fixed non-negative integer, $-\infty\leq a<b\leq\infty$ and $g(x)\geq 0$ an absolutely continuous function with $g'(x)\neq 0$ a.e. on $(a,\ b)$. Then the sequence of functions $\{(g(x))^n e^{-g(x)},\ \ n\geq n_0\}$ is complete in $L(a,\ b)$ iff $g(x)$ is strictly monotone on $(a,\ b)$.

Using the above Proposition we get a stronger version of Theorem 2.4.

5.1. Theorem. *A necessary and sufficient conditions for a random variable X to be distributed with pdf given by (1.1) is that*

$$\Big(1-\frac{\sigma j}{\alpha\beta\gamma_r}\Big)E[X^j(r,n,m,k)] = E[X^j(r-1,n,m,k)] + \frac{j\sigma^{\beta+1}}{\alpha\beta\gamma_r}E[X^{j-\beta}(r,n,m,k)]. \tag{5.1}$$

Proof The necessary part follows immediately from (2.14) on the other hand if the recurrence relation (5.1) is satisfied then on using (1.5), we have

$$\begin{aligned}\frac{C_{r-1}}{(r-1)!}&\int_0^\infty x^j[\bar F(x)]^{\gamma_r-1}f(x)g_m^{r-1}(F(x))dx\\ &= \frac{C_{r-1}}{\gamma_r(r-2)!}\int_0^\infty x^j[\bar F(x)]^{\gamma_r+m}f(x)g_m^{r-2}(F(x))dx\\ &+\frac{\sigma j C_{r-1}}{\alpha\beta\gamma_r(r-1)!}\int_0^\infty x^j[\bar F(x)]^{\gamma_r-1}f(x)g_m^{r-1}(F(x))dx\\ &+\frac{j\sigma^{\beta+1}C_{r-1}}{\alpha\beta\gamma_r(r-1)!}\int_0^\infty x^{j-\beta}[\bar F(x)]^{\gamma_r-1}f(x)g_m^{r-1}(F(x))dx.\end{aligned} \tag{5.2}$$

Integrating the first integral on the right-hand side of the above equation by parts and simplifying the resulting expression, we get

$$\frac{jC_{r-1}}{\gamma_r(r-1)!}\int_0^\infty x^{j-1}[\bar F(x)]^{\gamma_r-1}g_m^{r-1}(F(x)) \times\Big\{\bar F(x)-\frac{\sigma x}{\alpha\beta}f(x)-\frac{\sigma^{\beta+1}}{\alpha\beta x^{\beta-1}}f(x)\Big\}dx=0.$$

It now follows from Proposition 5.1, we get

$$\alpha\beta\bar F(x) = \sigma[1+(x/\sigma)^\beta]xf(x),$$

which proves that $f(x)$ has the form (1.1).

6. Concluding Remarks

In the study presented above, we established some new explicit expressions and recurrence relations between the single and product moments of *gos* from the type II exponentiated log-logistic distribution. In addition ratio and inverse moments of type II exponentiated log-logistic distribution are also established. Further, the conditional expectation of *gos* is used to characterize the distribution.

Acknowledgements
The author appreciates the comments and remarks of the referees which improved the original form of the paper.

References

[1] Aggarwala, R. and Balakrishnan, N. *Recurrence relations for progressive type II right censored order statistics from exponential truncated exponential distributions*, Ann. Instit. Statist. Math., **48**, 757-771, (1996).
[2] Ahmad, A.A. and Fawzy, M. *Recurrence relations for single moments of generalized order statistics from doubly truncated distribution*, J. Statist. Plann. Inference, **177**, 241-249, (2003).
[3] Ahmad, A.A. *Relations for single and product moments of generalized order statistics from doubly truncated Burr type XII distribution*, J. Egypt. Math. Soc., **15**, 117-128, (2007).
[4] Ahsanullah, M. *Record Statistics*, Nova Science Publishers, New York, (1995).

[5] Ahsanullah, M. *Generalized order statistics from exponential distribution*, J. Statist. Plann. Inference, **85**, 85-91, (2000).

[6] Ahsanullah, M. and Habibullah, M. *Estimation of parameters of a Pareto distribution by generalized order statistics*, Comm. Statist. Theory Methods, **29**, 1597-1609, (2000).

[7] Arnold, B.C. and Balakrishnan, N. *Relations, Bounds and Approximations for Order Statistics. Lecture Notes in Statistics*, Springer-Verlag, Berlin, **53**, (1989).

[8] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. *A First Course in Order Statistics*, John Wiley and Sons, New York (1992).

[9] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. *Records*, . John Wiley, New York, (1998).

[10] Balakrishnan, N. and Aggarwala, R. *Progressive censoring: Theory Methods and Applications*, Birkhauser Berlin, (2000).

[11] Balakrishnan, N. and Sandhu, R. A. *Best linear unbiased and maximum likelihood estimation for exponential distribution under general progressive type II censored samples*, Sankhya Ser. B, **58**, 1-9, (1996).

[12] Balakrishnan, N., Chan, P. S. and Ahsanullah, M. *Recurrence relations for moments of record values from generalized extreme value distribution*, Comm. Statis. Theory and Methods, **22**, 1471-1482, (1993).

[13] Balakrishnan, N., Cramer, E. and Kamps, U. *Bounds for means and variances of progressive type II censored order statistics*, Statist. Prob. Letters, **54**, 301-315, (2001).

[14] Balasooriya, U. and Saw, S. *Reliability sampling plans for the two parameter exponential distribution under progressive censoring*, J. Appl. Statist., **25**, 707-714, (1998).

[15] Bieniek, M. and Szynal, D. *Characterizations of distributions via linearity of regression of generalized order statistics*, Metrika, **58**, 259-271, (2003).

[16] Burkschat, M. Cramer, E. and Kamps, U. *Linear estimation of location and scale parameters based on generalized order statistics from generalized Pareto distribution*, In Recent Development in Ordered random Variables, 253-261, (2007).

[17] Cramer, E., Kamps, U. and Keseling, C. *Characterization via linear regression of ordered random variables: a unifying approach*, Comm. Statist. Theory Methods, **33**, 2885-2911, (2004).

[18] Cramer, E. *Contributions to generalized order statistics*, Ph. D. Thesis. University of Oldenburg, (2002).

[19] Cramer, E. and Kamps, U. *Sequential order statistics and k- out-of n systems with sequentially adjusted failure rates*, Ann. Instit. Statist. Math, **48**, 535-549, (1996).

[20] Cramer, E. and Kamps, U. *Relations for expectations of functions of generalized order statistics*, J. Statist. Plann. Inference, **89**, 79-89, (2000).

[21] Glick, N. *Breaking records and breaking boards*, Amer. Math. Monthly , **85**, 2-26, (1978).

[22] Kamps, U. and Cramer, E. *On distributions of generalized order statistics*, Statistics, **35**, 269-280, (2001).

[23] Kamps, U. and Gather, U. *Characteristic property of generalized order statistics for exponential distribution*, Appl. Math. (Warsaw), **24**, 383-391, (1997).

[24] Kamps, U. *A Concept of Generalized Order Statistics*, B.G. Teubner Stuttgart, (1995).

[25] Keseling, C. *Conditional distributions of generalized order statistics and some characterizations*, Metrika, **49**, 27-40, (1999).

[26] Khan, A.H., Khan, R.U. and Yaqub, M. *Characterization of continuous distributions through conditional expectation of functions of generalized order statistics*, J. Appl. Probab. Statist., **1**, 115-131, (2006).

[27] Kumar, D. *Recurrence relations for single and product moments of generalized order statistics from $p-th$ order exponential distribution and its characterization*, Journal of Statistical Research of Iran, **7**, 101-112, (2010).

[28] Kumar, D. *Generalized order statistics from Kumaraswamy distribution and its characterization*, Tamsui Oxford journal of Mathematical Sciences, **27**, 463-476, (2011).

[29] Kumar, D. *On moments of lower generalized order statistics from exponentiated lomax distribution and characterization*, American Journal of Mathematical and Management Sciences, **32**, 238-256, (2013).

[30] Lin, G.D. *On a moment problem*, Tohoku Math. Journal, **38**, 595-598, (1986).

[31] Mahmoud, M. A. W., Sultan, K. S. and Amer, A. M. *Order statistics from inverse Weibull distribution and associated inference*, Computational Statistics and data Analysis, **42**, 149-163, (2003).

[32] Malinowska, I., Pawlas, P. and Szynal, D. *Estimation of location and scale paremeters for the Burr XII distribution using generalized order statistics.* , Linear Algebra and its Applications, **417**, 150-162, (2006).

[33] Nevzorov, V.B. *Records*, Theory probab. Appl, **32**, (1987).

[34] Pawlas, P. and Szynal, D. *Recurrence relations for single and product moments of generalized order statistics from Pareto, generalized Pareto, and Burr distributions*, Comm. Statist. Theory Methods, **30**, 739-746, (2001).

[35] Resnick, S. I. *Record values and related statistics*, Ann. Probab, **2**, 650- 662, (1973).

[36] Ruiz, S. M. *An algebraic identity leading to Wilson's theorem*, Math. Gaz., **80**, 579-582, (1996).

[37] Schenk, N. *Point estimation with sequential order statistics from exponential distributions*, PhD. Thesis, Oldenburg University, Germany, (2001).

[38] Sultan, K. S. and Moshref, M. E. *Record value from generalized Pareto distribution and associated inference*, Metrika, **51**, 105-116, (2000).

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 735 – 746

An improved estimator of the distortion risk measure for heavy-tailed claims

Rassoul Abdelaziz*

Abstract

The main aim of this paper is to propose an alternative estimate of the distortion risk measure for heavy-tailed claims. Our approach is based on the result of Balkema and de Haan (1974) [3], and Pickands (1975) [22] for approximating the tail of the distribution by a generalized Pareto distribution. The asymptotic normality of the new estimator is established, and its performance illustrated by some results of simulation who shows the advantages of the new estimator over the estimator based on the classical extreme-value theory.

Keywords: Premium principle, Distortion risk measure, POT method, Extremes values theory, Generalized Pareto distribution, Loss distribution.

Received 11/02/2014 : Accepted 18/06/2014 Doi : 10.15672/HJMS.2014207471

1. Introduction

A number of risks measures found in finance and insurance literature are special cases of the distortion risk measure, defined by

$$H\left[F,g\right]=\int_0^{+\infty} g\left(\overline{F}(x)\right)dx. \tag{1.1}$$

where $X \geq 0$ is a loss random variable with cumulative distribution function (cdf) F and the de-cumulative distribution function (ddf) $\overline{F} = 1 - F$, which is also known as survival function. The distortion function $g : [0,1] \to [0,1]$ is assumed to be an increasing function such that $g(0) = 0$ and $g(1) = 1$.
Dhaene et al. (2012) [9] show that, when the distortion function g is right continuous on $[0,1)$, the formula (1.1) may be rewritten as follows

$$H\left[F,g\right]=\int_0^{1} \mathbb{Q}(1-s)dg\left(s\right), \tag{1.2}$$

*Laboratoire de Génie de l'eau et Environement (GEE)
Ecole Nationale Supérieure d'Hydraulique, Blida, Algérie,
Email: a.rassoul@ensh.dz

where $\mathbb{Q}$ is the quantile function corresponding the cdf F, that is

$$\mathbb{Q}(t) = \inf\{x : F(x) \geq t\} = F^{-1}(t), \text{ for } t \in]0,1[\,.$$

The risk measure $H[F,g]$, which can also be viewed as a premium calculation principle, has manifested in the econometric literature, particularly in Yaari's (1987) [31] dual theory of choice under risk, and has been introduced into actuarial literature by Wang (1996) [28]. A number of risk measures of this form have been discussed by Wirch and Hardy (1999) [30].
In Artzner (1999) [1] and Artzner et al. (1999) [2] a risk measure satisfying the four axioms of subadditivity, monotonicity, positive homogeneity and translation invariance is called Coherent, and also demonstrated that the risk measure $H[F,g]$ is coherent when g is concave. Note that the class of concave distortion risk measures is only a subset of the class of coherent risk measures.
Many special cases that have arisen in the finance and insurance literature are such:

- VaR: $g(x) = 1_{[1-q,1]}$ for some $q \in]0,1[$
- Tail-VaR: $g(x) = \min\{\frac{x}{1-q}, 1\}$ for some $q \in (0,1)$
- Proportional Hazard Transform: $g(x) = x^{1/\rho}$ for some $\rho > 1$
- Dual-Power Transform: $g(x) = 1 - (1-x)^{\rho}$ for some $\rho > 1$
- Gini principle: $g(x) = (1+\rho)x - \rho x^2$, with $0 < \rho \leq 1$.
- Lookback distortion: $g(x) = x^{\rho}(1 - \rho\ln(x))$, with $0 < \rho \leq 1$.

Detailed studies of distortion risk measures, also known as Wang's risk measures, can be found in, for example, Wang (1996) [28], Wang and Young (1998) [29], Hürlimann (1998) [12], and Hua and Joe, (2012) [13].

A number of authors have tackled the distortion risk measure from the statistical inferential point of view. A short survey and classification of papers in the area follows:

- Light-tailed distributions
  - Classical-type asymptotic results
  - Asymptotic results aimed at variance reduction
- Heavy-tailed distributions
  - Fisher-Tippett-Gnedenko type extreme-value methods
  - Pickands-Balkema-de Haan type Peack Over Threshold methods

Jones and Zitikis (2003) [16] noticed that the empirical counterpart of $H[F,g]$ is a linear combination of order statistics, commonly known as L-statistic. This opens up a fruitful venue for developing statistical inferential results, which have been actively investigated by a number of researchers. Specifically, let $X_1, ..., X_n$ be independent copies of X; and let $X_{1,n}, ..., X_{n,n}$ be the corresponding ascending order statistics. The empirical estimator of the risk premium $H[F,g]$ is obtained by substituting the quantile $\mathbb{Q}$ on the right-hand side of equation (1.2) by its empirical counterpart

$$\mathbb{Q}_n(s) := inf\{x : F_n(x) \geq s\} := F_n^{-1}(s)\,,$$

on the real line, defined by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leq x\}},$$

with $1_{\{.\}}$ being the indicator function. After straightforward computation, we obtain the formula

$$\widehat{H}_n[F_n, g] = \int_0^1 \widehat{Q}_n(1-s)dg(s)\,,$$

where $\widehat{Q}_n(1-s)$ is an empirical estimator of the quantile function, given by the formula

$$\widehat{Q}_n(1-s) := X_{n-k+1,n}, \text{ where } \frac{k-1}{n} < s \leq \frac{k}{n},$$

Then, the empirical estimator of $H[F,g]$ is given by the formula

$$\widehat{H}_n[F,g] = \sum_{i=1}^{n} \left(g\left(\frac{i}{n}\right) - g\left(\frac{i-1}{n}\right)\right) X_{n-i+1,n}.$$

For recent literature on statistical inference for distortion premiums, we refer to Jones and Zitikis (2003) [16], Jones and Zitikis (2007) [17], Centeno and Andrade (2005) [8], Furman and Zitikis (2008) [10], Brazauskas et al. (2008) [6], Greselin et al. (2009) [11], Necir et al. (2010) [20], Joseph H. T. Kim. (2010) [18], Peng et al. (2012) [21] and the references therein.
The asymptotic normality of the estimator $\widehat{H}_n[F,g]$ is established by Jones and Zitikis (2003) [16] as follows

$$\sqrt{n}\left(\widehat{H}_n[F_n,g] - H[F,g]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right),$$

in particular, if g is differentiable, we have

$$\sigma^2 := \int_0^1\int_0^1 \left(\min\{F(t),F(s)\} - F(t)F(s)\right) g\left(1-F(t)\right) g\left(1-F(s)\right) dtds,$$

by provided that the second moment are finite, that is $E(X^2) < \infty$. This is a very restrictive condition in the context of heavy-tailed distributions as the following considerations show. Assume that the rv X_1 follows the Fréchet law with index $\gamma > 0$, that is, $1 - F(x) = exp\left\{-x^{-1/\gamma}\right\}$ for $x > 1$. When $\gamma \in (0.5, 1]$, the mean exist, but the second moment $E\left(X^2\right)$ is infinite. Hence, the range is not covered by the CLT and thus, another approach to handle this situation is needed. Making use of the results of Balkema and de Haan (1974) [3], and Pickands (1975) [22] to approximate the tail of the distribution by the Generalized Pareto Distribution (GPD), this result is know by the Peack Over Threshold method (POT) to propose a alternative estimator for the distortion risk premiums. Moreover, under suitable assumptions we established its asymptotic normality, and we presente some results of simulation to illustrate the performance of our estimator applying to the proportional hazard premium PHP. Empirical studies have shown that Financial and actuarial data exhibit heavy tails or Pareto like distributions. The class of regularly varying cdf's is a major subclass of heavy-tailed distributions, it includes distributions such as Pareto, Burr, Student, Lévy-stable, and loggamma, which are known to be appropriate models for fitting large insurance claims, large fluctuations of prices, log-returns, etc. (see, e.g., Beirlant et al., 2001 [4]; Reiss and Thomas, 2007 [23] and Rolski et al., 1999 [25]).

Note that throughout this paper, the standard notations $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{D}}$ and $\overset{d}{=}$ respectively stand for convergence in probability, convergence in distribution and equality in distribution, $\mathcal{N}(a,b^2)$ denotes the normal distribution with mean a and variance b^2, and $N_2(\mu,\Sigma)$ denote the bivariate normal distribution with mean vector μ and matrix of variance-covariance Σ.

The paper is organized as follows. In section 2, we introduice the differents notions and definitions of the used tools and the mains assumptions. In sections 3 we introduice the new estimator of $H_{g,n}$, and presente the main result about the limiting behavior of the proposed estimator. Some results of simulation and illustration are given in section 4. The Proofs of the mains results are postponed until section 5.

2. Main assumptions, notations, and the POT method

Distortion functions. We assume that the distortion function g is regularly varying at infinity, with index of regular variation $r \in [0,1]$, that is,

$$g(x) = x^r \ell(x), \tag{2.1}$$

where ℓ is a slowly varying function, that is, $\ell(tx)/\ell(x) \to 1$ when $x \to \infty$ for any $t > 0$. For further properties of these functions, we refer to, for example, Resnick (1987) [24], Seneta (1976) [26]. Examples of such distortion functions are:

- VaR: $r = 0$ and $\ell(x) = 1_{[1-q,1]}(x)$
- Tail-VaR: $r = 1$ and $\ell(x) = 1/(1-q)$
- Proportional Hazard Transform: $r = 1/\rho$ and $\ell(x) = 1$
- Dual-Power Transform: $r = 1$ and $\ell(x) = \rho - \frac{\rho(\rho-1)}{2}x + o(x)$
- Gini Principle: $r = 1$ and $\ell(x) = 1 + \rho - \rho x$.
- Lookback distortion: $r = \rho$ and $\ell(x) = (1 - \rho \ln(x))$.

Distribution functions. We deal only with losses X that are heavy tailed. More specifically, we work within the class of of regularly varying cdf's. Namely, the survival function or the tail of cdf F is said to be with regulary varying at infinity, that is

$$\overline{F}(x) = cx^{-1/\xi}\left(1 + x^{-\delta}\mathbb{L}(x)\right) \text{ when } x \to \infty, \tag{2.2}$$

for $\xi \in (0,1)$, $\delta > 0$ and some real constant c, where $\mathbb{L}$ a slowly varying function.

The POT method. Let $X_1, ..., X_n$ be independent and identically distributed random variables, each with the same cdf F, and let u_n be some a large number, 'high level,' which we later let tend to infinity when $n \to \infty$. With the notation

$$\overline{F}_{u_n}(y) = \mathbf{P}[X_1 - u_n > y \mid X_1 > u_n],$$

we have that

$$\overline{F}_{u_n}(y) = \frac{\overline{F}(u_n + y)}{\overline{F}(u_n)},$$

and thus

$$\overline{F}_{u_n}(y) = \left(1 + \frac{y}{u_n}\right)^{-1/\xi}\left[\frac{1 + (u_n + y)^{-\delta}\mathbb{L}(u_n + y)}{1 + u_n^{-\delta}\mathbb{L}(u_n)}\right]. \tag{2.3}$$

Upon recalling the definition of the generalised Pareto distribution, we have that, for all parameter values $\beta > 0$ and $\xi > 0$,

$$\mathbb{G}_{\xi,\beta}(y) = 1 - \left(1 + \xi\frac{y}{\beta}\right)^{-1/\xi}, \quad 0 \le y < \infty. \tag{2.4}$$

We see that, the right-hand side of equation (2.3) is a perturbed version of $\mathbb{G}_{\xi,\beta_n}(y)$, with the notation $\beta_n = u_n\xi$. Balkema and de Haan (1974) [3], and Pickands (1975) [22] have shown that F_{u_n} is approximated by a generalized Pareto distribution GPD function $\mathbb{G}_{\xi,\beta_n}$ with shape parameter $\xi \in \mathbb{R}$ and scale parameter $\beta = \beta(u_n)$, in the following sense:

$$\sup_{y>0}|F_{u_n}(y) - \mathbb{G}_{\xi,\beta}(y)| = O(u_n^{-\delta}\mathbb{L}(u_n)), \tag{2.5}$$

where, for any $\delta > 0$, we have $u_n^{-\delta}\mathbb{L}(u_n) \to 0$ when $u_n \to \infty$.

Approximation (2.5) suggests to define an estimator of $\overline{F}_{u_n}(y)$ as follows:

$$\widehat{\overline{F}}_{u_n}(y) = \overline{\mathbb{G}}_{\widehat{\xi}_n,\widehat{\beta}_n}(y), \tag{2.6}$$

for appropriate estimates $\widehat{\xi}_n$ and $\widehat{\beta}_n$ of ξ and β, respectively. Note that β will be estimated separately, i.e. $\beta = \xi u_n$ will not be used. The reason for this is to achieve greater flexibility in the parameter fitting, compensating for the underlying distribution not being an exact GPD. Theorem 3.2 in Smith (1987) [27] gives us the asymptotic distribution of the tail parameters $\left(\widehat{\xi}_n, \widehat{\beta}_n\right)$ as follows

$$\text{(2.7)} \qquad \sqrt{np_n}\begin{pmatrix} \widehat{\beta}_n/\beta - 1 \\ \widehat{\xi}_n - \xi \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_2\left(0, \mathbf{\Sigma}^{-1}\right) \text{ when } n \to \infty,$$

provided that $\sqrt{np_n}u_n^{-\delta}\mathbb{L}(u_n) \to 0$ when $n \to \infty$ and the function $x \mapsto x^{-\delta}\mathbb{L}(x)$ is non-increasing for all sufficiently large x, where

$$\text{(2.8)} \qquad \mathbf{\Sigma}^{-1} = (1+\xi)\begin{pmatrix} 2 & -1 \\ -1 & 1+\xi \end{pmatrix}.$$

We note that when $\sqrt{np_n}u_n^{-\delta}\mathbb{L}(u_n) \nrightarrow 0$, then the limiting distribution in (2.7) is biased.

Next we define an estimator of $\overline{F}(u_n)$. For this, let $N \equiv N_n(u_n)$ be defined by

$$N = \#\left\{X_i : \ X_i > u_n : 1 \le i \le n \right\},$$

which is the number of those X_i's that exceed u_n. Since N follows the binomial distribution $\mathcal{B}(p_n, n)$ with the parameter $p_n = \mathbf{P}[X_1 > u_n]$, which is equal to $\overline{F}(u_n)$, we have a natural estimator of $\overline{F}(u_n)$ defined by

$$\widehat{p}_n = \frac{N}{n}.$$

From the definition of $\overline{F}_{u_n}(y)$ we have $\overline{F}(u_n+y) = \overline{F}(u_n)\overline{F}_{u_n}(y)$. Hence, with the above defined estimators for $\overline{F}_{u_n}(y)$ and $\overline{F}(u_n)$, we have the following estimator of $\overline{F}(u_n+y)$:

$$\begin{aligned} \widehat{\overline{F}}(u_n+y) &= \widehat{\overline{F}}(u_n)\widehat{\overline{F}}_{u_n}(y) \\ \text{(2.9)} \qquad &= \widehat{p}_n\overline{\mathbb{G}}_{\widehat{\xi}_n,\widehat{\beta}_n}(y). \end{aligned}$$

We shall use $\widehat{\overline{F}}(u_n+y)$ to construct an estimator for the distortion risk measure $H\left[F,g\right]$ and then show in a simulation study that in this way constructed empirical distortion risk measure outperforms the one constructed using Fisher-Tippett-Gnedenko type extreme-value methods.

3. The new estimator and the main result

We start constructing a POT-based estimator of $H\left[F,g\right]$ using the following lemma.

3.1. Lemma. *Assume that F and g satisfying (2.2) and (2.1) respectively, and u_n be some large level. Then, when $n \to \infty$, we have that*

$$\text{(3.1)} \qquad H_n[F,g] = \int_0^{u_n} g\left(\overline{F}(x)\right)dx + (p_n)^r\,\frac{\beta}{r-\xi} + r_n$$

with the remainder term

$$r_n = O\big(u_n^{1-r/\xi-\delta}\big),$$

which converges to 0 *when* $n \to \infty$ *because* $1 - r/\xi - \delta < 0$.

The proof of the lemma 3.1 is relegated to Section 5. With p_n, β and ξ on the right-hand side of equation (3.1) replaced by their estimators, we obtain an estimator of $H\left[F,g\right]$, defined as follows:

$$\text{(3.2)} \qquad \widehat{H}_n\left[F,g\right] = \int_0^{u_n} g\left(\overline{F}_n(x)\right)dx + \left(\widehat{p}_n\right)^r\,\frac{\widehat{\beta}_n}{r-\widehat{\xi}_n}.$$

The asymptotic normality of $\widehat{H}_n[F,g]$ is established in the following theorem.

3.2. Theorem. *Let F be a distribution function fulfilling (2.2) with $\xi\in(0.5,1)$ and the distortion function g is differentiable and regularly varying at infinity with index $0\leq r\leq 1$. Suppose that $\mathbb{L}$ is locally bounded in $[x_0,+\infty)$ for $x_0\geq 0$ and $x\to x^{-\delta}\mathbb{L}(x)$ is non-increasing near infinity, for some $\delta>0$. For any $u_n=O(n^{\alpha\xi})$ with $\alpha\in(0,1)$, we have*

$$\frac{\sqrt{n}}{\gamma_n\sigma_n}\left(\widehat{H}_n[F,g]-H[F,g]\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\,,\ as\ n\to\infty,$$

where

$$\sigma_n^2:=1+\frac{\theta_1^2}{\gamma_n^2}p_n(1-p_n)+\frac{2(1+\xi)\theta_2^2}{p_n\gamma_n^2}+\frac{(1+\xi)^2\theta_3^2}{p_n\gamma_n^2}-\frac{(1+\xi)\theta_2\theta_3}{p_n\gamma_n^2}.$$

and

$$\gamma_n^2=\mathbf{Var}\left[\int_0^{u_n}g'\left(\overline{F}(x)\right)\mathbf{1}(X\leq x)\,dx\right],$$

with

$$\theta_1=\frac{\beta g'(p_n)}{r-\xi},\theta_2=\frac{\beta g(p_n)}{r-\xi},\theta_3=\frac{\beta g(p_n)}{(r-\xi)^2},$$

and $\beta=u_n\xi$.

4. Simulation Study

To illustrate the result of the Theorem 3.2 , we carry out a simulation study (by means of the statistical software **R**, see Ihaka and Gentleman, 1996) [14], in this study we are interesting by a popular risks measure named Proportional Hazard Premium (PHP) where the distortion function is given by $g(x)=x^{1/\rho}$ with $\rho>1$, to illustrate the performance of our estimation and its comparison with the parametric estimator, through its application to sets of samples taken from two distinct Pareto distributions $\overline{F}(x)=x^{-1/\xi},x\geq 1$ (with tail index $\xi=2/3$ and $\xi=3/4$), we are interesting by the PHP risk measure, that is, the distortion function is given by $g(x)=x^{1/\rho}$ with the distortion paramater $\rho>1$, in this case the esimator of the PHP is given by

$$\widehat{H}_{\rho,n}=\int_0^{u_n}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{(X_j\geq x)}\right)^{1/\rho}dx+(\widehat{p}_n)^{1/\rho}\,\frac{\rho\widehat{\beta}_n}{1-\rho\widehat{\xi}_n}.$$

In the first part, we evaluate the root mean squared error (rmse), the accuracy of the confidence intervals via and their lengths (length) and the coverage probabilities (cprob), the confidence level $1-\zeta$ is fixed at 0.95, we generate 200 independent replicates of sizes 500, 1000 and 2000 from the selected parent distribution for $\xi=2/3$. For each simulated sample, we obtain an estimate of the estimators premium H_ρ for two distinct aversion index values $\rho=1.1$ and $\rho=1.2$. In each case we compute, by averaging over all samples, the confidence bounds and the coverage probability and length of the corresponding confidence interval. Note that lcb and ucb stand respectively for lower confidence bound and upper confidence bound.
To this end. We summarize the results in Table 1 for $\xi=2/3$, $\rho=1.1$, and Table 2 for $\xi=2/3$,$\rho=1.2$.

In this second part, we generate 200 independent replicate of size 1000 from the selected parent distribution $\overline{F}(x)=x^{-1/\xi},x\geq 1$ (with tail index $\xi=2/3$ and $\xi=3/4$) and estimate the PHP for two distinct aversion index values $\rho=1.1$ and $\rho=1.2$. We interesting by the comparison of our estimator $\widehat{H}_{\rho,n}$ with the old estimator constructed

Table 1. Point estimates and 95%-confidence intervals for H,based on200 samples of Pareto-distributed rv's with tail index $\xi = 2/3$ and $\rho = 1.1$.

$\rho = 1.1$			$H = 3.75$			
n	$\widehat{H}_{\rho,n}$	rmse	lcb	ucb	cprob	length
500	3.312	0.561	2.23	4.39	0.54	2.168
1000	4.037	0.286	3.139	4.934	0.71	1.793
2000	3.765	0.050	3.189	4.342	0.82	1.153

Table 2. Point estimates and 95%-confidence intervals for H,based on 200 samples of Pareto-distributed rv's with tail index $\xi = 2/3$ and $\rho = 1.2$.

$\rho = 1.2$			$H = 5$			
n	$\widehat{H}_{\rho,n}$	rmse	lcb	ucb	cprob	length
500	5.194	0.835	2.852	7.537	0.640	4.683
1000	5.069	0.355	3.444	6.696	0.815	3.252
2000	5.028	0.311	3.617	6.439	0.890	2.822

by the extreme values methods by (Necir and Meraghni 2009 [19]) and noted $\widetilde{H}_{\rho,n}$, this comparaison is in terms the **bias** and the mean squared error **(MSE).** We summarize the results in Table (3)

Table 3. Analog between the new estimator and the old estimator of the premium hazard proportional for two tail index and two risk aversions index

ξ	2/3		3/4	
ρ	1.1	1.2	1.1	1.2
H_ρ	3.75	5	5.714	10
$\widehat{H}_{\rho,n}$	3.752	5.071	5.815	10.036
bias	0.002	0.071	0.101	0.037
MSE	0.0998	0.256	0.340	1.796
$\widetilde{H}_{\rho,n}$	4.042	5.280	6.050	8.718
bias	0.292	0.280	0.336	-1.283
MSE	0.116	0.299	0.457	2.048

From these results, we observe that the new estimator has smaller bias and mean squared error than the old estimator in most cases, the new estimator performs worse, which may be explained by the Theorem 3.2.

5. Proofs

The following propositions are instrumental for the proof of Theorem 3.2.

5.1. Proposition. *Let F be a distribution function fulfilling (2.2) with $\xi \in (0,1)$, $\delta > 0, r \in [0,1]$ and some real c. Suppose that $\mathbb{L}$ is locally bounded in $[x_0, +\infty)$ for $x_0 \geq 0$.*

Then for n large enough, for any $u_n = O\left(n^{\alpha\xi}\right)$, $\alpha \in (0,1)$, we have that

$$p_n = cn^{-\alpha}(1+o(1)), \tag{5.1}$$

$$\gamma_n^2 = O\left(n^{2\alpha(\xi-r+1)}\right), \tag{5.2}$$

and

$$\sqrt{np_n}u_n^{-\delta}\mathbb{L}\left(u_n\right) = O\left(n^{-\alpha/2-\alpha\xi\delta+1/2}\right). \tag{5.3}$$

Proof of the proposition 5.1. We will now prove the result (5.1), let $\overline{F}\left(x\right) = cx^{-1/\xi}\left(1+x^{-\delta}\mathbb{L}\left(x\right)\right)$. Then for n large enough, we have

$$\begin{aligned} p_n &= \mathbf{P}\left(X > u_n\right) = \overline{F}\left(u_n\right) \\ &= cu_n^{-1/\xi}\left(1+u_n^{-\delta}\mathbb{L}\left(u_n\right)\right), \end{aligned}$$

with $u_n = O\left(n^{\alpha\xi}\right)$, then we obtain the statement (5.1). The result (5.3) are straightforward from the result (5.1). We shall next prove statement (5.2). Note that the quantity γ_n^2 defined the formulation of the theorem is equal to $\mathbf{Var}[Z]$, where

$$Z = \int_0^{u_n} g'\left(\overline{F}(x)\right)\mathbf{1}\left(X \le x\right)dx.$$

Since $\overline{F}(x) = x^{-1/\xi}O(1)$, $g(x) = x^r O(1)$ and $u_n = n^{\alpha\xi}O(1)$, we have that

$$\begin{aligned} \mathbf{E}\left[Z\right] &= \int_0^{u_n} g'\left(\overline{F}(x)\right)F(x)dx \\ &= \int_0^{u_n} g'\left(\overline{F}(x)\right)dx - \int_0^{u_n} g'\left(\overline{F}(x)\right)\overline{F}(x)dx \\ &= n^{\alpha(1+\xi-r)}O(1). \end{aligned}$$

Furthermore,

$$\begin{aligned} \mathbf{E}\left[Z^2\right] &= \int_0^{u_n}\int_0^{u_n} g'\left(\overline{F}(x)\right)g'\left(\overline{F}\left(y\right)\right)\min\left(F(x),F\left(y\right)\right)dxdy \\ &= \int_0^{u_n} g'\left(\overline{F}(x)\right)\left(\int_0^{x} g'\left(\overline{F}\left(y\right)\right)F\left(y\right)dy\right)dx \\ &\quad + \int_0^{u_n} g'\left(\overline{F}(x)\right)\left(\int_x^{u_n} g'\left(\overline{F}\left(y\right)\right)F\left(y\right)dy\right)dx \\ &= 2n^{2\alpha(\xi-r+1)}O(1). \end{aligned}$$

Consequently, statement (5.2) holds. □

Proof of Lemma 3.1. We start with the elementary equation

$$H_{g,n} = \int_0^{u_n} g\left(\overline{F}(x)\right)dx + \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx.$$

Hence, the remainder term r_n noted in the formulation of the Lemma 3.1 is

$$r_n = \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx - \left(p_n\right)^r\frac{\beta}{r-\xi}.$$

Next we express the integral in the definition of r_n as follows:

$$\begin{aligned} \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx &= \int_0^{\infty} g\left(\overline{F}\left(s+u_n\right)\right)ds \\ &= \int_0^{\infty} g\left(p_n\overline{F}_{u_n}(s)\right)ds. \end{aligned}$$

Since $\overline{F}_{u_n}(s) = \overline{F}(u_n + s) / \overline{F}(u_n)$, we have that

$$\overline{F}_{u_n}(s) = \left(1 + \frac{\xi}{\beta} s\right)^{-1/\xi} \frac{1 + (u_n + s)^{-\delta} \mathbb{L}(u_n + s)}{1 + u_n^{-\delta} \mathbb{L}(u_n)}.$$

Consequently,

$$\int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx = (p_n)^r \frac{\beta}{r - \xi} \left(\frac{1 + (u_n + s)^{-\delta} \mathbb{L}(u_n + s)}{1 + u_n^{-\delta} \mathbb{L}(u_n)}\right)^r.$$

Since function $\mathbb{L}$ is locally bounded in $[x_0, \infty)$ for $x_0 \geq 0$ and $x^{-\delta}\mathbb{L}(x)$ is non-increasing near infinity, then for all large n, we have that

$$u_n^{r/\xi} \int_{u_n}^{\infty} x^{-r/\xi - \delta} \mathbb{L}(x) dx = O(u_n^{-\delta}).$$

Consequently, for all large n,

$$\int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx = \frac{\beta}{\xi} \int_{1}^{\infty} g\left(p_n (z)^{-1/\xi}\right) dz \left(1 - u_n^{-\delta} \mathbb{L}(u_n) + O(u_n^{-\delta} \mathbb{L}(u_n)\right).$$

This implies that $r_n = O(u_n^{1 - r/\xi - \delta})$ and concludes the proof of Lemma 3.1. □

Proof of Theorem 3.2. We write

$$\sqrt{n}\big(\widehat{H}_{g,n} - H_g\big) = A_n + B_n,$$

where

$$A_n = \sqrt{n} \int_{0}^{u_n} \left(g\left(\overline{F}_n(x)\right) - g\left(\overline{F}(x)\right)\right) dx$$

and

$$B_n = \sqrt{n} \left((\widehat{p}_n)^r \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n} - \int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx\right).$$

Using Lemma 3.1 and the fact that $\sqrt{n}\ u_n^{1 - r/\xi - \delta} \to 0$, as $n \to \infty$, we have that

$$\begin{aligned} B_n &= \sqrt{n} \left((\widehat{p}_n)^r \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n} - \int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx\right) \\ &= \sqrt{n}\left(B_{n,1} + O\big(u_n^{1 - r/\xi - \delta}\big)\right) \\ &= \sqrt{n} B_{n,1} + o(1), \end{aligned}$$

where

$$\begin{aligned} B_{n,1} &= (\widehat{p}_n)^r \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n} - (p_n)^r \frac{\beta}{r - \xi} \\ &= \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n} (\widehat{p}_n^r - p_n^r) + \frac{(p_n)^r \beta}{(r - \widehat{\xi}_n)} (\widehat{\beta}_n / \beta - 1) + \frac{(p_n)^r}{(r - \widehat{\xi}_n)(r - \xi)} (\widehat{\xi}_n - \xi). \end{aligned}$$

By Smith (1987) [27], we have that

(5.4) $$\widehat{\beta}_n / \beta - 1 = O_{\mathbf{P}}\left(u_n^{-\delta} \mathbb{L}(u_n)\right)$$

and

(5.5) $$\widehat{\xi}_n - \xi = O_{\mathbf{P}}\left(u_n^{-\delta} \mathbb{L}(u_n)\right).$$

Furthermore, by the CLT, we have that

(5.6) $$\widehat{p}_n - p_n = O_{\mathbf{P}}(\sqrt{p_n / n}).$$

Consequently, we have that

$$B_{n,1} = \theta_1 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{p}_n - p_n) + \theta_2 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{\beta}_n/\beta - 1) \\ + \theta_3 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{\xi}_n - \xi),$$

where

$$\theta_1 = \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n}, \quad \theta_2 = \frac{(p_n)^r \beta}{(r - \widehat{\xi}_n)}, \quad \theta_3 = \frac{(p_n)^r}{(r - \xi)^2}.$$

We now examine A_n, and start with the equations

$$A_n = \frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(g\left(\overline{F}_n(x)\right) - g\left(\overline{F}(x)\right)\right) dx$$

$$= \frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(\overline{F}_n(x) - \overline{F}(x)\right) g'\left(\overline{F}(x)\right) dx + o_{\mathbf{P}}(1). \tag{5.7}$$

Continuing with (5.7), we have that

$$\begin{aligned} A_n &= -\frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(F_n(x) - F(x)\right) g'\left(\overline{F}(x)\right) dx + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(\frac{1}{n} \sum \mathbf{1}\left(X_i \le x\right) - F(x)\right) g'\left(\overline{F}\left(x\right)\right) dx + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \left(\frac{1}{n} \sum \int_0^{u_n} \mathbf{1}\left(X_i \le x\right) g'\left(\overline{F}(x)\right) dx - \int_0^{u_n} F(x) g'\left(\overline{F}(x)\right) dx\right) + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) + o_{\mathbf{P}}(1), \end{aligned}$$

where $\overline{Z}$ is the arithmetic average of the n random variables

$$Z_i := \int_0^{u_n} g'\left(\overline{F}(x)\right) \mathbf{1}\left(X_i \le x\right) dx.$$

Note that the quantity γ_n^2 defined in the formulation of the Theorem 3.2 is equal to $\mathbf{Var}[Z_1]$.

Next, we shall show that

$$\frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

when $n \to \infty$. We shall next employ the Lindeberg-Feller Theorem. For this, we write:

$$\begin{aligned} \frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) &= \sum_{k=1}^{n} \frac{\int_0^{u_n} g'\left(\overline{F}(x)\right) \mathbf{1}\left(X_k \le x\right) dx - \mathbf{E}\left[Z_1\right]}{\gamma_n \sqrt{n}} \\ &\equiv \sum_{k=1}^{n} \xi_{k,n}, \end{aligned}$$

where $\mathbf{E}\left(\xi_{k,n}\right) = 0$, $\mathbf{E}\left(\xi_{k,n}^2\right) = 1/n$, and $\sum_{k=1}^n \mathbf{E}\left(\xi_{k,n}^2\right) = 1$ for all $n \ge 1$. Furthermore, for all $\alpha \in (0,1), \xi \in (0,1)$ and $\epsilon > 0$, where $u_n = O\left(n^{\alpha\xi}\right)$ was used. This means that

$$\begin{aligned} \sum_{k=1}^{n} \mathbf{E}\left[\left|\xi_{k,n}\right|^2 1 \left|\xi_{k,n}\right| > \epsilon\right] &= \frac{1}{\gamma_n^2} \mathbf{E}\left[\left[Z_k - \mathbf{E}\left[Z_1\right]\right]^2 1 \left|Z_k - \mathbf{E}\left[Z_1\right]\right| > \epsilon \gamma_n \sqrt{n}\right] \\ &\le \frac{u_n^2}{\gamma_n^2} \mathbf{P}\left[\left|Z_k - \mathbf{E}\left[Z_1\right]\right| > \epsilon \gamma_n \sqrt{n}\right] \\ &\le \frac{u_n^2}{\gamma_n^4 \epsilon^2 n}. \end{aligned}$$

We have, from (5.2) with $u_n = n^{\alpha\xi}$, that

$$\frac{u_n^2}{\gamma_n^4 n} = n^{\alpha(4(r-1)-2\xi)-1} O(1)$$

As $\alpha(4(r-1)-2\xi)-1<0$, we conclut that

$$\sum_{k=1}^{n} \mathbf{E}\left[|\xi_{k,n}|^2 ; |\xi_{k,n}| > \epsilon\right] \to 0 \text{ as } n \to \infty.$$

Finally, we obtain that

$$\begin{aligned}\frac{\sqrt{n}}{\gamma_n}\left(\widehat{H}_{\rho,n} - H_\rho\right) \to -\frac{\sqrt{n}}{\gamma_n}\left(\overline{Z} - \mathbf{E}[Z_1]\right) + \theta_1 \frac{\sqrt{p_n(1-p_n)}}{\gamma_n} \frac{\sqrt{n}}{\sqrt{p_n(1-p_n)}}(\widehat{p}_n - p_n) \\ + \frac{\theta_2}{\sqrt{p_n}\gamma_n}\sqrt{np_n}\left(\widehat{\beta}_n/\beta - 1\right) + \frac{\theta_3}{\sqrt{p_n}\gamma_n}\sqrt{np_n}\left(\widehat{\xi}_n - \xi\right) + o_{\mathbf{P}}(1),\end{aligned}$$

From **Lemma A-2** of Johansson 2003 [15], under the assumptions of Theorem 3.2, for any real numbers, t_1, t_2, t_3 and t_4, we have

$$\begin{aligned}&\mathbf{E}\left[\exp\left\{it_1\frac{\sqrt{n}}{\gamma_n}\left(\overline{Z} - \mathbf{E}[Z_1]\right) + i\sqrt{np_n}\,(t_2, t_3)\begin{pmatrix}\widehat{\beta}_n/\beta - 1 \\ \widehat{\xi}_n - \xi\end{pmatrix} + it_4\frac{\sqrt{n}\,(\widehat{p}_n - p_n)}{\sqrt{p_n(1-p_n)}}\right\}\right] \\ &\to \exp\left\{-\frac{t_1^2}{2} - \frac{1}{2}(t_2, t_3)\,\mathbf{\Sigma}^{-1}\begin{pmatrix}t_2 \\ t_3\end{pmatrix} - \frac{t_4^2}{2}\right\}(1 + o_{\mathbf{P}}(1)).\end{aligned}$$

as $n \to \infty$, where $\mathbf{\Sigma}^{-1}$ is that in (2.8), $\gamma_n^2 = Var(Z_1)$ and $i^2 = -1$. It follows that, with this result that

$$\frac{\sqrt{n}}{\gamma_n \sigma_n}\left(\widehat{H}_{\rho,n} - H_\rho\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \text{ as } n \to \infty,$$

where

$$\sigma_n^2 = 1 + \frac{\theta_1^2}{\gamma_n^2} p_n(1-p_n) + \frac{2(1+\xi)\theta_2^2}{p_n\gamma_n^2} + \frac{(1+\xi)^2\theta_3^2}{p_n\gamma_n^2} - \frac{(1+\xi)\theta_2\theta_3}{p_n\gamma_n^2}.$$

This complet the proof of Theorem 3.2. □

Acknowledgement

The author are very grateful to the referees for a careful reading of the paper that led to significant improvements of the earlier draft.

References

[1] Artzner, P. *Application of coherent capital requirements*, North American Actuarial Journal **3**(2), 11-25, 1999.

[2] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. *Coherent measures of risk*, Mathematical Finance **9**, 203-228, 1999.

[3] Balkema, A. and de Haan L. *Residual Lifetime at Great Age*, Annals of Probability **2**, 792-804, 1974.

[4] Beirlant, J., Matthys, G. and Dierckx, G. *Heavy-tailed distributions and rating*, Astin Bulletin **31**, 37-58, 2001.

[5] Bingham, N., Goldie, C. and Teugels, J. *Regular Variation*, Encyclopedia of Mathematics and its Applications **27**, Cambridge University Press, 1987.

[6] Brazauskas, V., Jones, B. L., Puri, M. L. and Zitikis, R. *Estimating conditional tail expectation with actuarial applications in view*, J. Statist. Plann. Inference **138**, 3590-3604, 2008.

[7] Cebrian, A. C., Denuit, M. and Lambert, P. *Generalized Pareto Fitto the Society of Actuaries Large Claims Database*, N. Am. Actuar. J. **7**, 18-36, 2003.

[8] Centeno, M. L. and Andrade e Silva, J. *Applying the proportional hazard premium calculation principle*, Astin Bulletin **35**, 00-00, 2005.
[9] Dhaene, J., Kukush, A., Linders, D. and Tang, Q. *Remarks on quantiles and distortion risk measures*, European Actuarial Journal **2**(2), 319-328, 2012.
[10] Furman, E. and Zitikis, R. *Weighted risk capital allocations*, Insurance: Mathematics and Economics **43**, 263-269, 2008.
[11] Greselin, R., Puri, ML. and Zitikis, R. *L-functions, processes, and statistics in measuring economic inequality and actuarial risks*, Statistics and Its Interface **2** (2), 227-245, 2009.
[12] Hürlimann W. *Inequalities for Look Back Option Strategies and Exchange Risk Modelling.* Paper presented at the First Euro-Japanese Workshop on Stochastic Modelling in Insurance, Finance, Production and Reliability, Brussels, 1998.
[13] Hua, L. and Joe, H. *Tail comonotonicity: Properties, constructions, and asymptotic additivity of risk measures*, Insurance: Mathematics and Economics **51** (2), 492-503, 2012.
[14] Ihaka, R. and Gentleman, R. *R: A language for data analysis and graphics*, J. Comput. Graph. Statist. **5**, 299-314, 1996.
[15] Johansson, J. *Estimating the mean of heavy-tailed distributions*, *Extremes* 6, 91-131, 2003.
[16] Jones, B. L. and Zitikis, R. *Empirical estimation of risk measures and related quantities*, North American Actuarial Journal, **7** (4), 44-54, 2003.
[17] Jones, B. L. and Zitikis, R. *Risk measures, distortion parameters, and their empirical estimation*, Insurance: Mathematics and Economics **41** (2), 279-297, 2007.
[18] Joseph H. T. Kim. *Bias correction for estimated distortion risk measure using the bootstrap*, Insurance: Mathematics and Economics **47** (2), 198-205, 2010.
[19] Necir, A. and Meraghni, D. *Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts*, Insurance: Mathematics and Economics **45**, 49-58, 2009.
[20] Necir, A., Rassoul, A. and Zitikis, R. *Estimating the conditional tail expectation in the case of heavy-tailed losses*, Journal of Probability and Statistics, 2010.
[21] Peng, L., Qib, Y., Wang, R. and Yang, J. *Jackknife empirical likelihood method for some risk measures and related quantities*, Insurance: Mathematics & Economics, **51**, 142-150, 2012.
[22] Pickands, J. *Statistical Inference Using Extreme Order Statistics*, Annals of Statistics **3**, 119-131, 1975.
[23] Reiss, R.D. and Thomas, M. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhauser, Basel, 2007.
[24] Resnick, S. *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, 1987.
[25] Rolski, T., Schimidli, H., Schmidt, V. and Teugels, J. L. *StochasticProcesses for Insurance and Finance*, John Wiley & Sons, Chichester, 1999.
[26] Seneta, E. *Regularly Varying Functions*, Lecture Notes in Mathematics, Springer-Verlag, **508**, 1976.
[27] Smith, R. *Estimating Tails of Probability Distributions*, The Annals of Statistics, **15**, 1174-1207, 1987.
[28] Wang, S. S. *Premium Calculation by Transforming the LayerPremium Density*, ASTIN Bulletin **26**, 71-92, 1996.
[29] Wang, S. and Young, V.R. *Ordering risks. Expected utility theory versus Yaari' dual theory of risk*, Insurance: Mathematics & Economics. **22**, 145-161, 1998.
[30] Wirch, J.L. and Hardy, M.R. *A synthesis of risk measures for capital adequacy*, Insurance: Math. Econom. **25**, 337-347, 1999.
[31] Yaari, M. E. *The Dual Theory of Choice Under Risk*, Econometrica **55**, 95-115, 1987.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 747 – 754

Improved ratio-type estimators of finite population variance using quartiles

Ramkrishna S. Solanki*† Housila P. Singh‡ and Surya K. Pal§

Abstract

In this paper we have proposed some ratio-type estimators of finite population variance using known values of parameters related to an auxiliary variable such as quartiles with their properties in simple random sampling. The suggested estimators have been compared with the usual unbiased and ratio estimators and the estimators due to [2], [12, 13, 14] and [3]. An empirical study is also carried out to judge the merits of the proposed estimator over other existing estimators of population variance using natural data set.

2000 AMS Classification: 62D05

Keywords: Study variable, Auxiliary variable, Bias, Mean squared error, Quartiles, Simple random sampling.

Received 10/04/2014 : Accepted 30/07/2014 Doi : 10.15672/HJMS.2014448247

1. Introduction

Estimating the finite population variance has great significance in various fields such as industry, agriculture, medical and biological sciences where we come across the populations which are likely to be skewed. Variation is present everywhere in our day to day life. It is law of nature that no two things or individuals are exactly alike. For instance, a physician needs a full understanding of variation in the degree of human blood pressure, body temperature and pulse rate for adequate prescription. A manufacture needs constant knowledge of the level of variation in people's reaction to his product to be able to known whether to reduce or increase his price, or improve the quality of his product. An agriculturist needs an adequate understanding of variations in climate factors especially from place to place (or time to time) to be able to plan on when, how and where

*Department of Mathematics and Statistics, College of Agriculture, Balaghat, JNKVV, Jabalpur, M.P., India
Email: ramkssolanki@gmail.com
†Corresponding Author.
‡S.S. in Statistics, Vikram University, Ujjain, India, Email: hpsujn@gmail.com
§S.S. in Statistics, Vikram University, Ujjain, India, Email: suryakantpal6676@gmail.com

to plant his crop. In manufacturing industries and pharmaceutical laboratories some of times researchers are more interested about the variation of their products or yields. Many more situations can be encountered in practice where the estimation of population variance of the study variable assumes importance. For these reasons various authors have paid their attention towards the estimation of population variance. In sample surveys, auxiliary information on the finite population under study is quite often available from previous experience, census or administrative databases. The sampling literature describes a wide variety of techniques for using auxiliary information to improve the sampling design and/or obtain more efficient estimators of finite population variance. It is well known that when the auxiliary information is to be used at the estimation stage, the ratio method of estimation is extensively employed. The ratio estimation method has been extensively used because of its intuitive appeal, computational simplicity and applicability to a general design. Perhaps, this is why many researchers have directed their efforts toward to get more efficient ratio-type estimators of the population variance by modifying the structure of existing estimators. Such as, [2], [5], [6, 7], [8] and [11] have suggested some modified estimators of population variance using known values of coefficient of variation, coefficient of kurtosis, coefficient of skewness of an auxiliary variable together with their biases and mean squared errors. We have known that the value of quartiles and their functions are unaffected by the extreme values or the presence of outliers in the population values. For this reason, [3] and [12, 13, 14] have considered the problem of estimating the population variance of the study variable using information on variance, quartiles, inter-quartile range, semi-quartile range and semi-quartile average of an auxiliary variable. In this paper our main goal is to estimate the unknown population variance of the study variable by improving the estimators suggested previously using same information on an auxiliary variable such as quartiles, inter-quartile range, semi-quartile range, semi-quartile average etc. The remaining part of the paper is organized as follows: The Section 2 introduced the notations and some existing estimators of population variance in brief. In Section 3, the ratio-type estimator of population variance is suggested and the expressions of their asymptotic biases and the mean squared errors are obtained. In addition, some members of suggested ratio-type estimators are also generated with their properties. The Section 4 is addressed the problem of efficiency comparisons of proposed ratio-type estimators with the usual unbiased estimator and the estimator due to [1], while Section 5 is focused on empirical study of proposed ratio-type estimators for the real data set. We conclude with a brief discussion in Section 6.

2. Notations and literature review

Much literature has been produced on sampling from finite populations to address the issue of the efficient estimation of the variance of a survey variable when auxiliary variables are available. Our analysis refers to simple random sampling without replacement (SRSWOR) and considers, for brevity, the case when only a single auxiliary variable is used. Let $U = (U_1, U_2, ..., U_N)$ be finite population of size N and (y, x) are (study, auxiliary) variables taking values (y_i, x_i) respectively for the i^{th} unit U_i of the finite population U. Our quest is to estimate the unknown population variance $S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2$ of study variable y, where $\bar{Y} = N^{-1}\sum_{i=1}^{N} y_i$ is the population mean of y. Let a simple random sample (SRS) of size n be drawn without replacement (WOR) from the finite population U. The usual unbiased estimator of finite population variance S_y^2 is defined as

$$(2.1)\qquad s_y^2 = t_0 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

where $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$. [1] has suggested the usual ratio estimator of S_y^2 as

$$t_R = t_1 = s_y^2\left(\frac{S_x^2}{s_x^2}\right) \tag{2.2}$$

where $S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2$, $\bar{X} = N^{-1}\sum_{i=1}^{N} x_i$, $s_x^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ and $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$. Motivated by [10], [15] and [9], [2] have proposed following ratio-type estimators of the population variance as

$$t_2 = s_y^2\left(\frac{S_x^2 - C_x}{s_x^2 - C_x}\right) \tag{2.3}$$

$$t_3 = s_y^2\left(\frac{S_x^2 - \beta_2(x)}{s_x^2 - \beta_2(x)}\right) \tag{2.4}$$

$$t_4 = s_y^2\left(\frac{\beta_2(x)S_x^2 - C_x}{\beta_2(x)s_x^2 - C_x}\right) \tag{2.5}$$

$$t_5 = s_y^2\left(\frac{C_xS_x^2 - \beta_2(x)}{C_xs_x^2 - \beta_2(x)}\right) \tag{2.6}$$

where $C_x = \left(S_x/\bar{X}\right)$ and $\beta_2(x)$ are the known coefficients of variation and kurtosis of the auxiliary variable x respectively. Using the known value of population median Q_2 of the auxiliary variable x [12]have suggested the ratio-type estimator of population variance S_y^2 as

$$t_6 = s_y^2\left(\frac{S_x^2 + Q_2}{s_x^2 + Q_2}\right) \tag{2.7}$$

[13] have proposed the modified ratio-type estimators of population variance S_y^2 of the study variable y using the known quartiles and their functions of the auxiliary variable x as

$$t_7 = s_y^2\left(\frac{S_x^2 + Q_1}{s_x^2 + Q_1}\right) \tag{2.8}$$

$$t_8 = s_y^2\left(\frac{S_x^2 + Q_3}{s_x^2 + Q_3}\right) \tag{2.9}$$

$$t_9 = s_y^2\left(\frac{S_x^2 + Q_r}{s_x^2 + Q_r}\right) \tag{2.10}$$

$$t_{10} = s_y^2\left(\frac{S_x^2 + Q_d}{s_x^2 + Q_d}\right) \tag{2.11}$$

$$t_{11} = s_y^2\left(\frac{S_x^2 + Q_a}{s_x^2 + Q_a}\right) \tag{2.12}$$

where Q_i is the i^{th} quartile $(i = 1, 3)$, $Q_r = (Q_3 - Q_1)$ (inter-quartile range) $Q_d = \left(\frac{Q_3 - Q_1}{2}\right)$ (semi-quartile range) and $Q_a = \left(\frac{Q_3 + Q_1}{2}\right)$ (semi-quartile average). Taking motivation form [2] and [12]; [14] have suggested the ratio-type estimators of population variance S_y^2 using known values of coefficient of variation C_x and population median Q_2 of an auxiliary variable x as

$$t_{12} = s_y^2\left(\frac{C_xS_x^2 + Q_2}{C_xs_x^2 + Q_2}\right) \tag{2.13}$$

Recently [3] have proposed another ratio-type estimator of population variance S_y^2 using known values of coefficient of correlation ρ between the variables (y, x) and population quartile Q_3 of an auxiliary variable x as

$$t_{13} = s_y^2\left(\frac{\rho S_x^2 + Q_3}{\rho s_x^2 + Q_3}\right) \tag{2.14}$$

To the first degree of approximation the biases and mean squared errors ($MSEs$) of the estimators t_j, $(j = 0, 1, 2, ..., 13)$ are respectively given as

(2.15) $\quad Bias\,(t_j) = \Phi\tau_j\,(\tau_j - c)$

(2.16) $\quad MSE\,(t_j) = \gamma\,[\lambda^*_{40} + \tau_j\lambda^*_{04}\,(\tau_j - 2c)]$

where $\Phi = \gamma\lambda^*_{04}$, $\gamma = n^{-1}S^4_y, c = \left(\lambda^*_{22}\lambda^{*-1}_{04}\right)$, $\tau_0 = 0$, $\tau_1 = 1$, $\tau_2 = S^2_x\left(S^2_x - C_x\right)^{-1}$, $\tau_3 = S^2_x\left(S^2_x - \beta_2\,(x)\right)^{-1}$, $\tau_4 = \beta_2\,(x)\,S^2_x\left(\beta_2\,(x)\,S^2_x - C_x\right)^{-1}$, $\tau_5 = C_xS^2_x\left(C_xS^2_x - \beta_2\,(x)\right)^{-1}$, $\tau_6 = S^2_x\left(S^2_x + Q_2\right)^{-1}$, $\tau_7 = S^2_x\left(S^2_x + Q_1\right)^{-1}$, $\tau_8 = S^2_x\left(S^2_x + Q_3\right)^{-1}$, $\tau_9 = S^2_x\left(S^2_x + Q_r\right)^{-1}$, $\tau_{10} = S^2_x\left(S^2_x + Q_d\right)^{-1}$, $\tau_{11} = S^2_x\left(S^2_x + Q_a\right)^{-1}$, $\tau_{12} = C_xS^2_x\left(C_xS^2_x + Q_2\right)^{-1}$, $\tau_{13} = \rho S^2_x\left(\rho S^2_x + Q_3\right)^{-1}$, $\lambda^*_{rs} = (\lambda_{rs} - 1)$, $\lambda_{rs} = \mu_{rs}\left(\mu^{s/2}_{02}\mu^{r/2}_{20}\right)^{-1}$, $\mu_{rs} = N^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^r\left(x_i - \bar{X}\right)^s$ (r, s being non negative integers). It is observed that the estimators $(t_6, t_7, ..., t_{13})$ due to [12, 13, 14] and [3] have used the quartiles and their functions such as inter-quartile range Q_r , semi-quartile range Q_d and semi-quartile average Q_a and in additive form to sample and population variances s^2_x and S^2_x respectively of the auxiliary variable x. It is to be noted that the unit of the quartiles and their function as given above is of original variable x, while the unit of S^2_x and s^2_x are in the square of the unit of the original variable x. These lead authors to develop a more justified ratio-type estimators of the population variance S^2_y of the study variable y using known values of parameters related to the auxiliary variable x and study their properties in simple random sampling.

3. The proposed ratio-type estimator

We propose following ratio-type estimators of population variance S^2_y in simple random sampling as

(3.1) $$T = s^2_y\left(\frac{\delta S^2_x + \alpha L^2}{\delta s^2_x + \alpha L^2}\right)$$

where $\left(\delta S^2_x + \alpha L^2\right) > 0, \left(\delta s^2_x + \alpha L^2\right) > 0$ and (δ, L) are either real constants or function of known parameters of an auxiliary variable x with $0 \leq \alpha \leq 1$. To obtain the bias and MSE of the proposed ratio-type estimator T, we write $s^2_y = S^2_y\,(1 + e_0)$ and $s^2_x = S^2_x\,(1 + e_1)$ such that $E\,(e_0) = E\,(e_1) = 0$ and to the first degree of approximation (ignoring finite population correction (f.p.c.) term), we have $E\left(e^2_0\right) = n^{-1}\lambda^*_{40}$, $E\left(e^2_1\right) = n^{-1}\lambda^*_{04}$, $E\,(e_0e_1) = n^{-1}\lambda^*_{22}$. Now expressing (3.1) in terms of e's, we have

(3.2) $$T = S^2_y\,(1 + e_0)\left[\frac{\delta S^2_x + \alpha L^2}{\delta S^2_x\,(1 + e_1) + \alpha L^2}\right] = S^2_y\,(1 + e_0)\,(1 + \tau^*e_1)^{-1}$$

where $\tau^* = \delta S^2_x\left(\delta S^2_x + \alpha L^2\right)^{-1}$. We assume that $|\tau^*e_1| < 1$ so that $(1 + \tau^*e_1)^{-1}$ is expendable. Expanding the right hand side of (3.2) and multiplying out, we have

$T = S^2_y\,(1 + e_0)\left(1 - \tau^*e_1 + \tau^{*2}e^2_1 - ...\right)$

$= S^2_y\left(1 + e_0 - \tau^*e_1 - \tau^*e_0e_1 + \tau^{*2}e^2_1 + \tau^{*2}e_0e^2_1 - ...\right)$

Neglecting terms of e's having power greater than the two, we have

$T \cong S^2_y\left(1 + e_0 - \tau^*e_1 - \tau^*e_0e_1 + \tau^{*2}e^2_1\right)$

or

(3.3) $$\left(T - S^2_y\right) \cong S^2_y\left(e_0 - \tau^*e_1 - \tau^*e_0e_1 + \tau^{*2}e^2_1\right)$$

Taking expectation of both sides of (3.3),we get the bias of the estimator $T$to the first degree of approximation as

(3.4) $$Bias\,(T) = \Phi\tau^*\,(\tau^* - c)$$

Squaring both sides of (3.3) and neglecting terms of e's having power greater than two, we have

$$(3.5) \qquad \left(T - S_y^2\right)^2 \cong S_y^4 \left(e_0^2 + \tau^{*2} e_1^2 - 2\tau^* e_0 e_1\right)$$

Taking expectation of both sides of (3.5), we get the MSE of the estimator T to the first degree of approximation as

$$(3.6) \qquad MSE\,(T) = \gamma\left[\lambda_{40}^* + \tau^* \lambda_{04}^* \left(\tau^* - 2c\right)\right]$$

Below we have identified some members of proposed ratio type estimator T for different choices of (δ, L).

(i) The estimator based on coefficient of variation C_x and quartile Q_1:
If we set $(\delta, L) = (C_x, Q_1)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.7) \qquad T_1 = s_y^2 \left(\frac{C_x S_x^2 + \alpha Q_1^2}{C_x s_x^2 + \alpha Q_1^2}\right)$$

(ii) The estimator based on coefficient of kurtosis $\beta_2(x)$ and median Q_2:
If we set $(\delta, L) = (\beta_2(x), Q_2)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.8) \qquad T_2 = s_y^2 \left(\frac{\beta_2(x) S_x^2 + \alpha Q_2^2}{\beta_2(x) s_x^2 + \alpha Q_2^2}\right)$$

(iii) The estimator based on population mean $\bar{X}$ and quartile Q_3:
If we set $(\delta, L) = \left(\bar{X}, Q_3\right)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.9) \qquad T_3 = s_y^2 \left(\frac{\bar{X} S_x^2 + \alpha Q_3^2}{\bar{X} s_x^2 + \alpha Q_3^2}\right)$$

(iv) The estimator based on coefficient of kurtosis $\beta_2(x)$ and inter-quartile range Q_r:
If we set $(\delta, L) = (\beta_2(x), Q_r)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.10) \qquad T_4 = s_y^2 \left(\frac{\beta_2(x) S_x^2 + \alpha Q_r^2}{\beta_2(x) s_x^2 + \alpha Q_r^2}\right)$$

(v) The estimator based on correlation coefficient ρ and semi-quartile range Q_d:
If we set $(\delta, L) = (\rho, Q_d)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.11) \qquad T_5 = s_y^2 \left(\frac{\rho S_x^2 + \alpha Q_d^2}{\rho s_x^2 + \alpha Q_d^2}\right)$$

(vi) The estimator based on correlation coefficient ρ and semi-quartile average Q_a:
If we set $(\delta, L) = (\rho, Q_a)$ in (3.1), we get the estimator of S_y^2 as,

$$(3.12) \qquad T_6 = s_y^2 \left(\frac{\rho S_x^2 + \alpha Q_a^2}{\rho s_x^2 + \alpha Q_a^2}\right)$$

Similarly one can identify many other estimators from the proposed ratio-type estimator T for different combinations of (δ, L). To the first degree of approximation the biases and mean squared errors $(MSEs)$ of the estimators $T_k, (k = 1, 2, ..., 6)$ are respectively given by

$$(3.13) \qquad Bias\,(T_k) = \Phi \tau_k^* \left(\tau_k^* - c\right)$$

$$(3.14) \qquad MSE\,(T_k) = \gamma\left[\lambda_{40}^* + \tau_k^* \lambda_{04}^* \left(\tau_k^* - 2c\right)\right]$$

where $\tau_1^* = C_x S_x^2 \left(C_x S_x^2 + \alpha Q_1^2\right)^{-1}$, $\tau_2^* = \beta_2(x) S_x^2 \left(\beta_2(x) S_x^2 + \alpha Q_2^2\right)^{-1}$, $\tau_3^* = \bar{X} S_x^2 \left(\bar{X} S_x^2 + \alpha Q_3^2\right)^{-1}$, $\tau_4^* = \beta_2(x) S_x^2 \left(\beta_2(x) S_x^2 + \alpha Q_r^2\right)^{-1}$, $\tau_5^* = \rho S_x^2 \left(\rho S_x^2 + \alpha Q_d^2\right)^{-1}$, $\tau_6^* = \rho S_x^2 \left(\rho S_x^2 + \alpha Q_a^2\right)^{-1}$.

Table 1. The parameters of population data set

N	80	C_y	0.3542	Q_1	5.1500
n	20	S_x	8.4563	Q_2	10.300
$\bar{Y}$	51.8264	C_x	0.7507	Q_3	16.975
$\bar{X}$	11.2646	λ_{04}	2.8664	Q_r	11.825
ρ	0.9413	λ_{40}	2.2667	Q_d	5.9125
S_y	18.3569	λ_{22}	2.2209	Q_a	11.0625

Table 2. $PREs$ of estimators $t_j, (j = 0, 1, ..., 13)$ with respect to s_y^2

Percent relative efficiency (PRE)						
(t_0, s_y^2)	(t_1, s_y^2)	(t_2, s_y^2)	(t_3, s_y^2)	(t_4, s_y^2)	(t_5, s_y^2)	(t_6, s_y^2)
100.00	183.23	179.62	169.24	181.98	164.49	226.87
Percent relative efficiency (PRE)						
(t_7, s_y^2)	(t_8, s_y^2)	(t_9, s_y^2)	(t_{10}, s_y^2)	(t_{11}, s_y^2)	(t_{12}, s_y^2)	(t_{13}, s_y^2)
206.64	247.25	232.13	209.86	229.54	238.17	249.84

4. The theoretical evaluation

We have made some theoretical conditions under which the ratio-type estimators $T_k, (k = 1, 2, ..., 6)$ which are members of proposed ratio-type estimator T are more efficient than the other existing estimators $t_j, (j = 0, 1, ..., 13)$ which are due to [1], [2], [12, 13, 14] and [3] respectively. From (2.16) and (3.14), we have

$MSE\left(T_k\right) < MSE\left(T_j\right)$ if $\tau_k^*(\tau_k^* - 2c) < \tau_j(\tau_j - 2c)$

i.e. if either,

$\tau_k^* < \tau_j$ and $c < \left(\frac{\tau_k^* + \tau_j}{2}\right)$

or,

$\tau_k^* > \tau_j$ and $c > \left(\frac{\tau_k^* + \tau_j}{2}\right)$

or equivalently ,

$\min.[\tau_j, (2c - \tau_j)] \leq \tau_k^* \leq \max.[\tau_j, (2c - \tau_j)]\,, (j = 0, 1, ..., 13; k = 1, 2, ..., 6).$

5. Emprical study

The performance of the ratio-type estimators $T_k, (k = 1, 2, ..., 6)$ which are members of the suggested ratio-type estimator T are evaluated against the usual unbiased estimator s_y^2 and the estimators $t_j, (j = 1, 2, ..., 13)$ which are due to [1], [2], [12, 13, 14] and [3] respectively. for the population data set [Source: [4]] summarized in Table 1. We have computed the percent relative efficiencies ($PREs$) of the estimators $t_j, (j = 1, 2, ..., 13)$ and the suggested ratio-type estimators $T_k, (k = 1, 2, ..., 6)$ with respect to the usual unbiased estimator $t_0 = s_y^2$ in certain range of $\alpha \in (0.0, 1.0)$ by using following formulae respectively as

$$(5.1)\qquad PRE\left(t_j, s_y^2\right) = \frac{MSE\left(s_y^2\right)}{MSE\left(t_j\right)} \times 100 = \frac{\lambda_{40}^*}{[\lambda_{40}^* + \tau_j \lambda_{04}^*(\tau_j - 2c)]} \times 100$$

$$(5.2)\qquad PRE\left(T_k, s_y^2\right) = \frac{MSE\left(s_y^2\right)}{MSE\left(T_k\right)} \times 100 = \frac{\lambda_{40}^*}{[\lambda_{40}^* + \tau_k^* \lambda_{04}^*(\tau_k^* - 2c)]} \times 100$$

and finding are summarized in Tables 2 and 3. It is observed from Tables 2 and 3 that all the ratio-type estimators $T_k, (k = 1, 2, ..., 6)$ which are members of proposed ratio-type estimator T performed better than the usual unbiased estimator s_y^2, usual ratio estimator

Table 3. $PREs$ of estimators $T_k, (k = 1, 2, ..., 6)$ with respect to s_y^2

	Percent relative efficiency (PRE)					
α	(T_1, s_y^2)	(T_2, s_y^2)	(T_3, s_y^2)	(T_4, s_y^2)	(T_5, s_y^2)	(T_6, s_y^2)
0.0	183.23	183.23	183.23	183.23	183.23	183.23
0.1	199.59	200.34	195.20	205.48	200.39	235.94
0.2	214.59	215.93	206.50	224.91	216.03	264.03
0.3	227.93	229.68	217.01	240.89	229.81	270.58
0.4	239.41	241.38	226.62	253.18	241.52	264.67
0.5	248.96	250.94	235.28	261.89	251.09	253.47
0.6	256.58	258.41	242.92	267.36	258.55	240.92
0.7	262.37	263.91	249.54	270.08	264.02	228.80
0.8	266.47	267.63	255.15	270.58	267.71	217.77
0.9	269.08	269.79	259.79	269.35	269.83	207.99
1.0	270.38	270.61	263.50	266.84	270.61	199.40

t_1 due to [1] and the estimators $t_j, (j = 2, 3, 4, 5)$ due to [2] for all $\alpha \in (0.0, 1.0)$. However all the ratio-type estimators $T_k, (k = 1, 2, ..., 6)$ are more efficient than the estimators $t_j, (j = 6, 7, ..., 12)$ due to [12, 13, 14] and [3] for a specific value of α. The estimators T_2 and T_5 which utilize the information on $(\beta_2(x), Q_2)$ and (ρ, Q_d) respectively are the best in the sense of having largest percent relative efficiency among all the estimators discussed here for $\alpha = 1$.

6. Conclusion

In this paper we have suggested some ratio-type estimators of population variance S_y^2 of the study variable y using known parameters of an auxiliary variable such as coefficient of variation, coefficient of kurtosis, correlation coefficient and quartiles etc. The bias and mean squared error formulae of the proposed ratio-type estimators are obtained and compared with that of the usual unbiased estimator, traditional ratio estimator and the estimators due to [2], [12, 13, 14] and [3]. We have also assessed the performance of the proposed estimators for known natural population data set and found that the performances of the proposed estimators are better than the other existing estimator for certain cases.

References

[1] Isaki, C.T. *Variance estimation using auxiliary information*, Journal of the American Statistical Association **78** (381), 117-123, 1983.

[2] Kadilar, C. and Cingi, H. *Ratio estimators for the population variance in simple and stratified random sampling*, Applied Mathematics and Computation **173** (2), 1047-1059, 2006.

[3] Khan, M. and Shabbir, J. *A ratio-type estimator for the estimation of population variance using quartiles of an auxiliary variable*, Journal of Statistics Applications and Probability **2** (3), 319-325, 2013.

[4] Murthy, M.N. *Sampling: Theory and Methods*, Statistical Publishing Society, Calcutta, India, 1967.

[5] Singh, H.P. and Solanki, R.S. *Estimation of finite population variance using auxiliary information in presence of random non-response*, Gujarat Statistical Review **36-37** (1-2), 46-58, 2009-2010.

[6] Singh, H.P. and Solanki, R.S. *A new procedure for variance estimation in simple random sampling using auxiliary information*, Statistical Papers **54** (2), 479-497, 2013a.

[7] Singh, H.P. and Solanki, R.S. *Improved estimation of finite population variance using auxiliary information*, Communications in Statistics-Theory and Methods **42** (15), 2718-2730, 2013b.

[8] Singh, H.P., Pal, S.K. and Solanki, R.S. *Improved estimation of finite population variance using quartiles*, ISTATISTIK **6** (3), 116-121, 2013.

[9] Singh, H.P., Tailor, R., Tailor, R. and Kakran, M.S. *An improved estimator of population mean using power transformation*, Journal of the Indian Society of Agricultural Statistics **58** (2), 223-230, 2004.

[10] Sisodia, B.V.S. and Dwivedi, V.K. *A modified ratio estimator using coefficient of variation of auxiliary variable*, Journal of the Indian Society of Agricultural Statistics **33** (2), 13-18, 1981.

[11] Solanki, R.S. and Singh, H.P. *An improved class of estimators for the population variance*, Model Assisted Statistics and Applications **8** (3), 229-238, 2013.

[12] Subramani, J. and Kumarapandiyan, G. *Variance estimation using median of the auxiliary variable*, International Journal of Probability and Statistics **1** (3), 62-66, 2012a.

[13] Subramani, J. and Kumarapandiyan, G. *Variance estimation using quartiles and their functions of an auxiliary variable*, International Journal of Statistics and Applications **2** (5), 67-72, 2012b.

[14] Subramani, J. and Kumarapandiyan, G. *Estimation of variance using known co-efficient of variation and median of an auxiliary variable*, Journal of Modern Applied Statistical Methods **12** (1), 58-64, 2013.

[15] Upadhyaya, L.N. and Singh, H.P. *Use of transformed auxiliary variable in estimating the finite population mean*, Biometrical Journal **41** (5), 627-636, 1999.

Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 755 – 760

Specification test for fixed effects in binary panel data model: a simulation study

Gang Yu * , Weiguo Wang † , Maozu Lu ‡ and Shaoping Wang §

Abstract

In this paper, we examine the specification tests which have been proposed for fixed effects in binary panel data model, using several different data generating processes to evaluate the performance of the specification test in different situations. By simulations, we find the specification test based on moment conditions is able to outperform the Lagrange multiplier test proposed by Gurmu [5].

2000 AMS Classification: 62F03, 62F05.

Keywords: Specification test, Fixed effects, Binary panel data.

Received 16/12/2013 : Accepted 30.06.2014 Doi : 10.15672/HJMS.2014227473

---

*Center for Econometrics Analysis and Forecasting, School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Liaoning, 116025, P.R.China.
School of Economics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R.China
Email: yugang543@163.com Corresponding Author.

†Center for Econometrics Analysis and Forecasting, School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Liaoning, 116025, P.R.China
Email:wwguodufe@163.com

‡Center for Econometrics Analysis and Forecasting, School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Liaoning, 116025, P.R.China
Email:maozulu@gmail.com

§School of Economics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R.China
Email:wangspi@sina.com

1. Introduction

Binary panel data models remain of major interest in microeconometrics. This paper examines the specification test for fixed effects in binary panel data model. The binary panel data model is in the following form:

$$y_{it} = 1(x_{it}'\beta + \eta_i + v_{it} \geq 0), \qquad i = 1, \cdots, N; t = 1, \cdots, T, \tag{1}$$

where $1(\cdot)$ denotes the indicator function that equals one if $\cdot$ is true and zero otherwise, y_{it} is an observed dependent variable, x_{it} is a $k \times 1$ vector of exogenous regressors, η_i denotes the individual's fixed effects and v_{it} is unobservable error term which is independently identical distributed with cdf $F(x)$ across units and time periods, where $F(x)$ is known and symmetric around 0.

In the binary panel data model (1), fixed effects estimation suffers from inconsistency under the incidental parameters problem, first considered by Neyman and Scott [9]. The incidental parameters problem persists in the binary panel data case because the nuisance parameters η_i can not be separated from estimators of coefficients of interest. As both N and T increase, the increasing number of parameters to estimate means that the coefficients will have an asymptotic bias.

Baltagi [1] proposes an open problem in Econometric Theory, i.e. the following test for fixed effects in binary panel data model (1):

$$H_0 : \eta_i = 0 \quad \text{for} \quad i = 1, \cdots, N. \tag{2}$$

If H_0 is not reject, the estimation procedure is simple and utilizes the usual logit and probit procedures. However, if H_0 is rejected, the maximum likelihood procedure is complicated by the presence of the incidental parameters problem. Furthermore, Gurmu [5] solves the open problem and proposes the lagrange multiplier (LM) test for the test problem H_0 by artificial regression, which is analogous to those used for tests in binary response model regression(BRMR) proposed by Davidson and MacKinnon [4], and shows LM $\sim \chi^2(N)$ under the null hypothesis. Some discussions about test for fixed effects in binary panel data also can be found in Baltagi [2]. Both Gurmu [5] and Baltagi [2] do not present the Monte Carlo simulations studies, LM test's small sample performance is unknown and will be tested in this paper through the use of Monte Carlo simulations.

For test problem (2) in binary panel data proposed by Baltagi [1], this paper also derives a test based on moment conditions, which asymptotic null distribution is the $\chi^2(1)$ distribution. The test is applied to Monte Carlo simulations and its power is compared with LM test proposed by Gurmu [5].

The structure of the paper is organized as follows. Section 2 introduces the test statistic based on moment conditions and its large sample properties. In section 3, we report some Monte Carlo simulation results. Section 4 concludes the paper.

2. Specification test based on moment conditions

The framework of deriving the test statistic is similar to Mora and Moro-Egido [8]. We assume that independent and identically distributed (i.i.d) observations $(y_{it}, x_{it}')'$ are available, where $i = 1, \cdots, N; t = 1, \cdots, T$. The following notation will be used: $p_{1,it}(\theta) \equiv \Pr(y_{it} = 1|x_{it}') = F(x_{it}'\beta + \eta_i), p_{0,it}(\theta) \equiv \Pr(y_{it} = 0|x_{it}') = 1 - p_{1,it}(\theta), p_{it} \equiv [p_{1,it}(\theta)]^{y_{it}} \times [p_{0,it}(\theta)]^{1-y_{it}}$, where $\theta = (\beta', \eta')'$ and $\eta = (\eta_1, \cdots, \eta_N)'$. Conditioning on the observations, the MLE of θ, $\hat{\theta} = (\hat{\beta}', \hat{\eta}')'$ maximizes the following log-likelihood function $l(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T} \ln p_{it}$.

Define $m_{it}(\theta) \equiv y_{it} - F(x'_{it}\beta + \eta_i)$. From binary panel data model (1), we have $\mathrm{E}m_{it}(\theta) = 0$. To derive the test statistic, we consider the random variable $\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})$, where $\hat{\theta} = (\hat{\beta}', 0')'$ is a well-behaved maximum likelihood estimator of $\theta_0 = (\beta', 0')'$ and $\hat{\beta}$ is the vector of ML estimate subject to the restriction $H_0 : \eta_i = 0 \quad \text{for} \quad i = 1, \cdots, N$.

2.1. Theorem. *Consider model (1), assuming the following regularity conditions hold,*
(i) In the neighborhood of true value θ_0, $\partial \ln p_{it}/\partial\theta, \partial^2 \ln p_{it}/\partial\theta^2, \partial^3 \ln p_{it}/\partial\theta^3$ exist;
(ii) In the neighborhood of true value θ_0, $|\partial^3 \ln p_{it}/\partial\theta^3| \leq H(x)$, and $EH(x) < \infty$;
(iii) At the true value θ_0, $E_{\theta_0}[\partial \ln p_{it}/\partial\theta] = 0, E_{\theta_0}[p''_{it}/p_{it}] = 0$,
$I(\theta_0) = Var_{\theta_0}[\partial \ln p_{it}/\partial\theta] > 0$.
Under the null hypothesis given in equation (2), when $N, T \longrightarrow \infty$, the C^M_{NT} statistics

$$C^M_{NT} = (NT)^{-1}[\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})]^2/\hat{V} \tag{3}$$

converges to a chi squared distribution with one degree, where

$$\hat{V} = (NT)^{-1}\{\sum_{i=1}^{N}\sum_{t=1}^{T} m^2_{it}(\hat{\theta}) - [\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})g_{it}(\hat{\theta})]^2/\sum_{i=1}^{N}\sum_{t=1}^{T} g^2_{it}(\hat{\theta})\}. \tag{4}$$

3. Monte Carlo simulation study

In this section, we present a small Monte Carlo study to illustrate the performance of the above test statistic (3) proposed in Section 2. For comparison, we also report the finite sample sizes and powers of LM test proposed by Gurmu [5].

The simulation is based on the logit model

$$y_{it} = 1(x_{it}\beta + \eta_i + v_{it} \geq 0), \qquad i = 1, \cdots, N; t = 1, \cdots, T, \tag{5}$$

where the true parameter value is $\beta = 1$, x_{it} is an exogenous variable and independently identical distributed with distribution $N(0, 1)$, v_{it} is independently identical distributed with logistic distribution $P\{v_{it} < x\} = F(x) = e^x/(1+e^x)$, and $\eta_i = (\sum_{t=1}^{T} z_{it})/T$, z_{it} is an exogenous variable and independently identical distributed with distribution $N(\mu, \sigma^2)$, so that the fixed effects η_i are generated from normal distribution. In model (5), we use statistics (3) to test $H_0 : \eta_i = 0 \quad \text{for} \quad i = 1, \cdots, N$. Parameter β is estimated by ML estimate assuming that H_0 holds. Values of both μ and σ^2 different from 0 allow us to examine the ability of the test statistic to detect misspecification in binary panel data model.

Table 1a Empirical sizes for logit design with different N and T.

N	Test	T=5			T=10			T=15		
		1%test	5%test	10%test	1%test	5%test	10%test	1%test	5%test	10%test
50	LM	0.013	0.052	0.103	0.008	0.044	0.114	0.002	0.051	0.093
	C^M_{NT}	0.004	0.060	0.117	0.021	0.062	0.118	0.016	0.057	0.113
100	LM	0.008	0.031	0.090	0.001	0.049	0.102	0.011	0.052	0.103
	C^M_{NT}	0.020	0.050	0.108	0.015	0.061	0.103	0.017	0.063	0.124
200	LM	0.003	0.034	0.078	0.006	0.040	0.112	0.010	0.052	0.083
	C^M_{NT}	0.012	0.062	0.108	0.015	0.066	0.091	0.019	0.062	0.119

Table 1b Empirical powers for logit design with different T when N=50.

μ	σ	Test	T=5			T=10			T=15		
			1%test	5%test	10%test	1%test	5%test	10%test	1%test	5%test	10%test
0.2	0.2	LM	0.010	0.061	0.169	0.030	0.139	0.201	0.045	0.164	0.275
		C_{NT}^M	0.148	0.334	0.450	0.335	0.602	0.709	0.527	0.753	0.837
	0.4	LM	0.011	0.093	0.163	0.035	0.145	0.242	0.056	0.180	0.298
		C_{NT}^M	0.177	0.346	0.460	0.365	0.587	0.682	0.518	0.715	0.826
	0.6	LM	0.016	0.101	0.210	0.032	0.176	0.289	0.072	0.249	0.335
		C_{NT}^M	0.171	0.360	0.449	0.340	0.582	0.668	0.531	0.738	0.824
	0.8	LM	0.023	0.130	0.230	0.061	0.207	0.354	0.101	0.289	0.399
		C_{NT}^M	0.155	0.322	0.437	0.340	0.540	0.675	0.527	0.722	0.819
0.4	0.2	LM	0.031	0.202	0.320	0.202	0.432	0.610	0.387	0.642	0.760
		C_{NT}^M	0.677	0.839	0.893	0.965	0.989	0.997	0.995	0.999	0.999
	0.4	LM	0.045	0.213	0.350	0.221	0.471	0.609	0.365	0.609	0.777
		C_{NT}^M	0.677	0.834	0.896	0.948	0.987	0.994	0.993	0.994	1.000
	0.6	LM	0.068	0.222	0.356	0.224	0.506	0.626	0.495	0.692	0.820
		C_{NT}^M	0.689	0.826	0.892	0.955	0.984	0.992	0.988	0.996	1.000
	0.8	LM	0.081	0.275	0.437	0.287	0.555	0.712	0.509	0.746	0.838
		C_{NT}^M	0.646	0.827	0.874	0.935	0.980	0.989	0.992	1.000	0.997

Table 2a Empirical sizes for probit design with different N and T.

N	Test	T=5			T=10			T=15		
		1%test	5%test	10%test	1%test	5%test	10%test	1%test	5%test	10%test
50	LM	0.002	0.025	0.067	0.008	0.052	0.086	0.009	0.038	0.086
	C_{NT}^M	0.012	0.064	0.106	0.007	0.052	0.095	0.011	0.056	0.111
100	LM	0.002	0.045	0.068	0.006	0.037	0.087	0.014	0.053	0.091
	C_{NT}^M	0.012	0.065	0.107	0.014	0.053	0.097	0.011	0.063	0.116
200	LM	0.001	0.034	0.069	0.003	0.036	0.092	0.006	0.050	0.074
	C_{NT}^M	0.007	0.053	0.107	0.011	0.042	0.101	0.006	0.050	0.104

Table 2b Empirical powers for probit design with different T when N=50

μ	σ	Test	T=5			T=10			T=15		
			1%test	5%test	10%test	1%test	5%test	10%test	1%test	5%test	10%test
0.2	0.2	LM	0.011	0.072	0.139	0.053	0.210	0.298	0.132	0.320	0.487
		C_{NT}^M	0.333	0.586	0.676	0.710	0.876	0.930	0.899	0.963	0.981
	0.4	LM	0.022	0.116	0.234	0.086	0.248	0.393	0.179	0.431	0.525
		C_{NT}^M	0.359	0.578	0.695	0.677	0.876	0.921	0.883	0.965	0.976
	0.6	LM	0.035	0.190	0.333	0.170	0.342	0.516	0.270	0.509	0.646
		C_{NT}^M	0.347	0.545	0.622	0.670	0.878	0.908	0.871	0.955	0.982
	0.8	LM	0.092	0.291	0.481	0.274	0.512	0.685	0.396	0.661	0.770
		C_{NT}^M	0.358	0.554	0.690	0.635	0.819	0.886	0.829	0.938	0.963
0.4	0.2	LM	0.108	0.328	0.501	0.606	0.815	0.898	0.916	0.980	0.993
		C_{NT}^M	0.957	0.995	0.995	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	LM	0.146	0.396	0.575	0.651	0.861	0.918	0.940	0.983	0.993
		C_{NT}^M	0.945	0.989	0.995	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	LM	0.215	0.480	0.657	0.720	0.884	0.954	0.940	0.989	0.996
		C_{NT}^M	0.958	0.984	0.995	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	LM	0.285	0.594	0.736	0.816	0.938	0.967	0.973	0.995	0.997
		C_{NT}^M	0.919	0.970	0.988	0.999	1.000	1.000	1.000	1.000	1.000

The simulation results of our test based on moment(C_{NT}^M), and Gurmu's test (LM) are reported in Table 1a and Table 1b based on 1000 simulations, where the nominal sizes are set to be 0.01, 0.05 and 0.10. From Table 1a, the empirical sizes for both tests are very close to the nominal sizes, with the LM test having less size distortion in most cases. From Table 1b, the proposed test C_{NT}^M is more powerful than Gurmu's LM test in all designs, and the powers significantly increase demonstrated by increasing the panel length T.

On the above data generating process (DGP), if assuming $F(x)$ is the Normal cdf, we also report the simulation results for a probit model in Table 2a and Table 2b. The results are qualitatively similar to those for the logit model in the previous Table 1a and Table 1b.

4. Conclusion

Specification test is an important part of panel data econometrics. This paper focuses on examining the specification test for fixed effects in binary panel data model by Monte Carlo simulations. The simulation results of this paper, along with the earlier work, show that the proposed test C_{NT}^{M} is more powerful than Gurmu's LM test.

In economics, it is more realistic to consider dynamic binary panel data model with fixed effects, for example, Hsiao [6], Bartolucci and Nigro [3], Yu, Gao and Shi [11] . As a possible area of further research it would be interesting to investigate the specification test for fixed effects in dynamic binary panel data model by using the proposed test C_{NT}^{M}.

5. Appendix A: Proof of results

Proof of Theorem 2.1: Using a first-order Taylor expansion for $\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})$, where $\hat{\theta}$ is the maximum likelihood estimator under H_0, we have

$$(NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta}) = (NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\theta_0) + B_0 \times (NT)^{1/2}(\hat{\theta}-\theta_0) + o_p(1), \tag{A.1}$$

where $B_0 = \mathrm{E}\{\partial m_{it}(\theta_0)/\partial\theta'\}$.

Under the conditions (i)-(iii) of Theorem 2.1, see the detailed proof of Theorem 2.3 in page 415, Lehmann [7], the ML estimator $\hat{\theta}$ satisfies that

$$(NT)^{1/2}(\hat{\theta}-\theta_0) = A_0^{-1} \times (NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} g_{it}(\theta_0) + o_p(1), \tag{A.2}$$

where $g_{it}(\theta_0) = \partial \ln p_{it}/\partial\theta = 1(y_{it}=0) \times \dfrac{-f(x_{it}'\beta)x_{it}}{1-F(x_{it}'\beta)} + 1(y_{it}=1) \times \dfrac{f(x_{it}'\beta)x_{it}}{F(x_{it}'\beta)}$, $f(x)$ denotes the first derivative of $F(x)$, and $A_0 = E\{-\partial g_{it}(\theta_0)/\partial\theta'\} = I(\theta_0)$.

Inserting the asymptotic expansion of $(NT)^{1/2}(\hat{\theta}-\theta_0)$ into the Taylor expansion of $(NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})$, we have

$$(NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta}) = (NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}[m_{it}(\theta_0) + B_0A_0^{-1}g_{it}(\theta_0)] + o_p(1),$$

and we know that random variables $m_{it}(\theta_0)+B_0A_0^{-1}g_{it}(\theta_0)$ is independent and identically distributed, $\mathrm{E}(m_{it}(\theta_0) + B_0A_0^{-1}g_{it}(\theta_0)) = 0$, by central limit theorem (CLT),

$$(NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta}) \xrightarrow{d} N(0, V), \tag{A.3}$$

where $V = \mathrm{Var}(m_{it}(\theta_0) + B_0A_0^{-1}g_{it}(\theta_0)$.

We can find that $m_{it}(\theta_0) + B_0A_0^{-1}g_{it}(\theta_0) = (1 : B_0A_0^{-1})(m_{it}(\theta_0), g_{it}(\theta_0))'$, so

$$V = (1 : B_0A_0^{-1}) \begin{pmatrix} \mathrm{E}m_{it}^2(\theta_0) & \mathrm{E}[m_{it}(\theta_0)g_{it}(\theta_0)] \\ \mathrm{E}[m_{it}(\theta_0)g_{it}(\theta_0)] & \mathrm{E}g_{it}^2(\theta_0) \end{pmatrix} (1 : B_0A_0^{-1})'$$

$$= \mathrm{E}\{m_{it}^2(\theta_0)\} - \mathrm{E}^2\{m_{it}(\theta_0)g_{it}(\theta_0)\}/\mathrm{E}\{g_{it}^2(\theta_0)\},$$

where $\mathrm{E}[g_{it}^2(\theta_0)] = A_0, \mathrm{E}[m_{it}(\theta_0)g_{it}(\theta_0)] = -B_0$.

To obtain the a test statistic,a consistent estimator V must be proposed. The natural candidate for estimating V is

$$\hat{V} = (NT)^{-1}\{\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}^2(\hat{\theta}) - [\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})g_{it}(\hat{\theta})]^2 / \sum_{i=1}^{N}\sum_{t=1}^{T} g_{it}^2(\hat{\theta})\}, \tag{A.4}$$

we replace population moments by sample moments, it is a standard estimate of V following Newey-Tauchen methodology, detailed discussion can be found in Orme [10].

Based on (A.3), (A.4) and Slutsky' Theorm, the test statistic proposed

$$C_{NT}^{M} = (NT)^{-1}[\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}(\hat{\theta})]^2/\hat{V} \xrightarrow{d} \chi^2(1), \tag{A.5}$$

what justifies the use of C_{NT}^{M} as an asymptotically valid test statistics. This completes the proof.

Acknowledgments

The authors would like to thank the editor Cem kadilar and two anonymous referees for their constructive comments and suggestions. This research is supported by the MOE (Ministry of Education in China) Youth Foundation Project of Humanities and Social Sciences (No.13YJC790185) , National Natural Science Foundation of China under Grants (Nos.71171035,71471030), China Postdoctoral Science Foundation (Nos.2012M511598, 2013T60710), Liaoning Educational Committee Special Project of Humanities and Social Science Key Research Base(No.ZJ2013039) and Dongbei University of Finance and Economics Project of Excellent Scientific Research Innovative Talents (No.DUFE2014R11).

References

[1] Baltagi B.H. *Testing for fixed effects in logit and probit models using an artificial regression*, Problem 95.5.4, Econometric Theory, **11**, 1179-1179, 1995.

[2] Baltagi B.H. *Specification tests in panel data models using artificial regressions*, Annals of Economics and Statistics, 277-297, 1999.

[3] Bartolucci F., Nigro V. *A dynamic model for binary panel data with unobserved heterogeneity admitting a-$\sqrt{N}$ consistent conditional estimator*, Econometrica, **78**, 719-733, 2010.

[4] Davidson R., MacKinnon J. *Convenient specification tests for logit and probit models*, Journal of Econometrics, **25**, 241-262, 1984.

[5] Gurmu S. *Testing for fixed effects in logit and probit models using an artificial regression*, Solution 95.5.4, Econometric Theory, **12**, 872-874, 1996.

[6] Hsiao C. *Analysis of Panel Data*, (Cambridge University Press, New York, 2003).

[7] Lehmann E.L. *Theory of Point Estimation*, (Wiley and Sons, Inc, 1983).

[8] Mora J, Moro-Egido A.I. *On specification testing of ordered discrete choice models*, Journal of Econometrics, **143**, 191-205, 2008.

[9] Neyman J., Scott E.L. *Consistent estimates based on partially consistent observations*, Econometrica, **16**, 1-32, 1948.

[10] Orme C.D. *The small sample performance of the information matrix test*, Journal of Econometrics, **46**, 309-331, 1990.

[11] Yu G., Gao W., Shi N.-Z. *A note on the estimation problem of dynamic binary panel data model with fixed effects*, Pakistan Journal of Statistics, **28**(2), 271-278, 2012.

ADVICE TO AUTHORS

Hacettepe Journal of Mathematics and Statistics publishes short to medium length research papers and occasional survey articles written in English. All papers are refereed to international standards.

Address for Correspondence

Editorial Office,
Hacettepe Journal of Mathematics and Statistics,
Hacettepe University,
Faculty of Science,
Department of Mathematics,
06532 Beytepe,
Ankara,
Turkey.

E-mail: hjms@hacettepe.edu.tr
Tel: Editors: + 90 312 297 7859
+ 90 312 297 7898
Associate Editor: + 90 312 297 7880
Fax : + 90 312 299 2017
Web Page : http//www.hjms.hacettepe.edu.tr

Advise to Authors : The style of articles should be lucid but concise. The text should be preceded by a short descriptive title and an informative abstract of not more than 100 words. Keywords or phrases and the 2010 AMS Classification should also be included.The main body of the text should be divided into numbered sections with appropriate headings. Items should be numbered in the form 2.4. Lemma, 2.5. Definition. These items should be referred to in the text using the form Lemma 2.4., Definition 2.5. Figures and tables should be incorporated in the text. A numbered caption should be placed above them. References should be punctuated according to the following examples, be listed in alphabetical order according to in the (first) author' s surname, be numbered consecutively and referred to in the text by the same number enclosed in square brackets. Only recognized abbreviations of the names of journals should be used.

[1] Banaschewski, B. *Extensions of topological spaces*, Canad. Math. Bull. **7** (1), 1–22, 1964.
[2] Ehrig, H. and Herrlich, H. *The construct PRO of projection spaces: its internal structure*, in: Categorical methods in Computer Science, Lecture Notes in Computer Science **393** (Springer-Verlag, Berlin, 1989), 286–293.
[3] Hurvich, C. M. and Tsai, C. L. *Regression and time series model selection in small samples*, Biometrika **76** (2), 297–307, 1989.
[4] Papoulis, A. *Probability random variables and stochastic process* (McGraw-Hill, 1965).

INSTRUCTIONS FOR AUTHORS

Preparation of Manuscripts : Manuscripts will be typeset using the LATEX typesetting system. Authors should prepare the article using the HJMS style before submission by e-mail. The style bundle contains the styles for the article and the bibliography as well as example of formatted document as pdf file. The authors should send their paper directly to hjms@hacettepe.edu.tr with the suggestion of area editor and potential reviewers for their submission. The editorial office may not use these suggestions, but this may help to speed up the selection of appropriate reviewers.

Notice : The content of a paper published in this Journal is the sole responsibility of the author or authors, and its publication does not imply the concurrence of the Editors or the Publisher.