HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 43 Issue 6
2014

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 43 Issue 6
December 2014

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 43 Sayı 6 (2014)
ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2014 SAYI : 43 - 6 AY : Aralık
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 43 Issue 6 (2014)

ISSN 1303 – 5010

Published by Hacettepe University
Faculty of Science

CONTENTS

MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 891 – 898

On some variants of compactness

S. Bayhan * and I. L. Reilly †

Abstract

Three weak variants of compactness were introduced and studied by Kohli and Singh [Acta Math. Hungar. 106 (2005), 317-329]. These three properties are reconsidered from the change of topology perspective. In particular, it is shown that each of these properties is equivalent to compactness with respect to another topology on the underlying set. Some consequences of this situation are investigated.

2000 AMS Classification: 54D10, 54D20, 54D30, 54C08, 54C10.

Keywords: compactness, change of topology, d-compact, d^*-compact, D_δ-compact, continuous function, D-continuous function, D^*-continuous function, D_δ-continuous function, D-supercontinuous function, D-irresolute function

Received 12 : 07 : 2013 : Accepted 08 : 11 : 2013 Doi : 10.15672/HJMS.2014437532

1. Introduction

The notion of compactness is one of the most significant topological properties, and its importance reaches well beyond topology to several other branches of mathematics. Weaker variants of compactness have been considered in the topological literature for at least nine decades. For example, Hausdorff almost compact spaces (now known as H-closed spaces) were introduced by Alexandroff and Urysohn [1], and have subsequently been investigated by many researchers. The book [12] is a comprehensive source of references. Almost compactness is considered in the book [3]. Frolik [5] introduced quasicompact spaces. One reason for their significance is that functionally Hausdorff quasicompact spaces are the natural setting for the Stone-Weierstrass theorem, see Stephenson [17].

In 2005 Kohli and Singh [10] introduced three weak variants of compactness which lie between compactness and quasicompactness. They studied the basic properties of the

*Department of Mathematics, Mehmet Akif Ersoy University, 15030, Istiklal Campus, Burdur, Turkey
Email: bayhan@mehmetakif.edu.tr
†Department of Mathematics,University of Auckland,P. B. 92019, Auckland, New Zealand
Email:i.reilly@auckland.ac.nz
The authors gratefully acknowledge financial support of this research by the Scientific and Technological Research Council of Turkey (TUBITAK).

classes of spaces defined by these three weak variants of compactness, which have been named d-compactness, d^*-compactness and D_δ-compactness.

In this paper, we reconsider these three notions from the perspective of change of topology. In each case, we observe that there is an appropriate change of topology which reveals that the new concept is equivalent to the classical notion of compactness. The last two results of Kohli and Singh [10, Theorems 5.17 and 5.18] make this observation almost as a post script. They do not use it anywhere in their paper [10]. In our view, this is the fundamental defining characteristic of these three notions of weak compactness. For us, it is the starting point of the discussion. Each of these three new notions is compactness with respect to another topology on the underlying set. This observation provides the natural setting for the subsequent discussion of each of these three notions.

We are able to exploit this observation to produce more elegant alternative proofs of some of the results of Kohli and Singh [10], and to suggest other results.

Our notation and terminology are standard, see for example Dugundji [4]. In particular, we do not assume any separation properties for the spaces we consider, unless explicitly stated. We denote the interior of a subset B of the topological space (X, τ) by $\tau intB$, or just $intB$, and the closure of B by τclB or clB.

2. Preliminaries and definitions

In a topological space (X, τ) a set B is defined to be regular open if $B = \tau int(\tau clB)$. Since the intersection of two regular open sets is regular open, the collection of all τ regular open sets forms the base for a topology τ_s on X, smaller than τ, called the semi-regularization of (X, τ). Note that (X, τ) is semi-regular if and only if $\tau = \tau_s$.

In 1968 Veličko [18] made the following definition.

Let (X, τ) be a topological space and let $A \subset X$. A point $x \in X$ is called a θ-limit point of $A \subset X$ if every closed neighbourhood of X intersects A. Let θclA denote the set of all θ-limit points of A. The set A is θ-closed if $A = \theta clA$. The complement of a θ-closed set is called a θ-open set. The collection of all θ-open sets in (X, τ) forms a topology on X, denoted by τ_θ.

A subset B of (X, τ) is called a zero-set if there is a continuous real-valued function f defined on X such that $B = \{x \in X : f(x) = 0\}$. The complement of a zero-set is called a co-zero set. The collection of all co-zero sets of (X, τ) is the base for a topology τ_z on X, and $\tau_z \subset \tau$. Moreover, (X, τ) is completely regular if and only if $\tau_z = \tau$.

2.1. Definition A space (X, τ) is said to be
(1) almost compact [3] if every open cover of X has a finite subcollection the closures of whose members cover X;
(2) quasicompact [5] if every cover of X by co-zero sets has a finite subcover;
(3) nearly compact [14] if every open cover of X admits a finite subcollection the interiors of the closures of whose members cover X;
(4) θ-compact [13, Definition 3.19] if every cover of X by θ-open sets has a finite subcover.

The following result was proved by Carnahan [2, Theorem 4.1].

2.2. Theorem (X, τ) is nearly compact if and only if (X, τ_s) is compact.

Definition 2.1 (2) immediately implies that

2.3. Theorem
(1) (X, τ) is quasicompact if and only if (X, τ_z) is compact.
(2) (X, τ) is θ-compact if and only if (X, τ_θ) is compact [13, Remark 4.27].

Following Heldermann [6] we have two definitions.

2.4. Definition A collection β of subsets of a space (X, τ) is called an open complementary system if β consists of open sets such that for every $B \in \beta$, there exist $B_1, B_2, ... \in \beta$ with $B = \cup\{X/B_i : i \in \mathbb{N}\}$.

2.5. Definition A subset U of a space (X, τ) is called a strongly open F_σ-set if there exists a countable open complementary system $\beta(U)$ with $U \in \beta(U)$. The complement of a strongly open F_σ-set is called a strongly closed G_δ-set.

Mack [11] made the next definition in 1970.

2.6. Definition A subset H of a space (X, τ) is called a regular G_δ-set if H is the intersection of a sequence of closed sets whose interiors contain H, i.e., if $H = \bigcap_{n=1}^{\infty} F_n = \bigcap_{n=1}^{\infty} intF_n$, where each F_n is a closed subset of X. The complement of a regular G_δ-set is called a regular F_σ-set.

Kohli and Singh [10] introduced three weak variants of compactness which we now consider from the perspective of change of topology.

2.7. Definition A space (X, τ) is said to be d-compact (d^*-compact, D_δ-compact) if every cover of X by open F_σ-sets (strongly open F_σ-sets, regular F_σ-sets) has a finite subcover.

3. Three topologies

Let (X, τ) be a topological space, and denote by β the collection of all open F_σ-subsets of (X, τ). Now the intersection of two open F_σ-subsets is an open F_σ-subset. Therefore the collection β is a base for a topology on X, which we denote by τ_d. This topology τ_d is called the D-regularization of τ by Kohli and Singh [10].

Similarly, if we replace " open F_σ-subsets " in the paragraph immediately above by " strongly open F_σ-subsets ", we obtain a second topology on X, denoted by τ^*, and called the D-complete regularization of τ in [10].

Yet again, if we replace " open F_σ-subsets " by " regular F_σ-subsets " we obtain a third topology on X, denoted by $\tau^\#$, and called the D_δ-complete regularization of τ by Kohli and Singh [10].

There is an alternative way of defining these three topologies given by the next definition.

3.1. Definition A set G in a topological space (X, τ) is said to be τ_d-open [8] (τ^*-open, $\tau^\#$-open) if for each $x \in G$, there exists an open F_σ-set (strongly open F_σ-set, regular F_σ-set) H such that $x \in H \subset G$. The complement of a τ_d-open (τ^*-open, $\tau^\#$-open) set will be referred to as a τ_d-closed (τ^*-closed, $\tau^\#$-closed) set.

We note that the members of these topologies are denoted by d-open, d^*-open and $d^{\#}$-open sets in [8], [15] and [9] respectively.

4. Change of topology

The fundamental defining characteristic of each of the three weak variants of compactness that we are considering is given by the following result, which is proved immediately from the definitions. Kohli and Singh [10, Theorems 5.17 and 5.18] have made this observation.

4.1. Theorem Let (X,τ) be a topological space. Then

(1) (X,τ) is d-compact if and only if (X,τ_d) is compact,

(2) (X,τ) is d^*-compact if and only if (X,τ^*) is compact,

(3) (X,τ) is D_δ-compact if and only if $(X,\tau^{\#})$ is compact.

Kohli and Singh [10] have provided an impressive list of Examples (2.8 to 2.13) to show all the weak variants shown in their diagram of relationships are distinct. We reproduce their diagram of relationships here as Figure 1.

compact	$\longrightarrow$	nearly compact	$\longrightarrow$	almost compact
$\downarrow$				$\downarrow$
d-compact	$\longrightarrow$	D_δ-compact	$\longleftarrow$	θ-compact
$\downarrow$		$\downarrow$		
d^*-compact	$\longrightarrow$	quasicompact	$\longrightarrow$	pseudocompact

Figure 1.

We take this diagram to mean that one can find a topological space (X,τ) having one of these properties but not one of the stronger properties. For this interpretation of Figure 1 one must regard the topology on X as fixed. Theorems 2.2, 2.3 and 4.1 indicate that six of the concepts in Figure 1 are each separately equivalent to compactness provided an appropriate change of the topology on the underlying set X is made in each case. It seems that the two exceptions are almost compactness and pseudocompactness.

Claims of the kind that "d-compactness is independent of compactness" are confusing. In fact, d-compactness is a disguised form of compactness. It is compactness with respect to another topology on the underlying set. So d-compactness is not a new concept. It is equivalent to the classical notion of compactness, only with respect to a different topology (than the original topology) on the underlying set. The same comments apply to the notions of d^*-compactness and D_δ-compactness.

5. Some Basic Properties

The following definition is due to Kohli and Singh [10, Definition 3.2]

5.1. Definition A topological space (X,τ) said to be D-Hausdorff (D^*-Hausdorff, D_δ-Hausdorff) if each pair of distinct points is contained in disjoint open F_σ-sets (strongly open F_σ-sets, regular F_σ-sets).

The proof of the next result is immediate from Definitions 5.1 and 3.1.

5.2. Proposition Let (X, τ) be a topological space. Then

(1) (X, τ) is D-Hausdorff if and only if (X, τ_d) is Hausdorff.

(2) (X, τ) is D^*-Hausdorff if and only if (X, τ^*) is Hausdorff.

(3) (X, τ) is D_δ-Hausdorff if and only if $(X, \tau^\#)$ is Hausdorff.

One of the well-known standard results of a first course in topology, for example see Dugundji [4, page 224 Theorem 1.4 (2)], is the following proposition.

5.3. Proposition If (X, τ) is a Hausdorff space and A is a compact subset (X, τ), then A is closed in (X, τ).

Applying Theorem 4.1 and Propositions 5.2 and 5.3 we obtain the next result.

5.4. Proposition If A is a d-compact subset of the D-Hausdorff space (X, τ), then A is d-closed in X.

Proof. Observe that A is a compact subset of the Hausdorff space (X, τ_d), so that Proposition 5.3 implies that A is closed in (X, τ_d), or that A is d-closed in (X, τ).

Exactly parallel results can be obtained by analogous proofs for a d^*-compact (D_δ-compact) subset of a D^*-Hausdorff (D_δ-Hausdorff) space. These three results have been proved from first principles by Kohli and Singh [10, Theorem 3.3].

Another standard result concerning compact spaces is that a closed subset of a compact space is compact, for example see Dugundji [4, page 224 Theorem 1.4 (3)]. From this result we obtain the following result which generalizes Theorem 3.10 of [10].

5.5. Proposition Let (X, τ) be a d-compact topological space, and A be τ_d-closed in X. Then A is d-compact.

Again, we can provide exactly parallel results for the other two variants of compactness.

5.6. Proposition Let (X, τ) be a d^*-compact (D_δ-compact) topological space, and A be τ^*-closed ($\tau^\#$-closed) in X. Then A is d^*-compact (D_δ-compact).

It is well-known that compactness can be characterized in terms of the finite intersection property and adherence properties of filters and filterbases, see Dugundji [4, page 223 Theorem 1.3] for example. Kohli and Singh [10, Theorem 4.4] provide a version of these characterizations for d-compactness, d^*-compactness and D_δ-compactness. We note that their definitions of d-adherence and d-convergence of a filterbase T [10, Definitions 4.2 and 4.3] are equivalent to adherence and convergence of T with respect to the topology τ_d. A change of topology approach to this topic is an alternative to the discussion presented in Section 4 of [10].

In order to consider mapping properties we must define appropriate classes of functions between topological spaces.

5.7. Definition A function $f : (X, \tau) \longrightarrow (Y, \sigma)$ is defined to be D-continuous [7] (D^*-continuous [16], D_δ-continuous) if for each point $x \in X$ and each open F_σ-set (strongly open F_σ-set, regular F_σ-set) V containing $f(x)$ there is an open subset U of X such that $x \in U$ and $f(U) \subset V$.

5.8. Definition A function $f:(X,\tau)\longrightarrow(Y,\sigma)$ is defined to be D-supercontinuous [8] (D^*-supercontinuous [15], D_δ-supercontinuous [9]) if for each point $x\in X$ and each open V of X containing $f(x)$ there is an open F_σ-set (strongly open F_σ-set, regular F_σ-set) U in X such that $x\in U$ and $f(U)\subset V$.

5.9. Definition A function $f:(X,\tau)\longrightarrow(Y,\sigma)$ is defined to be D-irresolute (D^*-irresolute, D_δ-irresolute) if for each point $x\in X$ and each open F_σ-set (strongly open F_σ-set, regular F_σ-set) V containing $f(x)$ there is an open F_σ-set (strongly open F_σ-set, regular F_σ-set) U in X such that $x\in U$ and $f(U)\subset V$.

The following results are immediate from the preceding definitions and the discussion in Section 3.

5.10. Proposition Let $f:(X,\tau)\longrightarrow(Y,\sigma)$ be a function between topological spaces. Then

(1) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D-continuous if and only if $f:(X,\tau)\longrightarrow(Y,\sigma_d)$ is continuous.

(2) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D^*-continuous if and only if $f:(X,\tau)\longrightarrow(Y,\sigma^*)$ is continuous.

(3) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D_δ-continuous if and only if $f:(X,\tau)\longrightarrow(Y,\sigma^{\#})$ is continuous.

(4) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D-supercontinuous if and only if $f:(X,\tau_d)\longrightarrow(Y,\sigma)$ is continuous.

(5) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D^*-supercontinuous if and only if $f:(X,\tau^*)\longrightarrow(Y,\sigma)$ is continuous.

(6) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D_δ-supercontinuous if and only if $f:(X,\tau^{\#})\longrightarrow(Y,\sigma)$ is continuous.

(7) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D-irresolute if and only if $f:(X,\tau_d)\longrightarrow(Y,\sigma_d)$ is continuous.

(8) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D^*-irresolute if and only if $f:(X,\tau^*)\longrightarrow(Y,\sigma^*)$ is continuous.

(9) $f:(X,\tau)\longrightarrow(Y,\sigma)$ is D_δ-irresolute if and only if $f:(X,\tau^{\#})\longrightarrow(Y,\sigma^{\#})$ is continuous.

The standard result that compactness is preserved by continuous functions, Proposition 5.10 and Theorem 4.1 can be used to prove the next set of results.

5.11. Proposition Let $f:(X,\tau)\longrightarrow(Y,\sigma)$ be a surjection.

(A) If f is D-continuous (D^*-continuous, D_δ-continuous) and (X,τ) is compact then (Y,σ) is d-compact (d^*-compact, D_δ-compact).

(B) If f is D-supercontinuous (D^*-supercontinuous, D_δ-supercontinuous) and (X,τ) is d-compact (d^*-compact, D_δ-compact) then (Y,σ) is compact.

(C) If f is D-irresolute (D^*-irresolute, D_δ-irresolute) and (X,τ) is d-compact (d^*-compact, D_δ-compact) then (Y,σ) is d-compact (d^*-compact, D_δ-compact).

Proof. We prove one case of each part only. The other cases have exactly similar proofs.

(A) For D-continuity: Now $f : (X, \tau) \longrightarrow (Y, \sigma_d)$ is a continuous surjection and (X, τ) is compact, so that (Y, σ_d) is compact. Thus (Y, σ) is d-compact.

(B) For D^*-supercontinuity: Now $f : (X, \tau^*) \longrightarrow (Y, \sigma)$ is a continuous surjection and (X, τ^*) is compact. Hence (Y, σ) is compact.

(C) For D_δ-irresoluteness: Now $f : (X, \tau^\#) \longrightarrow (Y, \sigma^\#)$ is a continuous surjection and $(X, \tau^\#)$ is compact. Therefore $(Y, \sigma^\#)$ is compact, so that (Y, σ) is D_δ-compact.

Note that 5.11(A) is Theorems 5.2 and 5.3 of Kohli and Singh [10], while 5.11(B) is Theorems 5.5 and 5.6 of [10]. The proofs provided by Kohli and Singh [10] are from first principles, and quite different in character to the proofs given above.

The change of topology approach can be used to suggest new results. To illustrate, we provide two such results.

5.12. Proposition Let $f, g : (X, \tau) \longrightarrow (Y, \sigma)$ be D-irresolute, and (Y, σ) be D-Hausdorff. Then E, the equalizer of f and g, given by $E = \{x \in X : f(x) = g(x)\}$ is d-closed in (X, τ).

Proof. We have that $f, g : (X, \tau_d) \longrightarrow (Y, \sigma_d)$ are continuous, and that (Y, σ_d) is Hausdorff. Thus by a standard result, e.g. Dugundji [4, page 140, 1.5(1)], we have that E is closed in (X, τ_d), so that E is d-closed in (X, τ).

5.13. Proposition If $f : (X, \tau) \longrightarrow (Y, \sigma)$ is D_δ-irresolute and (Y, σ) is D_δ-Hausdorff, then $G(f)$, the graph of f, is closed in $(X \times Y, \tau^\# \times \sigma^\#)$.

Proof. Note that $f : (X, \tau^\#) \longrightarrow (Y, \sigma^\#)$ is continuous, and that $(Y, \sigma^\#)$ is Hausdorff. Then a standard result for continuous functions, e.g. Dugundji [4, page 140, 1.5(3)], implies that $G(f)$ is closed in $(X \times Y, \tau^\# \times \sigma^\#)$.

References

[1] P. S. Alexandroff and P. Urysohn, Zur theorie der topologischen raume, Math. Ann. **92** (1924), 258-266.

[2] D. Carnahan, Locally nearly compact spaces, Boll. U. M. I. **4** (1972), 146-153.

[3] A. Császár, General Topology, Adam Hilger Ltd. Bristol, 1978.

[4] J. Dugundji, Topology, Allyn and Bacon, Boston, Mass. 1966.

[5] Z. Frolik, Generalizations of compact and Lindelöf spaces, Czech. Math. J. **13** (1959), 172-217.

[6] N. C. Helderman, Developability and some new regularity axioms, Can. J. Math., **33** (1981), 641-663.

[7] J. K. Kohli, D-continuous functions, D-regular spaces and D-Hausdorff spaces, Bull. Cal. Math. Soc., **84** (1992), 39-46.

[8] J. K. Kohli and D. Singh, D-supercontinuous functions, Indian J. Pure Appl. Math., **32** (2001), 227-235.

[9] J. K. Kohli and D. Singh, D_δ-supercontinuous functions, Indian J. Pure Appl. Math., **34** (2003), 1089-1100.

[10] J. K. Kohli and D. Singh, Between compactness and quasicompactness, Acta Math. Hungar., **106** (2005), 317-329.

[11] J. Mack, Countable paracompactness and weak normality properties, Trans. Amer. Math. Soc., **148** (1970), 265-272.

[12] J. R. Porter and R. G. Woods, Extensions and Absolutes of Hausdorff spaces, Springer-Verlag, New York, 1978.

[13] M. Saleh, On θ-closed sets and some forms of continuity, Archivum Mathematicum (BRNO), **40** (2004), 383-393.

[14] M. K. Singal and A. Mather, On nearly-compact spaces, Boll. U. M. I., **2** (1969), 702-710.

[15] D. Singh, D^*-supercontinuous functions, Bull. Cal. Math. Soc., **94** (2002), 67-76.

[16] D. Singh, D^*-continuous functions, Bull. Cal. Math. Soc., **91** (1999), 385-390.

[17] R. M. Stephenson, Jr., Spaces for which the Stone-Weierstrass theorem holds, Trans. Amer. Math. Soc., **133** (1968), 537-546.

[18] N. Veličko, H-closed topological spaces, Amer. Math. Soc. Transl., **78** (2) (1968), 103-118.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 899 – 913

Minimality over free monoid presentations

A. Sinan Cevik* , Kinkar Ch. Das † , I. Naci Cangul ‡ and A. Dilek Maden§

Abstract

As a continues study of the paper [4], in here, we first state and prove the p-Cockcroft property (or, equivalently, efficiency) for a presentation, say $\mathcal{P}_E$, of the semi-direct product of a free abelian monoid rank two by a finite cyclic monoid. Then, in a separate section, we present sufficient conditions on a special case for $\mathcal{P}_E$ to be minimal whilst it is inefficient.

2000 AMS Classification: 20L05, 20M05, 20M15, 20M50.

Keywords: Minimality, Efficiency, p-Cockcroft property.

Received 18 : 10 : 2012 : Accepted 04 : 01 : 2013 Doi : 10.15672/HJMS.2014437522

1. Preliminaries

Suppose that $\mathcal{P} = [X; \mathbf{r}]$ is a finite presentation for a monoid M. Then the *Euler characteristic* is defined by $\chi(\mathcal{P}) = 1 - |X| + |\mathbf{r}|$. There also exists an upper bound over M which is defined by $\delta(M) = 1 - rk_{\mathbb{Z}}(H_1(M)) + d(H_2(M))$. In fact, as depicted in [2, 3, 4], S. Pride has shown that $\chi(\mathcal{P}) \geq \delta(M)$. With this background, we define the monoid presentation $\mathcal{P}$ to be *efficient* if $\chi(\mathcal{P}) = \delta(M)$ and then M is called *efficient* if it has an efficient presentation. Moreover a presentation $\mathcal{P}_0$ for M is called *minimal* if $\chi(\mathcal{P}_0) \leq \chi(\mathcal{P})$, for all presentations $\mathcal{P}$ of M. There is also interest in finding *inefficient* finitely presented monoids since if we can find a minimal presentation $\mathcal{P}_0$ for a monoid M such that $\mathcal{P}_0$ is not efficient then we have $\chi(\mathcal{P}') \geq \chi(\mathcal{P}_0) > \delta(M)$, for all presentations

*Department of Mathematics, Selcuk University, Faculty of Science, Campus, 42075, Konya-Turkey, Corrosponding Author.
Email: `sinan.cevik@selcuk.edu.tr`
†Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea, Email:`kinkar2003@googlemail.com`
‡Department of Mathematics, Faculty of Art and Science, Uludag University, Gorukle Campus, 16059, Bursa-Turkey,
Email:`ncangul@gmail.com`
§ Email:`drdilekgungor@gmail.com`
This paper is partially supported by TUBITAK and the Scientific Research Center (BAP) of Selcuk and Uludag Universities. Also some parts of this work has been prepared during K.Ch. Das's visit in Selcuk and Uludag Universities.

$\mathcal{P}'$ defining the same monoid M. Thus there is no efficient presentation for M, that is, M is not an efficient monoid.

Some of the fundamental material (for instance, *semi-direct products of monoids*, *Squier complex, a trivializer set* of the Squier complex, *p-Cockcroft property, monoid pictures*) which will be needed to construct the main results of this paper have been defined and referenced in detail in [1, 2, 3, 4].

The following theorem also proved by S. Pride which we will use it rather than making more direct computations of homology for monoids. In fact Kilgour and Pride showed the analogous result for groups in [8] and credit an earlier proof by Epstein ([5]).

1.1. Proposition. *Let $\mathcal{P}$ be a monoid presentation. Then $\mathcal{P}$ is efficient if and only if it is p-Cockcroft for some prime p.*

Let A and K be arbitrary monoids with associated presentations $\mathcal{P}_A = [X\,;\,\mathbf{r}]$ and $\mathcal{P}_K = [Y\,;\,\mathbf{s}]$, respectively. Also let $E = K \rtimes_\theta A$ be the corresponding semi-direct product of these two monoids. For every $x \in X$ and $y \in Y$, choose a word, which we denote by $y\theta_x$, on Y such that $[y\theta_x] = [y]\theta_{[x]}$ as an element of K. To establish notation, let us denote the relation $yx = x(y\theta_x)$ on $X \cup Y$ by T_{yx} and write $\mathbf{t}$ for the set of relations T_{yx}. Then, for any choice of the words $y\theta_x$,

$$\mathcal{P}_E = [Y,\, X\,;\, \mathbf{s},\, \mathbf{r},\, \mathbf{t}] \tag{1.1}$$

is a standard monoid presentation for the semi-direct product E. Then a trivializer set, $\mathbf{X_E}$, of the Squier complex $\mathcal{D}(\mathcal{P}_E)$ has been defined in [10] by J. Wang as the set

$$\mathbf{X_A} \cup \mathbf{X}_K \cup \mathbf{C}_1 \cup \mathbf{C}_2$$

(see also [4, Lemma 1.5]) where $\mathbf{X_A}$ and $\mathbf{X_K}$ are the trivializers of the Squier complexes $\mathcal{D}(\mathcal{P}_A)$ and $\mathcal{D}(\mathcal{P}_K)$, and also the subsets $\mathbf{C}_1$, $\mathbf{C}_2$ consist of the generating monoid pictures $\mathbb{P}_{S,x}$ $(S \in \mathbf{s},\ x \in X)$ and $\mathbb{P}_{R,y}$ $(R \in \mathbf{r},\ y \in Y)$. Hence, by using the set $\mathbf{X_E}$, Çevik proved the following result which will be used to proof of Theorem 2.4 below.

1.2. Theorem. [3, Theorem 3.1] *Let p be a prime or 0. Then the presentation $\mathcal{P}_E$ in (1.1) is p-Cockcroft if and only if the following conditions hold.*

- (*i*) *$\mathcal{P}_A$ and $\mathcal{P}_K$ are p-Cockcroft,*
- (*ii*) *$\exp_y(S) \equiv 0 \ (mod\ p)$ for all $S \in \mathbf{s}$, $y \in Y$,*
- (*iii*) *$\exp_S(\mathbb{B}_{S,x}) \equiv 1 \ (mod\ p)$ for all $S \in \mathbf{s}$, $x \in X$,*
- (*iv*) *$\exp_S(\mathbb{C}_{y,\theta_R}) \equiv 0 \ \ (mod\ p)$ for all $S \in \mathbf{s}$, $y \in Y$, $R \in \mathbf{r}$,*
- (*v*) *$\exp_{T_{yx}}(\mathbb{A}_{R_+,y}) \equiv \exp_{T_{yx}}(\mathbb{A}_{R_-,y}) \ \ (mod\ p)$ for all $R \in \mathbf{r}$, $y \in Y$ and $x \in X$.*

This paper has been divided into two main parts. In Section 2, we will investigate the efficiency (in fact, by Proposition 1.1, p-Cockcroft property for a prime p) for a standard presentation of the semi-direct product E of a free abelian monoid rank two, say K_2, by a finite cyclic monoid, say A, (see Theorem 2.4 below). Moreover, in Section 3, we will present the minimality of the monoid E while it has an inefficient presentation (see Theorem 3.1 below) by considering a special case.

2. Efficiency

2.1. The semi-direct product of K_2 by A. By the definition, to define a semi-direct product of K_2 by an arbitrary monoid A, we first need to define an endomorphism of K_2. To do that, let us start with $\mathbb{Z}^{+^n}$ which is the free abelian monoid rank n, say K_n. Also let $\mathcal{M}$ be an $n \times n$-matrix with non-negative integer entries. Then we get a mapping

$$\psi_{\mathcal{M}} : K_n \longrightarrow K_n,\ \ v \longmapsto v\mathcal{M},$$

where $v = (v_1, v_2, \cdots, v_n)$. Actually $\psi_{\mathcal{M}} \in End(K_n)$ (and so $\psi_{\mathcal{M}_1}\psi_{\mathcal{M}_2} = \psi_{\mathcal{M}_1\mathcal{M}_2}$). We note that if $\phi \in End(K_n)$ then there exist a matrix $\mathcal{M}$ (depending on ϕ) such that $\phi = \psi_{\mathcal{M}}$. By the mapping $\mathcal{M} \longmapsto \psi_{\mathcal{M}}$, we get an isomorphism from $Mat_n(\mathbb{Z}^+)$ to the monoid $End(K_n)$, where

$$Mat_n(\mathbb{Z}^+) = \{\mathcal{M} \ : \ \mathcal{M} \text{ is an } n \times n\text{-matrix with non-negative integer entries}\}$$

is a monoid under matrix multiplication.

Suppose $\mathcal{P}_{K_n} = [y_i \ (1 \leq i \leq n) \ ; \ y_iy_j = y_jy_i \ (1 \leq i < j \leq n)]$ is a presentation for K_n and $\mathcal{P}_A = [\mathbf{x} \ ; \ \mathbf{r}]$ is a presentation for A. Suppose also that, for each $x \in \mathbf{x}$, we have an endomorphism ψ_x of K. Since $End(K_n) \cong Mat_n(\mathbb{Z}^+)$, the endomorphism ψ_x $(x \in \mathbf{x})$ will be $\psi_{\mathcal{M}_x}$ for some matrix $\mathcal{M}_x$. For any positive word $W = x_1x_2\cdots x_n$ on $\mathbf{x}$, let $\mathcal{M}_W$ be the product $\mathcal{M}_{x_1}\mathcal{M}_{x_2}\cdots\mathcal{M}_{x_n}$ of the matrices $\mathcal{M}_{x_i}$, where $1 \leq i \leq n$. Then the mapping $x \longmapsto \psi_x$ $(x \in \mathbf{x})$ induces a homomorphism $\theta : A \longrightarrow End(K_n)$ if and only if $\mathcal{M}_{R_+} = \mathcal{M}_{R_-}$, for all $R \in \mathbf{r}$.

Now let A be the finite cyclic monoid with a presentation $\mathcal{P}_A = [x \ ; \ x^k = x^l]$ where $1 \leq l < k$ and $l, k \in \mathbb{Z}^+$. (We note that the fundamental material about finite cyclic monoids can be found in the book [6]).

2.1. Remark. Recall that the elements of the finite cyclic monoid A represented by equivalence classes $[x^i]$ $(0 \leq i \leq k)$. For $0 \leq i \leq l$, the equivalence class $[x^i]$ just consist of the single element x^i. However for $i \geq l$, the equivalence class $[x^i]$ consist of infinitely many elements which are defined by $[x^i] = \{x^{i+q(k-l)} \ ; \ q = 0, 1, 2, \cdots\}$.

Also let us consider K_2 and let us suppose that ψ is the endomorphism $\psi_{\mathcal{M}}$ of K_2, where

$$\mathcal{M} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$$

such that the entries α_{ij}'s are the positive integers given by

$$[y_1] \longmapsto [{y_1}^{\alpha_{11}}{y_2}^{\alpha_{12}}] \ \text{ and } \ [y_2] \longmapsto [{y_1}^{\alpha_{21}}{y_2}^{\alpha_{22}}].$$

Hence, by the previous explanation, the mapping $x \longmapsto \psi_x$ $(x \in \mathbf{x})$ induces a well-defined monoid homomorphism $\theta : A \longrightarrow End(K_2)$ if and only if $\mathcal{M}_{[x^k]} = \mathcal{M}_{[x^l]}$, or equivalently,

$$(2.1) \qquad \mathcal{M}^k \equiv \mathcal{M}^l \mod d,$$

where $d \mid (k - l)$.

2.2. Remark. By considering the elements of finite cyclic monoid A with its presentation $\mathcal{P}_A$ as defined in Remark *2.1*, there exits an inequality between the non-negative integers k and l such as $1 \leq l < k$. Thus to define an induces homomorphism $\theta : A \longrightarrow End(K_2)$, that is, to be able to define $K_2 \rtimes_\theta A$, we must take congruence relation between $\mathcal{M}^k$ and $\mathcal{M}^l$ as given in *(2.1)* with the assumption $d \mid (k - l)$.

In fact the kth and lth powers of the matrices can be written as follows. Initially, let us consider the matrices

$$\mathcal{M}^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ \text{ and } \ \mathcal{M}^1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix},$$

and then, for simplicity, let us rewrite them as the matrices

$$\begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \ \text{ and } \ \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix},$$

respectively. Then we clearly get

$$\mathcal{M}^2 = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} A_1\alpha_{11} + B_1\alpha_{21} & A_1\alpha_{12} + B_1\alpha_{22} \\ C_1\alpha_{11} + D_1\alpha_{21} & C_1\alpha_{12} + D_1\alpha_{22} \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}, \text{say.}$$

Therefore the kth ($k \in \mathbb{Z}^+$) power of $\mathcal{M}$ will be

$$\begin{aligned} \mathcal{M}^k &= \begin{bmatrix} A_{k-1} & B_{k-1} \\ C_{k-1} & D_{k-1} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{k-1}\alpha_{11} + B_{k-1}\alpha_{21} & A_{k-1}\alpha_{12} + B_{k-1}\alpha_{22} \\ C_{k-1}\alpha_{11} + D_{k-1}\alpha_{21} & C_{k-1}\alpha_{12} + D_{k-1}\alpha_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}, \text{say.} \end{aligned}$$

As a similar idea, the lth ($l \in \mathbb{Z}^+$) power of $\mathcal{M}$ will be

$$\mathcal{M}^l = \begin{bmatrix} A_l & B_l \\ C_l & D_l \end{bmatrix}.$$

Now we can present the following lemma which gives the importance of Equation (2.1). In fact this lemma will be needed in the proof of Theorem 2.4 below.

2.3. Lemma. *The function $\theta : A \longrightarrow End(K_2)$ defined by $[x] \longmapsto \theta_{[x]}$ is a well-defined monoid homomorphism if and only if $A_k \equiv A_l \mod d$, $B_k \equiv B_l \mod d$, $C_k \equiv C_l \mod d$ and $D_k \equiv D_l \mod d$, where $d \mid (k - l)$.*

Proof. This follows immediately from $\mathcal{M}^k \equiv \mathcal{M}^l \mod d$. □

Now suppose that (2.1) holds. Then, by Lemma 2.3, we obtain a semi-direct product $E = K_2 \rtimes_\theta A$ and have a presentation

(2.2) $\qquad \mathcal{P}_E = [y_1, y_2, x \ ; \ S, R, T_{y_1x}, T_{y_2x}]$,

as in (1.1), for the monoid E where

$$S : y_1y_2 = y_2y_1, \qquad R : x^k = x^l$$
$$T_{y_1x} : y_1x = xy_1^{\alpha_{11}}y_2^{\alpha_{12}}, \qquad T_{y_2x} : y_2x = xy_1^{\alpha_{21}}y_2^{\alpha_{22}},$$

respectively.

At the rest of this paper, we will assume that Equality (2.1) always holds when we talk about the semi-direct product E of K_2 by A.

We know that the trivializer set of $\mathbf{X_E}$ of $\mathcal{D}(\mathcal{P}_E)$ consists of the trivializer set $\mathbf{X_{K_2}}$ of $\mathcal{D}(\mathcal{P}_{K_2})$, $\mathbf{X_A}$ of $\mathcal{D}(\mathcal{P}_A)$ and the sets $\mathbf{C}_1$, $\mathbf{C}_2$ (see [4, Lemma 1.5]). In our case, $\mathbf{X_{K_2}}$ is equal to the empty set since, for the relator S, we have $\iota(S_+) \neq \iota(S_-)$ (or, equivalently, $\tau(S_+) \neq \tau(S_-)$) and so, by [7], $\mathcal{P}_{K_2}$ is aspherical then p-Cockcroft for any prime p. Newertheless, the trivializer set $\mathbf{X_A}$ of the Squier complex $\mathcal{D}(\mathcal{P}_A)$ is defined as in Figure 1 (cf. [3, Lemma 4.4]).

Finally the subsets $\mathbf{C}_1$ and $\mathbf{C}_2$ contain the generating monoid pictures $\mathbb{P}_{S,x}$ (which contains a non-spherical subpicture $\mathbb{B}_{S,x}$ as depicted in [3]), $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$ of the trivializer set $\mathbf{X_E}$. These pictures can be presented as in Figure 2-(a) and (b).

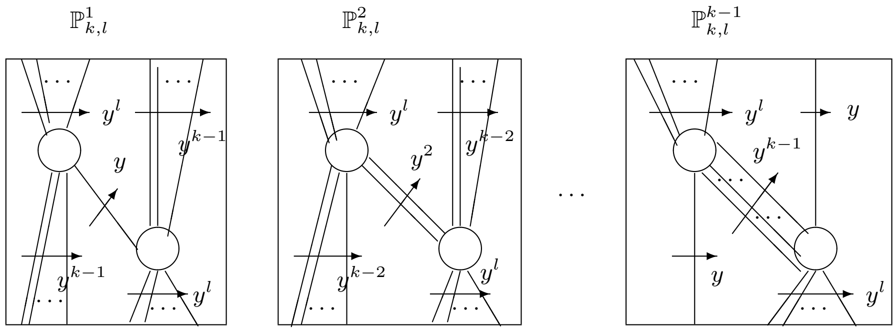
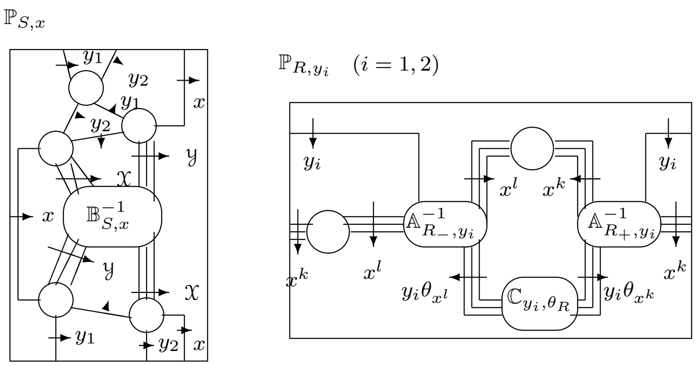


Figure 1. : Generating pictures of finite monogenic monoids



(a) (b)

Figure 2. : In the figure (a), $\mathcal{X} = y_1^{\alpha_{21}} y_2^{\alpha_{22}}$ and $\mathcal{Y} = y_1^{\alpha_{11}} y_2^{\alpha_{12}}$

2.2. The main theorem and its proof. For simplicity, let us replace the sum of coefficients

$$\left.\begin{array}{ll} A_0 + A_1 + \cdots + A_{k-1} \text{ as } \mathcal{A}_k \ , & A_0 + A_1 + \cdots + A_{l-1} \text{ as } \mathcal{A}_l \ , \\ B_0 + B_1 + \cdots + B_{k-1} \text{ as } \mathcal{B}_k \ , & B_0 + B_1 + \cdots + B_{l-1} \text{ as } \mathcal{B}_l \ , \\ C_0 + C_1 + \cdots + C_{k-1} \text{ as } \mathcal{C}_k \ , & C_0 + C_1 + \cdots + C_{l-1} \text{ as } \mathcal{C}_l \ , \\ D_0 + D_1 + \cdots + D_{k-1} \text{ as } \mathcal{D}_k \ , & D_0 + D_1 + \cdots + D_{l-1} \text{ as } \mathcal{D}_l \ . \end{array}\right\} \qquad (2.3)$$

Suppose that the positive integer d, defined in (2.1), is equal to a prime p such that $p \mid (k-l)$. Therefore the first main theorem of this paper can be given as in the following.

2.4. Theorem. *Let p be a prime or 0, and consider the replacements in* (2.3)*. Then the presentation $\mathcal{P}_E$, as in* (2.2)*, for the monoid $E = K_2 \rtimes_\theta A$ is p-Cockcroft if and only if*

a) $det\mathcal{M} \equiv 1 \mod p$,

b) $\mathcal{A}_k \equiv \mathcal{D}_l \mod p \ , \quad \mathcal{B}_k \equiv \mathcal{C}_l \mod p \ ,$
$\mathcal{C}_k \equiv \mathcal{B}_l \mod p \ , \quad \mathcal{D}_k \equiv \mathcal{A}_l \mod p \ .$

Proof. The proof will be given by checking the conditions of Theorem 1.2. By a part of prelimary material of this paper, it is clear that $\mathbf{X_{K_2}} = \emptyset$. Also, since the trivializer set $\mathbf{X_A}$ of the Squier complex $\mathcal{D}(\mathcal{P}_A)$ can be defined as in Figure 1, it is clear that $\mathcal{P}_A$ is p-Cockcroft (in fact Cockcroft). Moreover, by considering the picture $\mathbb{P}_{S,x}$ in Figure 2-(a), we see that $exp_{T_{y_1 x}}(\mathbb{P}_{S,x}) = 0 = exp_{T_{y_2 x}}(\mathbb{P}_{S,x})$ which is clear by $exp_{y_1}(S) = 0 = exp_{y_2}(S)$. Thus the conditions (i) and (ii) of Theorem 1.2 hold. Furthermore in the picture $\mathbb{B}_{S,x}$, we actually have $\alpha_{11}\,\alpha_{12}$-times positive and $\alpha_{12}\,\alpha_{21}$-times negative S-discs. Thus

$$exp_S(\mathbb{B}_{S,x}) = \alpha_{11}\,\alpha_{12} - \alpha_{12}\,\alpha_{21} = det\mathcal{M}.$$

So to condition (iii) be hold, we must have $det\mathcal{M} \equiv 1 \mod p$, as required.

Let us consider the generating pictures $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$ as drawn in Figure 2-(b). We always have $exp_R(\mathbb{P}_{R,y_1}) = 0 = exp_R(\mathbb{P}_{R,y_2})$. Recall that to define a semi-direct product $K_2 \rtimes_\theta A$, we assumed equality (2.1) be held. That means, for each $i \in \{1,2\}$, we must have

$$y_i\theta_{[x^k]} = y_i\theta_{[x^l]}.$$

But we know that this equality be hold if and only if the conditions in Lemma 2.3 are satisfied. Besides of that using the equality of the congruence classes gives us that there will be no $\mathbb{C}_{y_i,\theta_R}$ subpictures. In other words, all arcs in that part will be coincides to each other. So the condition (iv) will be directly held. Let us now consider the subpictures $\mathbb{A}_{R_+,y_i}$ and $\mathbb{A}_{R_-,y_i}$ which consist of only T_{y_ix} discs $(1 \leq i \leq 2)$. Since each of the generating pictures $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$ contains a single subpicture $\mathbb{A}_{R_+,y_i}$ and a single subpicture $\mathbb{A}^{-1}_{R_-,y_i}$, we must have

$$exp_{y_i}(\mathbb{A}_{R_+,y_i}) - exp_{y_i}(\mathbb{A}_{R_-,y_i}) = exp_{y_i}(\mathbb{P}_{R,y_i}).$$

Now let us take into account the matrices $\mathcal{M}^0, \mathcal{M}^1, \cdots, \mathcal{M}^{k-1}$. By using the endomorphism $\psi_\mathcal{M}$ of K defined by $[y_1] \longmapsto [{y_1}^{\alpha_{11}}{y_2}^{\alpha_{12}}]$ and $[y_2] \longmapsto [{y_1}^{\alpha_{21}}{y_2}^{\alpha_{22}}]$, a simple calculation shows that the sum of the first row and first column elements in these matrices gives the exponent sum of the T_{y_1x} discs in the subpicture $\mathbb{A}_{R_+,y_1}$. In other words

$$\mathcal{A}_k = exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_1}).$$

Similarly, we also get

$$\mathcal{B}_k = exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_1}), \mathcal{C}_k = exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_2}) \text{ and } \mathcal{D}_k = exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_2}).$$

On the other hand, again by considering the matrices $\mathcal{M}^0, \mathcal{M}^1, \cdots, \mathcal{M}^{l-1}$ with the same idea as above, we obtain

$$\mathcal{A}_l = exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_2})\,, \quad \mathcal{B}_l = exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_2})\,,$$
$$\mathcal{C}_l = exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_1})\,, \quad \mathcal{D}_l = exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_1})\,.$$

Therefore to p-Cockcroft property be hold, we need

$$exp_{T_{y_ix}}(\mathbb{A}_{R_+,y_i}) \equiv exp_{T_{y_ix}}(\mathbb{A}_{R_-,y_i}) \mod p,$$

for all $1 \leq i \leq 2$.

Conversely let the two conditions $a)$ and $b)$ of the theorem be hold. Then, by using the trivializer of the Squier complex $\mathcal{D}(\mathcal{P}_E)$, we can easily see that $\mathcal{P}_E$ is p-Cockcroft where p is a prime or 0.

Hence the result. □

2.5. Remark. The importance of the assumption $p \mid (k-l)$ seems much clear in the proof of Theorem *2.4*. Otherwise we could not have obtained Equality *(2.1)* and so could not have obtained the exponent sums of the disc T_{y_1x} and T_{y_2x} congruent to zero by modulo p in the subpictures $\mathbb{A}_{R_+,y_i}$ and $\mathbb{A}_{R_-,y_i}$, where $i \in \{1,2\}$, since these sums are directly related to the number of k-arcs and l-arcs, respectively.

2.3. Some applications.

2.6. Example. Let p be an odd prime and suppose that

$$\mathcal{M} = \begin{bmatrix} 1 & \alpha_{12} \\ 0 & 1 \end{bmatrix} \tag{2.4}$$

is a matrix representation for the endomorphism of free abelian monoid K_2 rank two. We then always have

$$\mathcal{M}^{p+1} \equiv \mathcal{M}^1 \mod p$$

and, by Lemma *2.3*, we also have $E = K_2 \rtimes_\theta A$. Hence we get a presentation

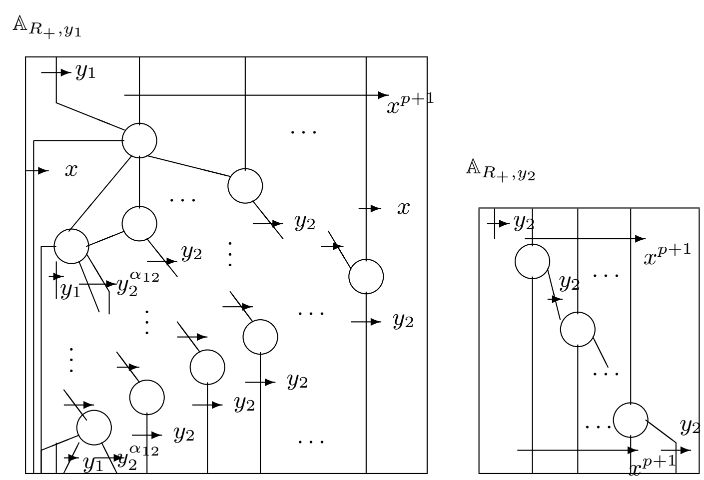
$$\mathcal{P}_E = [y_1, y_2, x \; ; \; y_1y_2 = y_2y_1, x^{p+1} = x, y_1x = xy_1y_2^{\alpha_{12}}, y_2x = xy_2], \tag{2.5}$$

as in *(2.2)*, for the monoid E.

Therefore we can give the following result as a consequence of Theorem 2.4.

2.7. Corollary. *For all odd prime p, the semi-direct product presentation $\mathcal{P}_E$ in* (2.5) *always p-Cockcroft.*

Proof. By considering the subpictures $\mathbb{A}_{R_+,y_1}$, $\mathbb{A}_{R_+,y_2}$, $\mathbb{A}_{R_-,y_1}$ and $\mathbb{A}_{R_-,y_2}$ given in Figures 3 and 4, the proof will be an easy application of Theorem 2.4. In fact the condition



(a) (b)

Figure 3

a) of Theorem 2.4 always holds since $det\mathcal{M} = 1$. Moreover we have

$$\begin{aligned} exp_{T_{y_1x}}(\mathbb{P}_{R,y_1}) &= exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_1}) - exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_1}) \\ &= A_0 + A_1 + \cdots + A_p - D_0 = (p+1) - 1 = p, \end{aligned}$$

which is obviously congruent to zero by modulo p, and

$$\begin{aligned} exp_{T_{y_2x}}(\mathbb{P}_{R,y_1}) &= exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_1}) - exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_1}) \\ &= B_0 + B_1 + \cdots + B_p - C_0 \\ &= \alpha_{12}\,\frac{p(p+1)}{2} - 0 = \alpha_{12}\,\frac{p(p+1)}{2} \equiv 0 \mod p. \end{aligned}$$

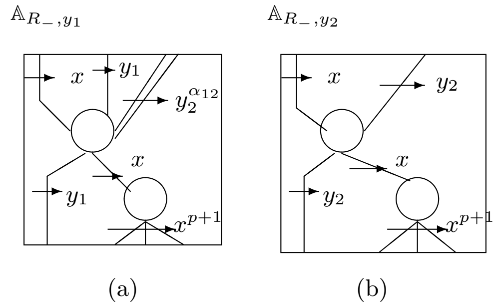


Figure 4

Similarly,

$$\begin{aligned} exp_{T_{y_1x}}(\mathbb{P}_{R,y_2}) &= exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_2}) - exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_2}) \\ &= C_0 + C_1 + \cdots + C_p - B_0 \equiv 0 \mod p, \\ exp_{T_{y_2x}}(\mathbb{P}_{R,y_2}) &= exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_2}) - exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_2}) \\ &= D_0 + D_1 + \cdots + D_p - A_0 = (p+1) - 1 = p \equiv 0 \mod p. \end{aligned}$$

Therefore, for all $i \in \{1,2\}$, $exp_{T_{y_ix}}(\mathbb{P}_{R,y_i}) \equiv 0 \mod p$. (We note that, by the explanation as in the proof of Theorem 2.4, we do not have $\mathbb{C}_{y_i,\theta_R}$ subpictures in $\mathbb{P}_{R,y_i}$). This completes the proof. □

2.8. Remark. In Example *2.6*, if we constructed the matrix $\mathcal{M}$, defined in *(2.4)*, for even prime p while $x^{p+1} = x$ then, by Lemma *2.3*, we would obtain a semi-direct product E for just $\alpha_{12} = 1$ or $\alpha_{12} = 0$ while $\mathcal{M}^3 \equiv \mathcal{M} \mod p$. However, for $\alpha_{12} = 1$, since

$$B_0 + B_1 + B_2 \neq C_0,$$

by Theorem *2.4*, the presentation $\mathcal{P}_E$ in *(2.5)* will be inefficient. Here, by Theorem *2.4*, one can show that $\mathcal{P}_E$ is efficient if and only if $\alpha_{12} = 0$. But $\alpha_{12} = 0$ gives the homomorphism θ is identitiy and so, $K_2 \rtimes_\theta A$ becomes $K_2 \times A$. In fact the efficiency for a presentation of the direct product of arbitrary two monoids has been investigated in [3, Theorem 4.1].

A similar case, as in Example 2.6, can be given by using the matrix

$$\mathcal{M} = \begin{bmatrix} 1 & 0 \\ \alpha_{21} & 1 \end{bmatrix}.$$

Then we obtain a semi-direct product E with a presentation

(2.6) $\qquad \mathcal{P}_E = [y_1, y_2, x\ ;\ y_1y_2 = y_2y_1, x^{p+1} = x, y_1x = xy_1, y_2x = xy_1^{\alpha_{21}}y_2].$

Thus we have the following result, as a consequence of Theorem 2.4, which can be proved quite similarly as in Corollary 2.7.

2.9. Corollary. *Let $\mathcal{P}_E$, as in* (2.6)*, be a presentation for the semi-direct product of K_2 by A. Then, for all odd prime p, $\mathcal{P}_E$ is p-Cockcroft.*

We note that Remark 2.8 is also valid for the above case.

2.10. Example. Suppose that p is a prime and the matrix $\mathcal{M}$ is equal to either $\begin{bmatrix} 1 & \alpha_{12} \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 \\ \alpha_{21} & 1 \end{bmatrix}$. Then, by applying a simple calculation as in the previous examples, we get an efficient semi-direct product presentation for $k = 2p+1$ and $l = 1$.

2.11. Example. Let p be any prime and let $\mathcal{M} = \begin{bmatrix} 1 & 0 \\ 0 & \alpha_{22} \end{bmatrix}$. Hence we get $\mathcal{M}^{2p+1} \equiv \mathcal{M} \mod p$ and, by Lemma *2.3*, we have a semi-direct product $E = K_2 \rtimes_\theta A$ with a presentation

$$\mathcal{P}_E = [y_1, y_2, x \ ; \ y_1y_2 = y_2y_1, x^{2p+1} = x, y_1x = xy_1, y_2x = xy_2^{\alpha_{22}}]. \tag{2.7}$$

As an application of Theorem 2.4, we also have the following corollary.

2.12. Corollary. *The presentation $\mathcal{P}_E$, as in* (2.7)*, is p-Cockcroft for all prime p, if $\alpha_{22} = 1 + pt$ where $t > 0$.*

Proof. In the proof, we will assume $\alpha_{22} = 1 + pt$, $t > 0$, and then just follow the same way as in the proof of Corollary 2.7. It is clear that $det\mathcal{M} \equiv 1 \pmod p$ by the assumption on α_{22}. So the condition *a*) in Theorem 2.4 holds. Now let us consider the subpictures $\mathbb{A}_{R_+,y_1}$, $\mathbb{A}_{R_+,y_2}$, $\mathbb{A}_{R_-,y_1}$ and $\mathbb{A}_{R_-,y_2}$ given in Figure 5. We note that, by fixing these subpictures into the pictures $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$ given in Figure 2-(b), we obtain similar $\mathbb{P}_{R,y_i}$ $(1 \leq i \leq 2)$ pictures for this case. Then we have

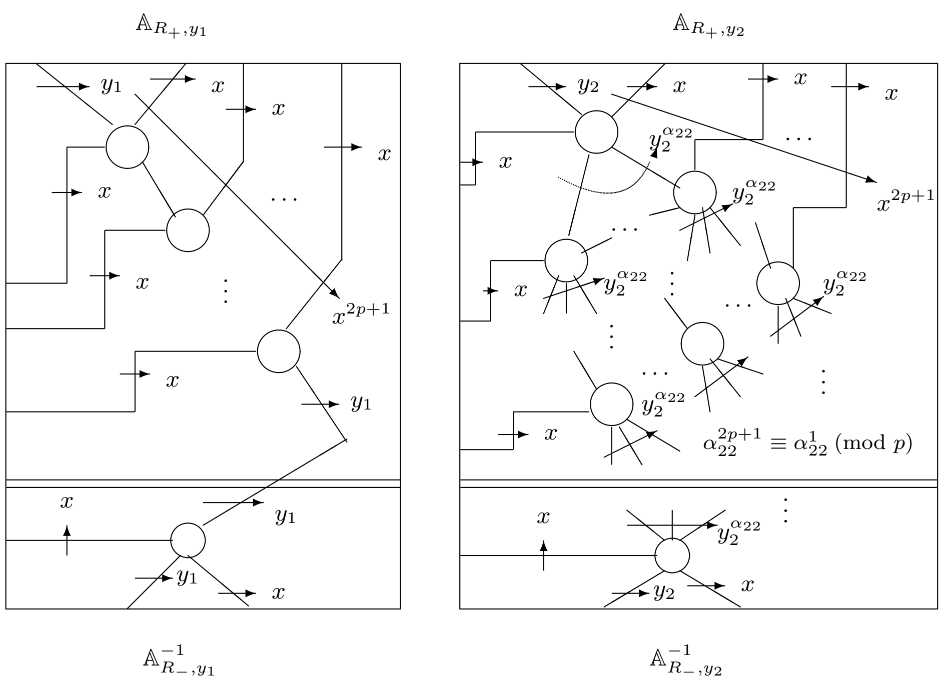


Figure 5

$$\begin{aligned} exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_1}) - exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_1}) &= (2p+1) - 1 = 2p \equiv 0 \mod p, \\ exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_1}) - exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_1}) &\equiv 0 \mod p \quad \text{and} \\ exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_2}) - exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_2}) &\equiv 0 \mod p. \end{aligned}$$

Furthermore, since

$$\begin{aligned} exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_2}) - exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_2}) &= 1 + \alpha_{22} + {\alpha_{22}}^2 + \cdots + {\alpha_{22}}^{2p} - 1 \\ &= \frac{{\alpha_{22}}^{2p+1} - 1}{\alpha_{22} - 1} - 1 \\ &= \frac{{\alpha_{22}}^{2p+1} - \alpha_{22}}{\alpha_{22} - 1} \equiv 0 \mod p, \end{aligned}$$

the condition $b)$ of Theorem 2.4 holds.

We should note that $\mathcal{M}^{2p+1} \equiv \mathcal{M} \pmod p$ implies ${\alpha_{22}}^{2p+1} \equiv \alpha_{22} \pmod p$ and this gives us that $\tau(\mathbb{A}_{R_+,y_2}) = \iota(\mathbb{A}^{-1}_{R_+,y_2})$, that is, there is no subpicture $\mathbb{C}_{y_2,\theta_R}$ in the picture $\mathbb{P}_{R,y_2}$ as expressed in the proof of Theorem 2.4. □

By choosing

$$\mathcal{M} = \begin{bmatrix} \alpha_{11} & 0 \\ 0 & 1 \end{bmatrix},$$

for any prime p, we get again $\mathcal{M}^{2p+1} \equiv \mathcal{M} \mod p$ as in Example 2.11, and so we obtain a presentation

(2.8) $\quad \mathcal{P}_E = [y_1, y_2, x \; ; \; y_1y_2 = y_2y_1, x^{2p+1} = x, y_1x = xy_1^{\alpha_{11}}, y_2x = xy_2],$

for the semi-direct product $E = K_2 \rtimes_\theta A$. Therefore, by drawing quite similar pictures as in Figure 5, we have the following consequence of Theorem 2.4.

2.13. Corollary. *The presentation $\mathcal{P}_E$, as in* (2.8)*, is p-Cockcroft for all prime p, if $\alpha_{11} = 1 + pt$ where $t > 0$.*

2.14. Remark. The examples and corrollories given in this subsection can also be true for the general case of $k = np + 1$ and $l = 1$ where n is the positive integer.

3. Minimality

3.1. The Main Theorem. Let K_2 be the free abelian monoid rank 2 with a presentation $\mathcal{P}_{K_2} = [y_1, y_2 \; ; \; y_1y_2 = y_2y_1]$ and let A be the finite cyclic monoid with a presentation $\mathcal{P}_A = [x \; ; \; x^{2p+1} = x]$. Also, suppose that ψ is the endomorphism $\psi_\mathcal{M}$ of K, where $\mathcal{M} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$ such that (2.1) holds with the assumption $d = p$. Then, by Lemma 2.3, we get a semi-direct product $E = K_2 \rtimes_\theta A$ with a presentation

$$\begin{aligned} \mathcal{P}_E = [\, y_1, y_2, x \quad ; \quad & y_1y_2 = y_2y_1,\; x^{2p+1} = x, \\ & y_1x = xy_1^{\alpha_{11}}y_2^{\alpha_{12}},\; y_2x = xy_1^{\alpha_{21}}y_2^{\alpha_{22}} \,]. \end{aligned} \tag{3.1}$$

Let us assume that

$$\begin{aligned} &\alpha_{11} = 1,\ \alpha_{12} = \alpha_{21} = 0 \quad \text{and} \quad \alpha_{22} = 1 + pt_1 \ (t_1 > 0) \quad \text{or} \\ &\alpha_{22} = 1,\ \alpha_{12} = \alpha_{21} = 0 \quad \text{and} \quad \alpha_{11} = 1 + pt_2 \ (t_2 > 0), \end{aligned}$$

where p is a prime. Then, by Corollary 2.12 or Corollary 2.13, the presentation $\mathcal{P}_E$ in (3.1) is p-Cockcroft for any prime p and so, by Proposition 1.1, it is efficient.

Suppose that p is an odd prime. Then, in particular, $\mathcal{P}_E$ is not efficient if

$$det\mathcal{M} = exp_S(\mathbb{B}_{S,x}) \equiv 0 \ \text{ or } \ p - 1 \mod p.$$

Therefore our another main result in this paper is the following.

3.1. Theorem. *The presentation* $\mathcal{P}_E$*, as in* (3.1)*, is minimal but inefficient if* p *is an odd prime and*

$$\text{either} \begin{cases} \alpha_{11} = p-1, \\ \alpha_{12} = \alpha_{21} = 0, \\ \alpha_{22} = 1, \end{cases} \qquad \text{or} \begin{cases} \alpha_{11} = 1, \\ \alpha_{12} = \alpha_{21} = 0, \\ \alpha_{22} = p-1. \end{cases}$$

3.2. Preliminaries for the minimality result. Let M be a monoid with a presentation $\mathcal{P} = [\mathbf{y}; \mathbf{s}]$, and let $P^{(l)} = \bigoplus_{S\in\mathbf{s}} \mathbb{Z}Me_S$ be the free left $\mathbb{Z}M$-module with bases $\{e_S : S \in \mathbf{s}\}$. For an atomic monoid picture, say $\mathbb{A} = (U, S, \varepsilon, V)$ where $U, V \in F(\mathbf{y})$, $S \in \mathbf{s}$, $\varepsilon = \pm 1$, the left evaluation of the positive atomic monoid picture $\mathbb{A}$ is defined by $eval^{(l)}(\mathbb{A}) = \varepsilon \hat{U} e_S \in P^{(l)}$, where $\hat{U} \in M$. For any spherical monoid picture $\mathbb{P} = \mathbb{A}_1\mathbb{A}_2 \cdots \mathbb{A}_n$, where each $\mathbb{A}_i$ is an atomic picture for $i = 1, 2, \cdots, n$, we then define $eval^{(l)}(\mathbb{P}) = \sum_{i=1}^{n} eval^{(l)}(\mathbb{A}_i) \in P^{(l)}$. Let $\delta_{\mathbb{P},S}$ be the coefficient of e_S in $eval^{(l)}(\mathbb{P})$. So we can write $eval^{(l)}(\mathbb{P}) = \sum_{S\in\mathbf{s}} \delta_{\mathbb{P},S} e_S \in P^{(l)}$. Let $I_2^{(l)}(\mathcal{P})$ be the 2-sided ideal of $\mathbb{Z}M$ generated by the set

$$\{\delta_{\mathbb{P},S} : \mathbb{P} \text{ is a spherical monoid picture}, S \in \mathbf{s}\}.$$

Then this ideal is called the *second Fox ideal* of $\mathcal{P}$.

The fact of the following lemma has also been discussed in [4].

3.2. Lemma. *If* $\mathbf{Y}$ *is a trivializer of* $\mathcal{D}(\mathcal{P})$ *then second Fox ideal is generated by the set* $\{\delta_{\mathbb{P},S} : \mathbb{P} \in \mathbf{Y}, S \in \mathbf{s}\}$.

The concept of the second Fox ideals is needed for a *test of minimality* for monoid presentations (see [4]). The group version of this test has been proved by M. Lustig ([9]).

3.3. Theorem. *Let* $\mathbf{Y}$ *be a trivializer of* $\mathcal{D}(\mathcal{P})$ *and let* ψ *be a ring homomorphism from* $\mathbb{Z}M$ *into the ring of all* $n \times n$ *martices over a comutative ring* L *with* 1*, for some* $n \geq 1$*, and suppose* $\psi(1) = I_{n\times n}$*. If* $\psi(\lambda_{\mathbb{P},S}) = 0$ *for all* $\mathbb{P} \in \mathbf{Y}$*,* $S \in \mathbf{s}$ *then* $\mathcal{P}$ *is minimal.*

3.3. Proof of Theorem 3.1. As previously, let K_2 denotes the free abelian monoid rank two with a presentation $\mathcal{P}_{K_2} = [y_1, y_2, \ ; \ y_1y_2 = y_2y_1]$ and, for an odd prime p, let A denotes the finite cyclic monoid with a presentation $\mathcal{P}_A = [x, \ ; \ x^{2p+1} = x]$. Moreover let $\mathcal{M}$ be the matrix representation of K_2 with the assumption $\mathcal{M}^{2p+1} \equiv \mathcal{M} \mod p$. Then we have a semi-direct product $E = K_2 \rtimes_\theta A$ with a presentation $\mathcal{P}_E$ as in (3.1).

Suppose that $\alpha_{11} = 1$, $\alpha_{12} = \alpha_{21} = 0$ and $\alpha_{22} = p-1$ in $\mathcal{P}_E$.

Let us consider the picture $\mathbb{P}_{S,x}$, as drawn in Figure 2-(a), and also consider the generating set $\{y_1, y_2\}$ of $\mathcal{P}_{K_2}$. For a fixed element y_i in this set, let us assume that $\frac{\partial}{\partial y_i}$ denotes the Fox derivation with respect to y_i, and let $\frac{\partial^E}{\partial y_i}$ be the composition

$$\mathbb{Z}F(\{y_1, y_2\}) \xrightarrow{\frac{\partial}{\partial y_i}} \mathbb{Z}F(\{y_1, y_2\}) \longrightarrow \mathbb{Z}E,$$

where $F(\{y_1, y_2\})$ is the free monoid on $\{y_1, y_2\}$. Furthermore, for the relator $S : y_1y_2 = y_2y_1$, let us define $\frac{\partial^E S}{\partial y_i}$ to be $\frac{\partial^E S_+}{\partial y_i} - \frac{\partial^E S_-}{\partial y_i}$. Thus, for a fixed $y_i \in \{y_1, y_2\}$, the coefficients of $e_{T_{y_i x}}$ in $eval^{(l)}(\mathbb{P}_{S,x})$ is $\frac{\partial^E S}{\partial y_i}$. In fact

$$\frac{\partial^E S}{\partial y_1} = y_2 - 1 \quad \text{and} \quad \frac{\partial^E S}{\partial y_2} = 1 - y_1.$$

We then have the following proposition.

3.4. Proposition. *The second Fox ideal $I_2^{(l)}(\mathcal{P}_E)$ of $\mathcal{P}_E$ is generated by the elements*

$$\begin{array}{cc} 1-x(eval^{(l)}(\mathbb{B}_{S,x})), & 1-x^{k-1},\ 1-x^{k-2}, \cdots, 1-x, \\ \dfrac{\partial^E S}{\partial y_1}, & \dfrac{\partial^E S}{\partial y_2}, \\ eval^{(l)}(\mathbb{A}_{R_+,y_1}) - eval^{(l)}(\mathbb{A}_{R_-,y_1}), & eval^{(l)}(\mathbb{A}_{R_+,y_2}) - eval^{(l)}(\mathbb{A}_{R_-,y_2}). \end{array}$$

Proof. Recall that $\mathcal{D}(\mathcal{P}_E)$ has a trivializer $\mathbf{X_E}$ consisting of the sets $\mathbf{X_A}$, $\mathbf{X_{K_2}}$, $\mathbf{C_1}$ and $\mathbf{C_2}$ where $\mathbf{X_A}$ (see Figure 1), $\mathbf{X_{K_2}}$ (which is equal to the empty set) are the trivializer sets of $\mathcal{D}(\mathcal{P}_A)$ and $\mathcal{D}(\mathcal{P}_{K_2})$, respectively and $\mathbf{C_1}$, $\mathbf{C_2}$ consist of the pictures $\mathbb{P}_{S,x}$ (see Figure 2-(a) by assuming $\alpha_{11}=1$, $\alpha_{12}=\alpha_{21}=0$, $\alpha_{22}=p-1$), $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$ (see Figure 2-(b) by fixing $\mathbb{A}_{R_+,y_i}$ and $\mathbb{A}_{R,y_i}$ given in Figure 5), respectively. Now we need to calculate $eval^{(l)}(\mathbb{P}_{S,x})$, $eval^{(l)}(\mathbb{P}_{R,y_1})$, $eval^{(l)}(\mathbb{P}_{R,y_2})$, and $eval^{(l)}(\mathbb{P}^m_{k,l})$ $(1 \leq m \leq k-1)$. So we have

$$\begin{aligned} eval^{(l)}(\mathbb{P}_{S,x}) &= \delta_{\mathbb{P}_{S,x},S}e_S + \delta_{\mathbb{P}_{S,x},T_{y_1x}}e_{T_{y_1x}} + \delta_{\mathbb{P}_{S,x},T_{y_2x}}e_{T_{y_2x}} \\ &= (1-x(eval^{(l)}(\mathbb{B}_{S,x})))e_S + (\frac{\partial^E S}{\partial y_1})e_{T_{y_1x}} + (\frac{\partial^E S}{\partial y_2})e_{T_{y_2x}} \\ eval^{(l)}(\mathbb{P}_{R,y_1}) &= \delta_{\mathbb{P}_{R,y_1},R}e_R + \delta_{\mathbb{P}_{R,y_1},T_{y_1x}}e_{T_{y_1x}} + \delta_{\mathbb{P}_{R,y_1},T_{y_2x}}e_{T_{y_2x}} \\ &= (1-y_1)e_R + (1+x+x^2+\cdots+x^{2p}-1)e_{T_{y_1x}} + 0e_{T_{y_2x}} \\ &= (1-y_1)e_R + (eval^{(l)}(\mathbb{A}_{R_+,y_1}) - eval^{(l)}(\mathbb{A}_{R_-,y_1}))e_{T_{y_1x}}. \end{aligned}$$

$$\begin{aligned} eval^{(l)}(\mathbb{P}_{R,y_2}) &= \delta_{\mathbb{P}_{R,y_2},R}e_R + \delta_{\mathbb{P}_{R,y_2},T_{y_1x}}e_{T_{y_1x}} + \delta_{\mathbb{P}_{R,y_2},T_{y_2x}}e_{T_{y_2x}} \\ &= (1-y_2)e_R + 0e_{T_{y_1x}} + (1+x+xy_2+x^2y_2^2+\cdots+x^2y_2^{\alpha_{22}-1}+ \\ &\quad \cdots + x^{2p} + x^{2p}y_2 + x^{2p}y_2^2 + \cdots + x^{2p}y_2^{\alpha_{22}{}^{2p}})e_{T_{y_2x}} \\ &= (1-y_2)e_R + (eval^{(l)}(\mathbb{A}_{R_+,y_2}) - eval^{(l)}(\mathbb{A}_{R_-,y_2}))e_{T_{y_2x}}. \end{aligned}$$

Also, for each $1 \leq m \leq k-1$, $eval^{(l)}(\mathbb{P}^m_{k,l}) = \delta_{\mathbb{P}^m_{k,l},R}e_R$, where $\delta_{\mathbb{P}^m_{k,l},R} = 1-x^{k-m}$.

Thus, by Lemma 3.2, we get the result as required. □

Let $aug : \mathbb{Z}E \longrightarrow \mathbb{Z}$, $s \longmapsto 1$ be the augmentation map.

3.5. Lemma. *We have the following equalities.*

1) $aug(eval^{(l)}(\mathbb{B}_{S,x})) = \exp_S(\mathbb{B}_{S,x})$.

2) *i)* $aug(\frac{\partial^E S}{\partial y_1}) = aug(y_2-1) = \exp_{y_1}(S)$,

ii) $aug(\frac{\partial^E S}{\partial y_2}) = aug(1-y_1) = \exp_{y_2}(S)$.

3) $\left.\begin{array}{lcl} i)\ aug(eval^{(l)}(\mathbb{A}_{R_+,y_i})) &=& exp_{T_{y_ix}}(\mathbb{A}_{R_+,y_i}), \\ ii)\ aug(eval^{(l)}(\mathbb{A}_{R_-,y_i})) &=& exp_{T_{y_ix}}(\mathbb{A}_{R_-,y_i}), \end{array}\right\}$ *for* $i \in \{1,2\}$.

4) $aug(eval^{(l)}(\mathbb{P}^m_{k,l})) = aug(1-x^m) = exp_R(\mathbb{P}^m_{k,l}))$, $1 \leq m \leq k-1$.

Proof. Since similar proofs of 1) and 2) can be found in [4], we will only show the remaining conditions.

<u>*Proof of* 3)</u>:

We will just consider $i)$ since the proof of $ii)$ is completely same with the first one. We can write

$$eval^{(l)}(\mathbb{A}_{R_+,y_i}) = \varepsilon_1 W_1 e_{T_{y_i x}} + \varepsilon_2 W_2 e_{T_{y_i x}} + \cdots + \varepsilon_n W_n e_{T_{y_i x}},$$

where, for $1 \leq j \leq n$, $\varepsilon_j = \pm 1$ and each W_j is the certain word on the set $\{y_1, y_2\}$. In the right hand side of the above equality, each term $\varepsilon_j W_j e_{T_{y_i x}}$ corresponds to a single $T_{y_i x}$ disc and, in fact, the value of each ε_j gives the sign of this single $T_{y_i x}$ disc. Therefore, since the $T_{y_i x}$ discs can only be occured in the subpictures $\mathbb{A}_{R_+,y_i}$ and $\mathbb{A}_{R_-,y_i}$, the sum of each ε_j (which is equal to the $aug(eval^{(l)}(\mathbb{A}_{R_+,y_i}))$ must give the exponent sum of the $T_{y_i x}$ discs in the picture $\mathbb{P}_{R,y_i}$, as required.

Proof of 4):

For each $1 \leq m \leq k-1$, since each $\mathbb{P}^m_{k,l}$ contains just two R-discs (one is positive and the other is negative), we write

$$eval^{(l)}(\mathbb{P}^m_{k,l}) = -{W_1}^m e_R + {W_2}^m e_R,$$

where each ${W_j}^m$ is the word on x $(1 \leq j \leq 2)$. As in the previous case, by considering the each term in above equalitiy, we get the sign of this single R-disc. Then the sum of the whole these signs (i.e the augmentation of the evaluation of each picture) must give the exponent sum of R-discs. That is,

$$aug(eval^{(l)}(\mathbb{P}^m_{k,l})) = aug(1 - x^m) = \exp_R(\mathbb{P}^m_{k,l}),$$

as required. Hence the result. □

We note that $det\mathcal{M} = \exp_S(\mathbb{B}_{S,x}) = p - 1$, where p is an odd prime, for the picture $\mathbb{P}_{S,x}$ in Figure 2-(a).

Also let us consider the homomorphism from E onto the finite cyclic monoid $M_{k,l}$ generated by x, defined by $y_1,\ y_2 \longmapsto 1$, $x \longmapsto x$. This induces a ring homomorphism

$$\gamma : \mathbb{Z}E \longrightarrow M_{k,l}[x].$$

Let η be the composition of γ and the mapping

$$M_{k,l}[x] \longrightarrow \mathbb{Z}_p[x], \quad x \longmapsto x,\ \ n \longmapsto \overline{n}\ (n \in \mathbb{Z}),$$

where $\overline{n}$ is n (mod p) and $p \mid (k - l)$.

We note that the restriction of η to the subring $\mathbb{Z}K_2$ of $\mathbb{Z}E$ is just the augmentation map $aug_p : \mathbb{Z}K_2 \longrightarrow \mathbb{Z}_p$ by modulo p. Therefore the following lemma is valid.

3.6. Lemma. *We have the following equalities.*

i) $aug_p(eval^{(l)}(\mathbb{P}^m_{k,l})) \equiv 0\ (mod\ p)$.

ii) $aug_p(\frac{\partial^E S}{\partial y_1}) = aug_p(\frac{\partial^E S}{\partial y_2}) \equiv 0\ (mod\ p)$.

iii) $aug_p(eval^{(l)}(\mathbb{P}_{R,y_1})) \equiv 0\ (mod\ p)$ *and* $aug_p(eval^{(l)}(\mathbb{P}_{R,y_2})) \equiv 0\ (mod\ p)$.

Proof. By Lemma 3.5-4), for $1 \leq m \leq k-1$, since $aug(eval^{(l)}(\mathbb{P}^m_{k,l})) = aug(1 - x^m) = exp_R(\mathbb{P}^m_{k,l}))$ and since, by Figure 1, $exp_R(\mathbb{P}^m_{k,l})) = 0$, it is obvious that the condition $i)$ holds. Similarly, by Lemma 3.5-2), $aug(\frac{\partial^E S}{\partial y_1}) = \exp_{y_1}(S) = 0 = \exp_{y_2}(S) = aug(\frac{\partial^E S}{\partial y_2})$. Then the condition $ii)$ clearly holds.

Let us consider the generating pictures $\mathbb{P}_{R,y_1}$ and $\mathbb{P}_{R,y_2}$, as drawn in Figure 2-(b) (by fixing the subpictures $\mathbb{A}_{R_+,y_i}$ and $\mathbb{A}_{R_-,y_i}$ given in Figure 5 into them). By Lemma 3.5-3), we then have

$$\begin{aligned} aug(eval^{(l)}(\mathbb{P}_{R,y_1})) &= aug[eval^{(l)}(\mathbb{A}_{R_+,y_1}) - eval^{(l)}(\mathbb{A}_{R_-,y_1})]e_{T_{y_1x}} \\ &\quad +aug(1-y_1)e_R \\ &= exp_{T_{y_1x}}(\mathbb{A}_{R_+,y_1}) - exp_{T_{y_1x}}(\mathbb{A}_{R_-,y_1}) + 0 \\ &= (2p+1) - 1 = 2p \end{aligned}$$

which is congruent to zero by modulo p. Moreover

$$\begin{aligned} aug(eval^{(l)}(\mathbb{P}_{R,y_2})) &= aug[eval^{(l)}(\mathbb{A}_{R_+,y_2}) - eval^{(l)}(\mathbb{A}_{R_-,y_2})]e_{T_{y_2x}} \\ &\quad +aug(1-y_2)e_R \\ &= exp_{T_{y_2x}}(\mathbb{A}_{R_+,y_2}) - exp_{T_{y_2x}}(\mathbb{A}_{R_-,y_2}) + 0 \\ &= \frac{\alpha_{22}^{2p+1}-1}{\alpha_{22}-1} - 1 = \frac{\alpha_{22}^{2p+1}-\alpha_{22}}{\alpha_{22}-1} \\ &= \frac{(p-1)^{2p+1}-(p-1)}{(p-2)} \end{aligned}$$

which is congruent to zero by modulo p. Hence the result. $\square$

Thus, by Lemmas 3.5 and 3.6, the image of $I_2^{(l)}(\mathcal{P}_E)$ under η is the ideal of $\mathbb{Z}_p[x]$ that is generated by the element $1-x(\exp_S(\mathbb{B}_{S,x})) = 1-\overline{(p-1)}x$ since $\exp_S(\mathbb{B}_{S,x}) = det\mathcal{M} = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = p-1$. In other words,

$$\eta(I_2^{(l)}(\mathcal{P}_E)) \quad = \quad <1-\overline{(p-1)}x>= I, \text{ say.}$$

3.7. Remark. A simple calculation shows that $I \neq \mathbb{Z}_p[x]$ since $1 \notin I$.

Let ψ be the composition

$$\mathbb{Z}E \xrightarrow{\eta} \mathbb{Z}_p[x] \xrightarrow{\phi} \mathbb{Z}_p[x]/I,$$

where ϕ is the natural epimorphism. Then

$$\begin{aligned} \psi(1-\hat{x}(eval^{(l)}(\mathbb{B}_{S,x}))) &= \phi\eta(1-\hat{x}(eval^{(l)}(\mathbb{B}_{S,x}))) \\ &= \phi(1-\hat{x}\overline{(\exp_S(\mathbb{B}_{S,x})}) \text{ since } \eta \text{ is a ring} \\ &\quad \text{homomorphism and by Lemma } 3.5-1) \\ &= \phi(1-\hat{x}\overline{(p-1)}) \text{ since } \exp_S(\mathbb{B}_{S,x}) = p-1 \\ &= 0. \end{aligned}$$

Moreover, by Lemmas 3.5 and 3.6, the images of $1-x^{k-1}$, $1-x^{k-2}$, $\cdots$, $1-x$, $\frac{\partial^E S}{\partial y_1}$, $\frac{\partial^E S}{\partial y_2}$, $eval^{(l)}(\mathbb{A}_{R_+,y_1}) - eval^{(l)}(\mathbb{A}_{R_-,y_1})$, $eval^{(l)}(\mathbb{A}_{R_+,y_2}) - eval^{(l)}(\mathbb{A}_{R_-,y_2})$ under ψ are all equal to 0 since the related exponent sums are all congruent to zero by modulo p. That means the images of the generators $I_2^{(l)}(\mathcal{P}_E)$ are all 0 under ψ. Therefore, by Theorem 3.3 (Pride), $\mathcal{P}_E$ is minimal and so $E = K_2 \rtimes_\theta A$ is a minimal but inefficient monoid.

We note that, by using the same method as in this proof, one can see that E is a minimal but inefficient monoid if p is an odd prime and

$$\alpha_{11} = p-1, \ \alpha_{22} = 1 \text{ and } \alpha_{12} = 0 = \alpha_{21}.$$

These all above progress complete the proof of Theorem 3.1. $\diamondsuit$

3.8. Example. Let $p = 3$ and $\mathcal{M} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. Thus we have $\mathcal{M}^7 \equiv \mathcal{M} \mod 3$ and, by Lemma *2.3*, we have $E = K_2 \rtimes_\theta A$ with a presentation $\mathcal{P}_E = [y_1, y_2, x\ ;\ y_1y_2 = y_2y_1, x^7 = x, y_1x = xy_1, y_2x = xy_2^2]$, as in *(3.1)*, for the monoid E. It is clear that $det\mathcal{M} = 2$ so, by Theorem *2.4*, $\mathcal{P}_E$ is inefficient and also, by Theorem *3.1*, $\mathcal{P}_E$ is minimal. Moreover, by taking the matrix $\mathcal{M} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, it can also be obtained a minimal but inefficient presentation.

3.9. Remark. *1)* By using same progress as in the proof of Theorem *3.1*, one can see that if $det\mathcal{M} = 0$ then $1 \in I$, that is,

$$\eta(I_2^{(l)}(\mathcal{P}_E)) =< 1 >= I$$

and so $I = \mathbb{Z}_p[x]$ (see Remark *3.7*). In fact this equality holds for any prime p. That means the minimality test (Theorem *3.3*) used in this paper cannot work for this case. Therefore it can be remained as a conjecture whether the presentation obtained by this case is minimal.

2) For $p = 2$, we have $det\mathcal{M} = 0$ or 1. In the case of $det\mathcal{M} = 1$, we know that $\mathcal{P}_E$ is efficient (see Corollary *2.12* or Corollary *2.13*) and so we cannot apply Theorem *3.1*. Furthermore if $det\mathcal{M} = 0$ then we need to turn back condition *1)*.

References

[1] Ateş, F. and Cevik, A.S. *Separability and Efficiency under Standard Wreath Product in terms of Cayley Graphs*, Rocky Mont. J. Math. **38**, 779-800, 2008.

[2] Ateş, F. and Cevik, A.S. *The p-Cockcroft Property of Central Extensions of Groups II*, Monatshefte für Math. **150**, 181-191, 2007.

[3] Cevik, A.S. *The p-Cockcroft property of the semi-direct products of monoids*, Int. J. Algebra Comput. **13** (1), 1-16, 2003.

[4] Cevik, A.S. *Minimal but inefficient presentations of the semi-direct products of some monoids*, Semigroup Forum **66**, 1-17, 2003.

[5] Epstein, D.B.A. *Finite presentations of groups and* 3*-manifolds*, Quart. J. Math. Oxford **12**, 205-212, 1961.

[6] Howie, J.M. *Fundamentals of Semigroup Theory*, Oxford University Press, 1995.

[7] Ivanov, S.V. *Relation modules and relation bimodules of groups, semigroups and associative algebras*, Int. J. Algebra and Comput. **1**, 89-114, 1991.

[8] Kilgour, C.W. and Pride, S.J. *Cockcroft presentations*, Journal of Pure and Appl. Algebra **106** (3), 275-295, 1996.

[9] Lustig, M. *Fox ideals, $\mathcal{N}$-torsion and applications to groups and* 3*-monifolds*, in "Two-dimensional homotopy and combinatorial group theory" (C. Hog-Angeloni, W. Metzler and A.J. Sieradski, editors), Cambridge University Press, 219-250, 1993.

[10] Wang, J. *Finite derivation type for semi-direct products of monoids*, Theoretical Computer Science **191** (1-2), 219-228, 1998.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 915 – 922

FG-morphisms and FG-extensions

Ceren Sultan ELMALI* and Tamer UGUR†

Abstract

We investigate the relations between Fan-Gottesman compactification and categories. We deal with maps having an extension to a homeomorphism between the Fan-Gottesman compactification of their domains and ranges.

2000 AMS Classification: 54D35, 18A40.

Keywords: Fan-Gottesman compactification, categories.

Received 31 : 05 : 2012 : Accepted 25 : 10 : 2013 Doi : 10.15672/HJMS.2014437526

The first section of this paper contains some preliminaries about categories. Category theory provides the language and mathematical foundations for discussing properties of large classes of mathematical objects such as the class of all sets or all groups while circumventing problems such as Russell's paradox. In fact S.Eilenberg and S. MacLane [10,11] give a lot of informations about categories and functors. Category theory has also played a foundational role for formalizing new concepts such as schemes which are fundamental to major areas of contemporary research. Pioneering work of this nature was done by A.Grothendieck [7], K. Morita [12,13,14,15] and others.

The second section of this paper contains some preliminaries about the Fan-Gottesman compactification. In 1952, Ky Fan and Noel Gottesman defined a compactification that is similar to the Wallman compactification, introduced by Henry Wallman in 1938 [17], and afterwards called Fan-Gottesman compactification of regular spaces with a normal base [5]. We investigated the relations between the Fan-Gottesman and Wallman compactification and showed that Fan-Gottesman compactification of some specific and interesting spaces such as normal A_2 and T_4 is Wallman-type compactification [4]. In this section we show that Fan-Gottesman compactification can be obtained via base consisting of open ultrafilters.

*Department of Mathematics, Faculty of Science,Erzurum Technic University, Erzurum, 25240, Turkey.
Email: ceren.elmali@erzurum.edu.tr
†Department of Mathematics, Faculty of Science, Ataturk University, Erzurum, 25240, Turkey.
Email:tugur@atauni.edu.tr

In [9], Herrlich has stated that it is of interest to determine if the Wallman compactification may be regarded as a functor, especially as an epireflection functor, on a suitable category of spaces. This problem was solved affirmatively by Harris in [8].

In [3], Belaid and Echi characterize when Wallman extensions of maps are homeomorphisms.

The third section of this paper, we define $FG-$morphism and $FG-$extension. Let X, Y be two T_3 spaces and $q : X \rightarrow Y$ a continuous map. An $FG-$extension of q is a continuous map $F(q) : FX \rightarrow FY$ such that $F(q) \circ f_X = f_Y \circ q$, where FX is the Fan-Gottesman compactification of X and $f_X : X \rightarrow FX$ is the canonical embedding of X into its Fan-Gottesman compactification FX. We will characterize when Fan-Gottesman extensions of maps are homeomorphisms.

1. Categories

A category C consist of a certain collection of object $Ob(C)$ and for any two object $b, c \in Ob(C)$, there is a set $morph(b, c)$ of morphism (function) between b and c. This collection may be empty, but an identity morphism 1_b must be contained in $morph(b, b)$. Furthermore if there are morphism $morph(b, c)$ and morphism $morph(a, b)$, then their composition must be in $morph(a, c)$. Given two categories C and D, then a map can be defined between these, the so called functor, $F : C \rightarrow D$. A functor send object of C to object of D and morphism in C to morphism in D subject to certain condition. Furthermore, it is possible to define maps between functors, the so called natural transformation [11].

One usually requires the morphisms to preserve the mathematical structure of the objects. So if the objects are all groups, a good choice for a morphism would be a group homomorphism. Similarly, for vector spaces, one would choose linear maps, and for differentiable manifolds, one would choose differentiable maps.

In the category of topological spaces, morphisms are usually continuous maps between topological spaces. However, there are also other category structures having topological spaces as objects, but they are not nearly as important as the "standard" category of topological spaces and continuous maps.

We denote by Top the category of topological spaces with continuous maps as morphisms, and by Top_i the full subcategory of Top whose objects are the T_i spaces. There are several ways to generalize the usual separation properties T_0, T_1, T_2, T_3 and T_4 of topology to topological categories [1,2]. All the above categories are full reflective subcategories of Top. There is a universal T_i -space for every topological space X, we denote it by $\mathbf{T}_i(X)$. The assignment $X \rightarrow \mathbf{T}_i(X)$ defines a functor $\mathbf{T}_i$ from Top onto Top_i, which is a left adjoint functor of the inclusion functor $Top_i \rightarrow Top$.

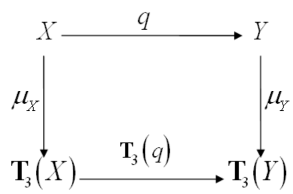
It is recalled that a continuous map $q : Y \rightarrow Z$ is said to be a *quasihomeomorphism*, if $U \rightarrow q^{-1}(U)$ defines a bijection $O(Z) \rightarrow O(Y)$ [7], where $O(Y)$ is the set of all open subsets of the space Y. If Z is T_2 space and, q is not onto, thus q is not a quasihomeomorphism. As showed by the open sets Z, $Z \backslash \{z\}$ for some $z \in Z$. On the other hand, if Z is $\mathbb{R}$, with open sets $\{(-\infty, c) : c \in (-\infty, \infty]\}$ and Y is its subspace $\mathbb{Q}$, then the embedding is a quasihomeomorphism. A subset S of a topological space X is said to be strongly dense in X, if S meets every nonempty locally closed subset of X [9]. In here, locally closed means that every point x of S has a neighbourhood such that $V_x \cap S$ is a closed subset of V_x. In other words, S is locally closed if and only if $S = O \cap F$ for some open subset O of X and some closed subset F of X. In addition, one most evident definition is equivalent to closedness. Thus, a subset S of X is strongly dense if and only if the canonical injection $S \rightarrow X$ is a quasihomeomorphism. Besides, a continuous map

$q : X \to Y$ is a quasihomeomorphism if and only if the topology of X is the inverse image of Y by q and the subset $q(X)$ is strongly dense in Y [7].

It is known that T_0 -identification of a topological space is done by Stone [17].

Now, we will construct T_3 reflection for X in Top. Firstly, we construct regular reflection by taking the supremum of all regular topologies which are coarser than the topology of X. This is a bireflection in Top, in other words, the underlying set stays the same. Then, apply it to the T_0- reflection. We get a space which is regular and T_0, hence regular and T_1. The composite of the two reflection is T_3-reflection.

Let X be a topological space and define $\sim$ on X by $x \sim y$ if and only if $cl_X \{x\} = cl_X \{y\}$. Then, $\sim$ is an equivalence relation on X and the resulting quotient space $X/\sim$ is T_0-space. This procedure and the space it produces are referred to as the T_0-identification of X. Clearly $\mathbf{T}_0 (X) = X/\sim$. $\mathbf{T}_0 (X)$ is called T_0- reflection. The canonical onto map from X onto its T_0- identification $\mathbf{T}_0 (X)$ will be denoted by μ_X. It is clear that μ_X is an onto quasihomeomorphism. If $q : X \to Y$ is a continuous map,



then the diagram is commutative. $\mathbf{T}_0$ defines a (covariant) functor from Top to itself. Thus, we get a space which is regular and T_0, hence regular and T_1. The composite of the two reflections is T_3-reflection.

2. Fan-Gottesman Compactification

A compactification of a topological space X is a compact Hausdorff space Y containing X as a subspace such that $cl_Y X = Y$. In addition there are a lot of compactification methods applying different topological space such as Aleksandrov (one-point), Wallman, Stone-Cech. But, we study with Fan-Gottesman compactification.

Let β be a class of open sets in X. If it satisfies the following three conditions, it is called a *normal base*.

(1) β is closed under finite intersections
(2) If $B \in \beta$, then $X - cl_X B \in \beta$, where $cl_X B$ denotes the closure of B in X.
(3) For every open set U in X and every $B \in \beta$ such that $cl_X B \subset U$, there exists a set $D \in \beta$ such that $cl_X B \subset D \subset cl_X D \subset U$.

We consider a regular space having a normal base for open sets i.e., which satisfies the above three properties of normal base. A *chain family on* β is a non-empty family of sets of β such that

$$cl_X B_1 \cap cl_X B_2 \cap ... \cap cl_X B_n \neq \varnothing$$

for any finite number of sets B_i of the family. Every chain family on β is contained in at least one maximal chain family on β by Zorn's lemma. Maximal chain families on β will be denoted by letters as $a^*, b^*, ...$, and also the set of all maximal chain families on β will be denoted by $(X, \beta)^*$. Whose topology is defined as follow. For each $B \in \beta$, let

$$\tau (B) = \{b^* \in (X, \beta)^* : \text{there exists a } A \in b^* \text{ with } cl_X B \subset A\}$$

Then, the topology of $(X,\beta)^*$ is defined by taking

$$\beta^* = \{\tau(B) : B \in \beta\}$$

as a base of open sets. $(X,\beta)^*$ is a compact Hausdorff space and is a compactification of our regular space. Afterwards this compactification is called Fan-Gottesman compactification [6].

Now, we determine the Fan-Gottesman compactification via open ultrafilters.

2.1. Definition. Let X be a T_3 space and FX the subcollection of all maximal ultrafilter of closed subsets on X. For each open set $O \subset X$, define $O^* \subset FX$ to be the set

$$O^* = \left\{\hat{G} \in FX : \hat{G} \text{ consists of } cl_X O\right\}$$

Let Φ be the family of O^*. It is clear that Φ is the base for open sets of topology on FX. FX is a compact space and it is called the Fan-Gottesman compactifications of X.

In order to avoid the confusion between FX and $(X,\beta)^*$, we will use FX when it regarded as Fan-Gottesman compactification of X.

On the other hand, for each closed set $D \subset X$, we define $D^* \subset FX$ by $D^* = \left\{\hat{G} \in FX : \hat{G} \text{ consists of } G \subseteq D \text{ for some } G\right\}$. The following properties of FX are useful;

(i) If $U \subset X$ is open, then $FX - U^* = (FX - U)^*$

(ii) If $D \subset X$ is closed, then $FX - D^* = (FX - D)^*$

(iii) If U_1 and U_2 are open in X, then $(U_1 \cap U_2)^* = U_1^* \cap U_2^*$ and $(U_1 \cup U_2)^* = U_1^* \cup U_2^*$

Properties We consider the map $f_X : X \to FX$ defined by $f_X(x) = \hat{G}_x$, the closed ultrafilter converging to x in X. In order to avoid the confusing between $\hat{G}_x$ and $\hat{G}$, we will use $\hat{G}_x$ when it regarded as the maximal filter of closures of open sets containing x. Then the following properties hold.

1. If U is open in X, then $\overline{f_X(U)} = U^*$. In particular $f_X(X)$ is dense in FX.
2. f_X is continuous and it is an embedding of X in FX if and only if X is a T_3-space.
3. If $U_1$and U_2 are open subsets of X, then $\overline{f_X(U_1 \cap U_2)} = \overline{f_X(U_1)} \cap \overline{f_X(U_2)}$.
4. FX is a compact T_2-space .

For a T_3 space, we define $FGX = F(X)$ and we call it the Fan-Gottesman compactification of X. The notation FX is reserved only for T_3 spaces so that it is better to use some other notation for topological spaces. The same for f_X : f_X is reserved for topological space; for $\mathbf{T}_3$ space, we define $F_X = f_X \circ \mu_x$ where μ_x is the canonical onto map from X onto its $T_3 - reflection$, $\mathbf{T}_3(X)$.

Since μ_x is an onto quasihomeomorphism, one obtains immediately that FGX can be described exactly as FX for T_3 space. The above properties are also true for a $\mathbf{T}_3$ space.

2.2. Remark. Let X be a T_3 space. Then, the following properties hold:

1. For each open subset U of X, we have $F_X(U) \subseteq U^*$
2. For each closed subset C of X, we have $F_X(C) \subseteq C^*$
3. Let U be open and C closed in a T_3 space. Then, $U \cap C \neq \varnothing$ if and only if $U^* \cap C^* \neq \varnothing$

2.3. Proposition. *Let X be a T_3 space and*

1. U be an open subset of X. If U is compact, then $U^* = F_X(U)$.
2. V be a closed subset of X. If V is compact, then $V^* = F_X(V)$.

Proof. Suppose that V is closed in X. We have $F_X(V) \subseteq V^*$ from Remark 1. If $\hat{G} \in V^*$, then there exists $G \in \hat{G}$ such that $G \subseteq V$. Then $V - G$ is compact by compactness of V. Thus $\cap \left\{ H \cap (V - G) : H \in \hat{G} \right\} \neq \varnothing$. If $x \in \cap \left\{ H \cap (V - G) : H \in \hat{G} \right\}$, then $\hat{G} = F_X(x)$. Hence, $\hat{G} \in F_X(V)$. Thus, $V^* \subseteq F_X(V)$. Therefore, $V^* = F_X(V)$.

Now, suppose that U is open in X. Let $\hat{G} \in U^*$. Thus, $U \in \hat{G}$. Since, $\cap \left\{ H : H \in \hat{G} \right\} \neq \varnothing$, we take an $x \in \cap \left\{ H : H \in \hat{G} \right\}$. It is seen that $\hat{G} = F_X(x)$. Therefore, according to Remark 1, $U^* = F_X(U)$. □

3. FG-morphisms and FG -extensions

Recall from [3] that a subset S of a topological space X is said to be *sufficiently dense* if S meets each nonempty closed subset and each nonempty open subset of X. By an *almost -homeomorphism* (*α-homeomorphism*, for short), we mean a continuous map $q : X \to Y$ such that $q(X)$ is sufficiently dense in Y and the topology of X is the inverse image of Y by q.

3.1. Definition. i) A subset C of a topological space is said to be openly dense if C meets each nonemty open subset of X.

Thus we have the following implications:

$$\begin{array}{c} \text{Strongly dense} \Rightarrow \text{Sufficiently dense} \Rightarrow \text{openly dense} \\ \Downarrow \\ \text{Dense} \end{array}$$

3.2. Definition. By a Fan-Gottesman morphism (FG-morphism, for short), we mean a continuous map $q : X \to Y$ such that $q(X)$ is openly dense in Y and the topology of X is the inverse image of Y by q.We conclude that

homeomorphism ⇒quasihomeomorphism ⇒ α-homeomorphism ⇒ FG-morphism

3.3. Theorem.

(1) *The composition of two FG-morphisms is an FG-morphism.*
(2) *If $q : X \to Y$ is an FG-morphism and X is T_0, then q is injective.*
(3) *If $q : X \to Y$ is an FG-morphism and Y is T_1, then q is an onto homeomorphism.*
(4) *If $q : X \to Y$ is an FG-morphism, X is T_0 and Y is T_1, q is a homeomorphism.*

Proof. We show that (1). Let $p : X \to Y$ and $q : Y \to Z$ be two FG-morphisms. Clearly, the topology of X is the inverse image of Z by $q \circ p$. Let A be open subset of Z. Since $q^{-1}(A)$ is open in Y, the $p(x) \cap q^{-1}(A) \neq \varnothing$, so that $A \cap q(p(X)) \neq \varnothing$. Hence, $q \circ p$ is an FG-morphism.

(2) Let x_1, x_2 be two points of X with $q(x_1) = q(x_2)$. Suppose that $x_1 \neq x_2$. Then, there exists an open subset U of X such that $x_1 \in U$, $x_2 \notin U$, since X is T_0. Because there exists an open subset H of Y satisfying $q^{-1}(H) = U$, we get $q(x_1) \in H$ and $q(x_2) \notin H$, which is impossible. It follows that q is injective.

(3) Let $y \in Y$. Then, $\{y\}$ is a locally closed subset of Y. Hence, $\{y\} \cap q(X) \neq \varnothing$, since $q(X)$ is strongly dense in Y. Thus, $y \in q(X)$, hence q is an onto map.

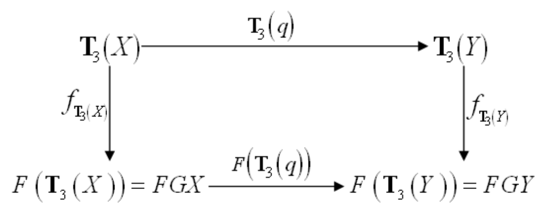
(4) It is clear that q is homeomorphism from (2) and (3). □

Now, we define FG-extensions.

3.4. Definition. A continuous map $q : X \to Y$ between T_3 spaces is said to be an FG-extension, if there is a continuous map $F(q) : FX \to FY$ such that $f_Y \circ q = F(q) \circ f_X$.

3.5. Theorem. *Let X, Y be two T_3 spaces and $q : X \to Y$ an FG-morphism. Then, q has an FG-extension which is a homeomorphism.*

Proof. We remark that diagram in the introduction commutes. Hence, $\mathbf{T}_3(q) \circ \mu_x = \mu_y \circ q$. Thus, $\mathbf{T}_3(q) \circ \mu_x$ is an FG-morphism. Now, $\mathbf{T}_3(q)$ is an FG-morphism from Proposition 2.1, since μ_x is a quasihomeomorphism. Therefore, $\mathbf{T}_3(q)$ is a homeomorphism by Proposition 2.1. It follows that $\mathbf{T}_3(q)$ has a canonical FG-extension $F(\mathbf{T}_3(q))$ which is a homeomorphism. Thus, the diagram commutes. If we denote $FG(q) = F(\mathbf{T}_3(q))$, then the diagram indicates clearly that $FG(q)$ is an FG-extension of q which is a homeomorphism.



□

It is known that if X is a T_4 space, then $FGX = wX = \beta(X)$ (the Wallman and Stone-Čech compactification, respectively)[4].

3.6. Corollary. *If* $\mathbf{T}_3(X)$ *is a T_4 space, then* $FGX = w(\mathbf{T}_3(q)) = \beta(\mathbf{T}_3(q))$.

3.7. Definition. Let X be a T_3 space and Y a subspace of X.

(1) Y is called a Fan-Gottesman generator ($FG-$generator) of X, if FGY is homeomorphic to FGX.
(2) Y is called a strong Fan-Gottesman generator (sFG-generator) of X, if the canonical embedding $i : Y \to X$ has an FG-extension $FG(i)$ which is a homeomorphism.

Clearly, sFG-generator$\Rightarrow$ $FG-$generator

3.8. Theorem. *Let X, Y be two T_3 spaces and $q : X \to Y$ a continuous map. Then, the following statements are equivalent:*

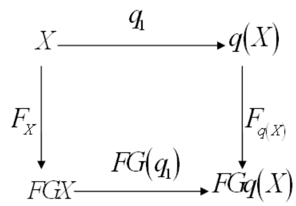
(1) q has an FG-extension which is a homeomorphism.
(2) $q(X)$ is an sFG-generator of Y and the topology of X is the inverse image of Y by q.

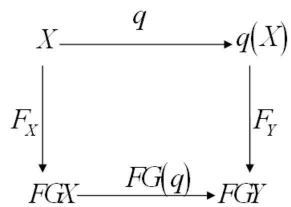
Proof. $(i) \Rightarrow (ii)$ Firstly, we show that the topology of X is the inverse image of Y by q. Let U be an open subset of X. Since $FG(q)$ is a homeomorphism, $FG(q)(U^*) = V$ is a closed subset of wY. Set $G = F_y^{-1}(V)$. We prove that $U = q^{-1}(G)$.

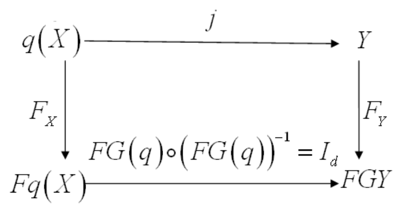
(a) Let $x \in U$. Then, $F_X(x) \in F_X(U) \subseteq U^*$. Hence, $FG(q)(F_X(x)) \in FG(q)(U^*) = V$ which gives $F_Y(q(x)) \in V$. It follows that $q(x) \in F_Y^{-1}(V) = G$. Therefore, $x \in q^{-1}(G)$.

(b) Conversely, let $x \in q^{-1}(G)$. Then, $q(x) \in G = F_X^{-1}(V)$; this means that $(F_Y \circ q)(x) \in V$, so that $FG(q)(F_X(x)) \in V = FG(q)(U^*)$. Since $FG(q)$ is bijective, $F_X(x) \in U^*$. Hence, $x \in F_X^{-1}(U^*) = U$. We have proved that $U = q^{-1}(G)$. In other words, the topology of X is the inverse image of Y by q.

Secondly, we show that $q(X)$ is an sFG-generator of Y. According to (1), the map $q_1 : X \to q(X)$ induced by q is an FG-morphism. Hence, q_1 has an FG-extension $F(q_1)$ which is a homeomorphism, by Proposition 2.1. Thus, the diagrams commute.

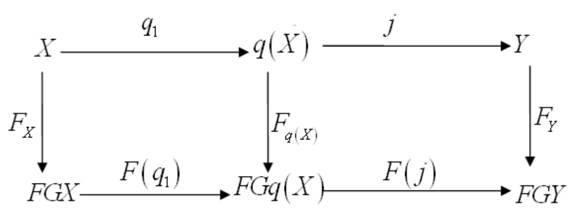






Let $j : q(X) \to Y$ be the canonical embedding. Clearly, the diagram commutes.

Therefore, j has $FG(q) \circ (FG(q))^{-1} = I_d$ as an FG-extension which is a homeomorphism. This means that $q(X)$ is an sFG generator of Y. $(ii) \Rightarrow (i)$ We assume (ii). The map $q_1 : X \to q(X)$ induced by q is an FG-morphism. Thus, according to Proposition 2.1, q_1 has an FG-extension $F(q_1)$ which is a homeomorphism. On the other hand, the canonical embedding $j : q(X) \to Y$ has an FG-extension which is a homeomorphism, by Proposition 2.1. It follows that the two diagrams commute.



Therefore, $F(j) \circ F(q_1)$ is an FG-extension of $q : X \to Y$ which is a homeomorphism. $\square$

Theorem 3.4 seems us to the following classical fact about the Stone-Ćech compactification $e_X : X \to \beta X$ of a Tychonoff space X .

Consider any continuous mapping $p : X \to Y$, where Y is also Tychonoff. Then, the map $\beta(p) : \beta X \to \beta Y$ is a homeomorphism if and only if p is a dense C^*-embedding.

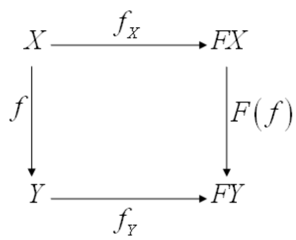
We can mention this analogy in our paper.

3.9. Theorem. *If X and Y are Tychonoff spaces, then the following are equivalent for a map $f : X \to Y$;*

(1) $F(f)[FX \backslash X]$ is contained in $FY \backslash Y$.
(2) The diagram

is pullback.

Proof. (1) $\Rightarrow$ (2) Suppose that $h : Z \to FX$ and $g : Z \to Y$ are mapping such that $F(f) \circ h = f_Y \circ g$. Since $f_Y \circ g[Z]$ is contained in FY and $F(f)$ sends $FX \backslash X$ into $FY \backslash Y$, we have that $h[Z]$ is contained in X. Hence, defining $I : Z \to X$ by $I(z) = h(z)$,



it is shown that the square is pullback.(2) $\Rightarrow$ (1) Choose p in FX and assume that $F(f)(p) = y$ belongs to Y. Then, let h be the map which embeds $\{p\}$ into FX and g be the map from the subspace $\{p\}$ which sends p to $F(f)(p)$. Then, $F(f) \circ h = f_Y \circ g$ so that there exist a map $I : \{p\} \to X$ such that $h = f_x \circ I$. Hence, p belongs to X. $\square$

References

[1] M. Baran, Separation Properties, Indian J. Pure and Appl. Math. 23, 333-342 (1992).

[2] M. Baran and H. Altındiş, T_2-objects in topological categories, Acta Math. Hungarica 71, 41-48 (1996).

[3] K. Belaid and O. Echi, $T_{(\alpha,\beta)}$-Space and the Wallman Compactification, International Journal of Mathematics and Mathematical Sciences Vol.2004, Issue 68, 3717-3735 (2004).

[4] C.Elmalı and T.Uğur, Fan-Gottesman Compactification of some specific space is Wallman-type compactification, Chaos,Soliton and Fractals, Vol.42, no.1, 17-19 (2009).

[5] K. Fan and N. Gottesman, On compactifications of Freudenthal and Wallman, Indag. Math. 14, 504-510 (1952).

[6] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. W. Mislove and D. S. Scott, Continuous Lattices and Domains, Cambridge University Press, Cambridge, England (2003).

[7] A. Grothendieck and J. Dieudonné, Éléments de Géometrie Algébrique, Die Grundlehren der mathematischen Wissenchaften, vol.166, Springer-Verlag, New York, (1971).

[8] D. Harris, The Wallman compactification is an epireflection, Proc. Amer. Math. Soc. 31, 265-267 (1972).

[9] H. Herrlich, On the concept of reflections in general topology, Contributions to Extension Theory of Topological Structures (Proc. Sympos., Berlin, 1967), Deutsch. Verlag Wissensch., Berlin, 105-114 (1969).

[10] S. Maclane, Categories for the Working Mathematician, Springer-Verlag, New York, (1971).

[11] S.Maclane and S.Eilenberg, Eilenberg-Maclane:collected works, Academic Press Inc., Orlando, FL, (1986).

[12] K.Morita, Category-isomorphism and endomorphism rings of modules, Trans. Amer. Math. Soc. 103, 451-469 (1962).

[13] K.Morita, Localization in categories of modules I. Math. Z. 114, 121-144 (1970).

[14] K.Morita, Localization in categories of modules II. J. Reine Angew. Math. 242, 163-169 (1970).

[15] K.Morita, Localization in categories of modules III. Math. Z. 119, 313-320 (1971).

[16] M. H. Stone, Applications of Boolean algebras to general topology, Bull. Amer. Math. Soc. 41, 375-481 (1997).

[17] H. Wallman, Lattices and topological spaces, Ann. of Math. (2) 39, 112-126 (1938).

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 923 – 933

On second-order linear recurrent homogeneous differential equations with period k

Julius Fergy Tiongson Rabago* †

Abstract

We say that $w(x)\colon \mathbb{R} \to \mathbb{C}$ is a solution to a second-order linear recurrent homogeneous differential equation with period k ($k \in \mathbb{N}$), if it satisfies a homogeneous differential equation of the form

$$w^{(2k)}(x) = pw^{(k)}(x) + qw(x), \quad \forall x \in \mathbb{R},$$

where $p, q \in \mathbb{R}^+$ and $w^{(k)}(x)$ is the k^{th} derivative of $w(x)$ with respect to x. On the other hand, $w(x)$ is a solution to an odd second-order linear recurrent homogeneous differential equation with period k if it satisfies

$$w^{(2k)}(x) = -pw^{(k)}(x) + qw(x), \quad \forall x \in \mathbb{R}.$$

In the present paper, we give some properties of the solutions of differential equations of these types. We also show that if $w(x)$ is the general solution to a second-order linear recurrent homogeneous differential equation with period k (resp. odd second-order linear recurrent homogeneous differential equation with period k), then the limit of the quotient $w^{((n+1)k)}(x)/w^{(n)}(x)$ as n tends to infinity exists and is equal to the positive (resp. negative) dominant root of the quadratic equation $x^2 - px - q = 0$ as x increases (resp. decreases) without bound.

Keywords: Homogenous differential equations, second-order linear recurrence sequences, solutions.

2000 AMS Classification: Primary: 34 B 05, 11 B 37. Secondary: 11 B 39.

Received 22 : 07 : 2013 : Accepted 05 : 10 : 2013 Doi : 10.15672/HJMS.2014437531

*Institute of Mathematics, College of Science, University of the Philippines Diliman, Quezon City 1101, PHILIPPINES

†Department of Mathematics and Computer Science, College of Science, University of the Philippines Baguio, Governor Pack Road, Baguio City 2600, PHILIPPINES

Email: `jtrabago@upd.edu.ph, jfrabago@gmail.com`

1. Introduction

Problems involving Fibonacci numbers and its various generalizations have been extensively studied by many authors. Its beauty and applications have been greatly appreciated since its introduction. In 1965, a certain generalization of the sequence of Fibonacci numbers was introduced by A. F. Horadam in [1], which is called as a second-order linear recurrence sequence and is now known as *Horadam sequence*. Properties of these type of sequences have also been studied by Horadam in [1]. In [2], J. S. Han, H. S. Kim, and J. Neggers studied a Fibonacci norm of positive integers. These authors [3] have also studied Fibonacci sequences in groupoids and introduced the concept of Fibonacci functions in [4]. They developed the notion of this type of functions using the concept of f-even and f-odd functions. Later on, a certain generalization of Fibonacci function has been investigated by B. Sroysang in [5]. In particular, Sroysang defined a function $f(x)\colon \mathbb{R} \to \mathbb{R}$ as a Fibonacci function of period $k, (k \in \mathbb{N})$ if it satisfies the equation $f(x+2k) = f(x+k) + f(x)$ for all $x \in \mathbb{R}$. Recently, the notion of Fibonacci function has been further generalized by the author in [6]. The concept of second-order linear recurrent functions with period k which has been introduced by the author in [6] gave rise to the concept of Pell and Jacobsthal functions with period k, which are analogues of Fibonacci functions. Some elementary properties of these newly defined functions were also presented by the author in [6]. Now, inspired by these results, we present in this work the concept of second-order (resp. odd second-order) linear recurrent homogeneous differential equations with period k, or simply SOLRHDE-k (resp. oSOLRHDE-k), and study some of its properties.

The next section, which discusses our main results, is organized as follows. First, we present some elementary results on second-order (and odd second-order) linear recurrent homogeneous differential equation with period k, and then provide the form of its general solution. Afterwards, we investigate the quotient $w^{((n+1)k)}(x)/w^{(n)}(x)$, where $w(x)$ is the general solution to a SOLRHDE-k (or an oSOLRHDE-k), and find its limit as n tends to infinity. Each of our results is accompanied by an example for validation and illustration.

2. Main Results

We start-off this section with the following definition.

2.1. Definition. Let $k \in \mathbb{N}$, $p, q \in \mathbb{R}^+$ and $w\colon \mathbb{R} \to \mathbb{C}$ be differentiable on $\mathbb{R}$ infinitely many times. We say that $w(x)$ is a solution to a SOLRHDE-k if it satisfies a differential equation of the form given by

$$w^{(2k)}(x) = pw^{(k)}(x) + qw(x), \tag{2.1}$$

for all $x \in \mathbb{R}$, where $w^{(k)}(x)$ is the k^{th} derivative of $w(x)$ with respect to x. If $(p,q) = (1,1), (1,2), (2,1)$, then w is a solution to a Fibonacci-like, Jacobsthal-like, and Pell-like homogeneous differential equation with period k, respectively.

2.2. Example. Let $p, q \in \mathbb{R}^+$ and $0 \neq t \in \mathbb{R}$. Define $w(x) = a^{tx}$, where $a > 0$. Suppose that $w(x)$ is a solution to a SOLRHDE-k then $(t \ln a)^{2k} a^{tx} = p(t \ln a)^k a^{tx} + qa^{tx}$. Hence, $r^2 - pr - q = 0$ where $r = (t \ln a)^k$. Solving for r, we have $r = (p \pm \sqrt{p^2 + 4q})/2$. So, $a = \exp\left(t^{-1}\Phi_\pm^{1/k}\right)$, where $\Phi_\pm = (p \pm \sqrt{p^2+4q})/2$. Thus, $w(x) = A\exp\left(\alpha^{1/k}x\right) + B\exp\left(\beta^{1/k}x\right)$, where $\alpha = \Phi_+$ and $\beta = \Phi_-$ and, A, B are any arbitrary real numbers. If we set $k = 1$, and $w(0) = 0$ and $w'(0) = 1$, then we get $A + B = 0$ and $\alpha A + \beta B = 1$. Here we obtain,

$$w(x) = \frac{1}{\alpha - \beta}\left(e^{\alpha x} - e^{\beta x}\right). \tag{2.2}$$

Thus, (2.2) is a solution to a SOLRHDE-k, with $k = 1$ and initial boundary conditions $w(0) = 0$ and $w'(0) = 1$. Using the identity $e^X = \sum_{n=0}^{\infty}(X^n/n!)$, we can express (2.2) in terms of power series, i.e. we have

$$w(x) = \frac{e^{\alpha x} - e^{\beta x}}{\alpha - \beta} = \sum_{n=0}^{\infty} \left(\frac{\alpha^n - \beta^n}{\alpha - \beta} \right) \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{W_n}{n!} x^n,$$

where W_n is the number sequence obtained from the recurrence relation given by

$$W_0 = 0, \quad W_1 = 1, \quad W_{n+1} = pW_n + qW_{n-1}, \quad \forall n \in \mathbb{N}. \tag{2.3}$$

We note that $\alpha + \beta = p, \alpha - \beta = \sqrt{p^2 + 4q}$, and $\alpha\beta = -q$. Hence, for some particular values of p and q, we have the following examples.

(1) For $(p, q) = (1, 1)$, the function defined by

$$f(x) = \frac{1}{\sqrt{5}} \left(e^{\phi x} - e^{(1-\phi)x} \right) = \sum_{n=0}^{\infty} \frac{F_n}{n!} x^n,$$

where ϕ is the golden ratio and F_n is the n^{th} Fibonacci number, is a solution to a Fibonacci-like homogeneous differential equation. By letting $x = 1$, we obtain the identity

$$\sum_{n=0}^{\infty} \frac{F_n}{n!} = \frac{e^{\phi} - e^{1-\phi}}{\sqrt{5}}.$$

(2) For $(p, q) = (1, 2)$, the function defined by

$$j(x) = \frac{1}{3} \left(e^{2x} - e^{-x} \right) = \sum_{n=0}^{\infty} \frac{J_n}{n!} x^n,$$

where J_n is the n^{th} Jacobsthal number, is a solution to a Jacobsthal-like homogeneous differential equation. By letting $x = 1$, we obtain the identity

$$\sum_{n=0}^{\infty} \frac{J_n}{n!} = \frac{e^2 - e^{-1}}{3}.$$

(3) For $(p, q) = (2, 1)$, the function defined by

$$p(x) = \frac{1}{2\sqrt{2}} \left(e^{\sigma x} - e^{(2-\sigma)x} \right) = \sum_{n=0}^{\infty} \frac{P_n}{n!} x^n,$$

where σ is the silver ratio and P_n is the n^{th} Pell number, is a solution to a Pell-like homogeneous differential equation. By letting $x = 1$, we obtain the identity

$$\sum_{n=0}^{\infty} \frac{P_n}{n!} = \frac{e^{\sigma} - e^{2-\sigma}}{2\sqrt{2}}.$$

2.3. Proposition. *Let $k \in \mathbb{N}$, $p, q, \in \mathbb{R}^+$ and $w(x)$ be a solution to the differential equation* (2.1)*. If $g_m(x) := w^{(m)}(x)$, then $g(x)$ is also a solution to* (2.1).

Proof. Let $k \in \mathbb{N}$ and $p, q, \in \mathbb{R}^+$. Suppose $g_m(x) = w^{(m)}(x)$ where $w(x)$ is a solution to (2.1). Then,

$$g_m^{(2k)}(x) = \frac{d^{2k} \left[w^{(m)}(x) \right]}{dx^{2k}} = p \frac{d^m \left[w^{(k)}(x) \right]}{dx^m} + q \frac{d^m \left[w(x) \right]}{dx^m} = p g_m^{(k)}(x) + q g_m(x),$$

proving the proposition. $\square$

2.4. Example. Let $j(x) = e^{(-1)^{1/k}x}$ where $k \in \mathbb{N}$. It can be verified easily that $j(x) = e^{(-1)^{1/2}x} = e^{\pm ix}$ is a solution to a Jacobsthal-like homogeneous differential equation with period 2, *i.e.*

$$j^{(4)}(x) = e^{\pm ix} = -e^{\pm ix} + 2e^{\pm ix} = j''(x) + 2j(x), \quad \forall x \in \mathbb{R}.$$

Now, define $g(x) = \pm ie^{\pm ix}$. We show that $g(x)$ is also a solution to a Jacobsthal-like homogeneous differential equation with period 2, *i.e.*

$$g^{(4)}(x) = g''(x) + 2g(x), \quad \forall x \in \mathbb{R}.$$

We note that,

$$g'(x) = -e^{\pm ix}, \quad g''(x) = \mp ie^{\pm ix}, \quad g'''(x) = e^{\pm ix}, \quad g^{(4)}(x) = \pm ie^{\pm ix}.$$

Hence,

$$g^{(4)}(x) = \pm ie^{\pm ix} = \mp ie^{\pm ix} + 2 \pm ie^{\pm ix} = g''(x) + 2g(x).$$

We can also show this via Proposition (2.3). Since $g(x) = j'(x)$, and $j(x)$ is a solution to a Jacosthal-like homogeneous differential equation with period 2, then so is $g(x)$ by Proposition (2.3).

2.5. Proposition. *Let $k \in \mathbb{N}$, $p, q, \in \mathbb{R}^+$ and, $g(x)$ and $h(x)$ be any two solutions of the differential equation* (2.1). *Then, any linear combination of $g(x)$ and $h(x)$, say $w(x) = Ag(x) + Bh(x)$ where $A, B \in \mathbb{R}$, is again a solution to* (2.1).

Proof. The proof is straightforward. Let $k \in \mathbb{N}, p, q, \in \mathbb{R}^+$, and $g(x)$ and $h(x)$ be any two solutions to the differential equation (2.1). Consider the function $w(x) = Ag(x) + Bh(x)$ where $A, B \in \mathbb{R}$. Then,

$$\begin{aligned} w^{(2k)}(x) &= Ag^{(2k)}(x) + Bh^{(2k)}(x) \\ &= p\left[Ag^{(k)}(x) + Bh^{(k)}(x)\right] + q\left[Ag(x) + Bh(x)\right] \\ &= pw^{(k)}(x) + qw(x). \end{aligned}$$

This proves the proposition. □

2.6. Example. Let $j(x) = e^{(-1)^{1/k}x}$ where $k \in \mathbb{N}$. It can be verified diretly that the function $j(x) = e^{(-1)^{1/3}x} = e^{tx}$, where $t \in \{-1, (1 \pm \sqrt{3}i)/2\}$, is a solution to a Jacobsthal-like homogeneous differential equation with period 3, *i.e.*

$$j^{(6)}(x) = j'''(x) + 2j(x), \quad \forall x \in \mathbb{R}, \tag{2.4}$$

Define $w(x) = Ae^{-x} + Be^{\frac{1}{2}(1\pm\sqrt{3})ix}$, where $A, B \in \mathbb{R}$. Then,

$$\begin{aligned} w^{(6)}(x) &= Ae^{-x} + Be^{\frac{1}{2}(1\pm\sqrt{3})ix} \\ &= -\left[Ae^{-x} + Be^{\frac{1}{2}(1\pm\sqrt{3}i)x}\right] + 2\left[Ae^{-x} + Be^{\frac{1}{2}(1\pm\sqrt{3}i)x}\right] \\ &= w'''(x) + 2w(x). \end{aligned}$$

In fact, this can also be shown using Proposition (2.5). Since $g(x) = e^{-x}$ and $h(x) = \exp(\frac{1}{2}(1 \pm \sqrt{3})ix)$ are solutions of (2.4), then the function defined by $w(x) = Ag(x) + Bh(x)$, where $A, B \in \mathbb{R}$, is also a solution to (2.4) by Proposition (2.5).

2.7. Theorem. *Let $k \in \mathbb{N}$, $p, q, \in \mathbb{R}^+$ and $w(x)$ be a solution to the differential equation* (2.1). *Furthermore, let $\{W_n\}_{n=0}^{\infty}$ be a number sequence obtained from a second-order linear recurrence relation defined by* (2.3). *Then,*

$$w^{(nk)}(x) = W_n w^{(k)}(x) + qW_{n-1}w(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}. \tag{2.5}$$

Proof. We prove this using induction on n. Let $k \in \mathbb{N},\ p, q \in \mathbb{R}^+$, and $w(x)$ be a solution to the differential equation (2.1). Then,

$$\begin{aligned}
w^{(k)}(x) &= (1)w^{(k)}(x) + q(0)w(x) = W_1 w^{(k)}(x) + qW_0 w(x),\\
w^{(2k)}(x) &= pw^{(k)}(x) + q(1)w(x) = W_2 w^{(k)}(x) + qW_1 w(x),\\
w^{(3k)}(x) &= \frac{d^k}{dx^k}\left(w^{(2k)}(x)\right) = pw^{(2k)}(x) + qw^{(k)}(x)\\
&= p\left[pw^{(k)}(x) + qw(x)\right] + qw^{(k)}(x)\\
&= (p^2+q)w^{(k)}(x) + qpw(x)\\
&= W_3 w^{(k)}(x) + qW_2 w(x).
\end{aligned}$$

Now we assume that the following equation is true for some natural number n,

$$w^{(nk)}(x) = W_n w^{(k)}(x) + qW_{n-1} w(x).$$

Hence,

$$\begin{aligned}
w^{((n+1)k)}(x) &= \frac{d^k}{dx^k}\left[w^{(nk)}\right] = \frac{d^k}{dx^k}\left[W_n w^{(k)}(x) + qW_{n-1}w(x)\right]\\
&= W_n w^{(2k)}(x) + qW_{n-1} w^{(k)}(x)\\
&= W_n\left[pw^{(k)}(x) + qw(x)\right] + qW_{n-1}w^{(k)}(x)\\
&= \left(pW_n + qW_{n-1}\right) w^{(k)}(x) + qW_n w(x)\\
&= W_{n+1} w^{(k)}(x) + qW_n w(x).
\end{aligned}$$

This proves the theorem. □

2.8. Corollary. *Let $k \in \mathbb{N}$ and $f(x)$ be a solution to a Fibonacci-like differential equation with period k. If $\{F_n\}_{n=0}^{\infty}$ is the sequence of Fibonacci numbers, then*

$$f^{(nk)}(x) = F_n f^{(k)}(x) + F_{n-1} f(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.9. Example. Consider the solution $f(x) = e^{\sqrt[4]{\phi}x}$ to a Fibonacci-like differential equation with period 4 given by the equation

$$f^{(8)}(x) = f^{(4)}(x) + f(x), \quad \forall x \in \mathbb{R}.$$

Furthermore, let $\{F_n\}$ be the sequence of Fibonacci numbers. By Corollary (2.8), we see that

$$\begin{aligned}
f^{(12)}(x) &= (2+\sqrt{5})e^{\sqrt[4]{\phi}x} = 2\phi e^{\sqrt[4]{\phi}x} + e^{\sqrt[4]{\phi}x} = F_3 f^{(4)}(x) + F_2 f(x),\\
f^{(16)}(x) &= \frac{1}{2}(7+3\sqrt{5})e^{\sqrt[4]{\phi}x} = 3\phi e^{\sqrt[4]{\phi}x} + 2e^{\sqrt[4]{\phi}x} = F_4 f^{(4)}(x) + F_3 f(x).
\end{aligned}$$

Similarly, for Jacobsthal-like and Pell-like differential equations with period k we have the following corollaries.

2.10. Corollary. *Let $k \in \mathbb{N}$ and $j(x)$ be a solution to a Jacobsthal-like differential equation with period k. If $\{J_n\}_{n=0}^{\infty}$ is the sequence of Jacobsthal numbers, then*

$$j^{(nk)}(x) = J_n j^{(k)}(x) + 2J_{n-1} j(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.11. Example. Consider the solution $j(x) = e^{-x}$ to a Jacobsthal-like differential equation given by

$$j''(x) = j'(x) + 2j(x), \quad \forall x \in \mathbb{R}.$$

Furthermore, let $\{J_n\}_{n=0}^{\infty}$ be the sequence of Jacobsthal numbers, *i.e.* $\{J_n\} = \{0, 1, 1, 3, 5, 11, 21, 43, 85, 171, \dots\}$. By Corollary (2.10), we see that

$$j^{(7)}(x) = -e^{-x} = 43(-e^{-x}) + 2(21)e^{-x} = J_7 j'(x) + 2J_6 j(x),$$
$$j^{(8)}(x) = e^{-x} = 85(-e^{-x}) + 2(43)e^{-x} = J_8 j'(x) + 2J_7 j(x),$$
$$j^{(9)}(x) = -e^{-x} = 171(-e^{-x}) + 2(85)e^{-x} = J_9 j'(x) + 2J_8 j(x).$$

2.12. Corollary. *Let $k \in \mathbb{N}$ and $p(x)$ be a solution to a Pell-like differential equation with period k. If $\{P_n\}_{n=0}^{\infty}$ is the sequence of Pell numbers, then*

$$p^{(nk)}(x) = P_n p^{(k)}(x) + P_{n-1} p(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.13. Example. Consider the solution $p(x) = e^{\sqrt[3]{\sigma}x}$ to a Pell-like differential equation with period 3 given by the equation

$$p^{(6)}(x) = 2p'''(x) + p(x), \quad \forall x \in \mathbb{R}. \tag{2.6}$$

Furthermore, let $\{P_n\}_{n=0}^{\infty}$ be the sequence of Pell numbers, *i.e.* $\{P_n\} = \{0, 1, 2, 5, 12, 29, \dots\}$. By Corollary (2.12), we see that

$$p^{(9)}(x) = (7 + 5\sqrt{2})e^{\sqrt[3]{\sigma}x} = 5\sigma e^{\sqrt[3]{\sigma}x} + 2e^{\sqrt[3]{\sigma}x} = P_3 p'''(x) + P_2 p(x),$$
$$p^{(12)}(x) = (17 + 12\sqrt{2})e^{\sqrt[3]{\sigma}x} = 12\sigma e^{\sqrt[3]{\sigma}x} + 5e^{\sqrt[3]{\sigma}x} = P_4 p'''(x) + P_3 p(x),$$
$$p^{(15)}(x) = (41 + 29\sqrt{2})e^{\sqrt[3]{\sigma}x} = 29\sigma e^{\sqrt[3]{\sigma}x} + 12e^{\sqrt[3]{\sigma}x} = P_5 p'''(x) + P_4 p(x).$$

In solving for the solution of equation (2.6), we obtain an approximation of the golden ratio involving the silver ratio σ. In particular, we obtain

$$\phi \approx 10\left(\sqrt[3]{\sigma}\sin(2\pi/3) - 1\right).$$

This gives us a motivation to obtain a better approximation which is given by

$$\phi \approx 10\left(\sqrt[3]{\sigma}\sin\left(\frac{2^{20}\cdot 5^6 - 315611}{2^{19}\cdot 3\cdot 5^6}\pi\right) - 1\right).$$

Looking at this approximation, it might be interesting to get a better approximation of ϕ in terms of σ by altering the coefficient of π inside the sine function.

2.14. Corollary. *Let $k = 1$, $p, q, \in \mathbb{R}^+$ and $w(x) = e^{\alpha x}$ be a solution to* (2.1)*. Furthermore, let $\{W_n\}_{n=0}^{\infty}$ be a number sequence obtained from* (2.3)*. Then,*

$$\alpha^n = \alpha W_n + qW_{n-1}, \quad \forall n \in \mathbb{N}. \tag{2.7}$$

Furthermore, if $\{F_n\}, \{J_n\}$, and $\{P_n\}$ are the sequence of Fibonacci, Jacobsthal and Pell numbers, respectively, then

$$\phi^n = \phi F_n + F_{n-1}, \quad \forall n \in \mathbb{N}, \tag{2.8}$$
$$2^{n-1} = J_n + J_{n-1}, \quad \forall n \in \mathbb{N}, \tag{2.9}$$
$$\sigma^n = 2\sigma P_n + P_{n-1}, \quad \forall n \in \mathbb{N}, \tag{2.10}$$

where ϕ and σ are the golden and silver ratio, respectively.

Proof. We note that $w(x) = e^{\alpha x}$ is a solution to equation (2.1) with period $k = 1$. So, by Theorem (2.7), we have

$$\alpha^n e^{\alpha x} = \alpha W_n e^{\alpha x} + qW_{n-1}e^{\alpha x},$$

proving equation (2.7). By letting $(p, q) = (1, 1), (1, 2), (2, 1)$, we obtain equations (2.8), (2.9), and (2.10), respectively. □

In the following discussion, we study differential equations of the form

$$w^{(2k)}(x) = -pw^{(k)} + qw(x), \quad \forall x \in \mathbb{R}, \tag{2.11}$$

where $k \in \mathbb{N}$ and $p, q \in \mathbb{R}^+$. We call such equation as an *odd second-order linear recurrent homogeneous differential equation with period* k, or simply, oSOLRHDE-k.

Solving equation (2.11) we obtain the solution

$$w(x) = ae^{\alpha^{1/k}\zeta_n x} + be^{\beta^{1/k}\zeta_n x},$$

where $\zeta_n = \cos\left(\frac{\pi+2n\pi}{k}\right) + i\sin\left(\frac{\pi+2n\pi}{k}\right),\ n = 0, 1, \ldots, k-1$, and $a, b \in \mathbb{R}$. If $(p, q, k) = (1, 1, 1)$, then we see that $f(x) = e^{-\phi x}$ is a solution to the following differential equation

$$w''(x) = -w'(x) + w(x), \quad \forall x \in \mathbb{R}.$$

Similarly, for $(p, q, k) = (1, 2, 1), (2, 1, 1)$, we see that the functions $j(x) = e^{-2x}$ and $p(x) = e^{-\sigma x}$ are solutions to the differential equations

$$j''(x) = -j'(x) + 2j(x), \quad \forall x \in \mathbb{R},$$
$$p''(x) = -2p'(x) + p(x), \quad \forall x \in \mathbb{R},$$

respectively. Also, if $(p, q, k) = (1, 1, 3)$, then the function defined by $f(x) = e^{tx}$, where $t \in \{-\sqrt[3]{\phi}, \sqrt[3]{\phi}(1 \pm \sqrt{3}i)/2\}$, is a solution to an odd Fibonacci-like homogeneous differential equation with period 3. *i.e.*, $f(x) = e^{tx}$ is a solution to

$$f^{(6)}(x) = -f^{(3)}(x) + f(x), \quad \forall x \in \mathbb{R}. \tag{2.12}$$

2.15. Theorem. *Let $k \in \mathbb{N},\ p, q, \in \mathbb{R}^+$ and $w(x)$ be a solution to the differential equation (2.11). Furthermore, let $\{W_{-n}\}_{n=0}^{\infty}$, where $W_{-n} = (-1)^{n+1}W_n$ be a number sequence obtained from a second-order linear recurrence relation defined by*

$$W_0 = 0, \quad W_{-1} = 1, \quad W_{-(n+1)} = -pW_{-n} + qW_{-n+1}, \quad \forall n \in \mathbb{N}. \tag{2.13}$$

Then,

$$w^{(nk)}(x) = W_{-n}w^{(k)}(x) + qW_{-n+1}w(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}. \tag{2.14}$$

Proof. We follow the proof of Theorem (2.7). Let $k \in \mathbb{N},\ p, q, \in \mathbb{R}^+$, and $w(x)$ be a solution to the differential equation (2.11). Then,

$$\begin{aligned}
w^{(k)}(x) &= (1)w^{(k)}(x) + q(0)w(x) = W_{-1}w^{(k)}(x) + qW_0 w(x),\\
w^{(2k)}(x) &= -pw^{(k)}(x) + q(1)w(x) = W_{-2}w^{(k)}(x) + qW_{-1}w(x),\\
w^{(3k)}(x) &= \frac{d^k}{dx^k}\left(w^{(2k)}(x)\right) = -pw^{(2k)}(x) + qw^{(k)}(x)\\
&= -p\left[-pw^{(k)}(x) + qw(x)\right] + qw^{(k)}(x)\\
&= (p^2 + q)w^{(k)}(x) + qpw(x)\\
&= W_{-3}w^{(k)}(x) + qW_{-2}w(x).
\end{aligned}$$

Now we assume that the following equation is true for some natural number n,

$$w^{(nk)}(x) = W_{-n}w^{(k)}(x) + qW_{-n+1}w(x).$$

Hence,

$$\begin{aligned} w^{((n+1)k)}(x) &= \frac{d^k}{dx^k}\left[w^{(nk)}\right] = \frac{d^k}{dx^k}\left[W_{-n}w^{(k)}(x) + qW_{-n+1}w(x)\right] \\ &= W_{-n}w^{(2k)}(x) + qW_{-n+1}w^{(k)}(x) \\ &= W_{-n}\left[-pw^{(k)}(x) + qw(x)\right] + qW_{-n+1}w^{(k)}(x) \\ &= \left(-pW_{-n} + qW_{-n+1}\right)w^{(k)}(x) + qW_{-n}w(x) \\ &= W_{-(n+1)}w^{(k)}(x) + qW_{-n}w(x), \end{aligned}$$

proving the theorem. □

2.16. Corollary. *Let $k \in \mathbb{N}$ and $f(x)$ be a solution to an odd Fibonacci-like differential equation with period k. If $\{F_n\}_{n=0}^{\infty}$ is the sequence of Fibonacci numbers then,*

$$f^{(nk)}(x) = F_{-n}f^{(k)}(x) + F_{-n+1}f(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.17. Example. Consider the solution $f(x) = e^{(\sqrt[3]{\phi}/2)(1+\sqrt{3}i)x}$ to the differential equation (2.12). By Corollary (2.16), we see that

$$\begin{aligned} f^{(15)}(x) &= -\frac{1}{2}(11 + 5\sqrt{5})e^{(\sqrt[3]{\phi}/2)(1+\sqrt{3}i)x} \\ &= -5\phi e^{(\sqrt[3]{\phi}/2)(1+\sqrt{3}i)x} + -3e^{(\sqrt[3]{\phi}/2)(1+\sqrt{3}i)x} \\ &= F_{-5}f^{(3)}(x) + F_{-4}f(x). \end{aligned}$$

2.18. Corollary. *Let $k \in \mathbb{N}$ and $j(x)$ be a solution to an odd Jacobsthal-like differential equation with period k. If $\{J_n\}_{n=0}^{\infty}$ is the sequence of Jacobsthal numbers then,*

$$j^{(nk)}(x) = J_{-n}j^{(k)}(x) + 2J_{-n+1}j(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.19. Example. Consider the solution $j(x) = e^{-\sqrt[5]{2}x}$ to the odd Jacobsthal-like differential equation with period 5 given by

$$j^{(10)}(x) = -j^{(5)}(x) + 2j(x), \quad \forall x \in \mathbb{R}.$$

By Corollary (2.18), we see that

$$j^{(25)}(x) = -32e^{-\sqrt[5]{2}x} = 11(-2e^{-\sqrt[5]{2}x}) + 2(-5)e^{-\sqrt[5]{2}x} = J_{-5}j^{(3)}(x) + 2J_{-4}f(x).$$

2.20. Corollary. *Let $k \in \mathbb{N}$ and $p(x)$ be a solution to an odd Pell-like differential equation with period k. If $\{P_n\}_{n=0}^{\infty}$ is the sequence of Pell numbers then,*

$$p^{(nk)}(x) = P_{-n}p^{(k)}(x) + P_{-n+1}p(x), \quad \forall x \in \mathbb{R},\ n \in \mathbb{N}.$$

2.21. Theorem. *Let $k \in \mathbb{N}$, $p, q \in \mathbb{R}^+$, and consider the SOLRHDE-k defined by (2.1). Then,*

$$\Omega_{W,k}(x) = \sum_{j=1}^{k}\left(c_j e^{r_j x} + \bar{c}_j e^{t_j x}\right), \quad \forall x \in \mathbb{R}, \tag{2.15}$$

where $c_j, \bar{c}_j \in \mathbb{R}$ and, r_j and t_j, for all $j = 1, 2, \ldots, k$ are roots of α and β, respectively, is the general solution of the given homogeneous differential equation.

Proof. Let $\{r_j\}_{j=1}^{k}$ and $\{t_j\}_{j=1}^{k}$ be the set of k^{th} roots of α and β, *i.e.*

$$r_j = |\alpha|^{1/k}\left[\cos\left(\frac{\theta_r + 2\pi j}{k}\right) + i\sin\left(\frac{\theta_r + 2\pi j}{k}\right)\right],$$

and

$$t_j = |\beta|^{1/k}\left[\cos\left(\frac{\theta_t + 2\pi j}{k}\right) + i\sin\left(\frac{\theta_t + 2\pi j}{k}\right)\right],$$

where $j = 1, 2, \ldots, k$, $\theta_r = \arg(\alpha)$ and $\theta_t = \arg(\beta)$. Note that $r_{j's}$ and $t_{j's}$ are all distinct then, $\{e^{r_1 x}, e^{r_2 x}, \ldots, e^{r_k x}\}$ and $\{e^{t_1 x}, e^{t_2 x}, \ldots, e^{t_k x}\}$ are linearly independent sets of solutions of the homogeneous linear equation defined in (2.1). Hence, by Proposition (2.5), conclusion follows. □

2.22. Example. Consider the Jacobsthal-like homogeneous differential equation (2.4) with period 3. By Theorem (2.21), we have the general solution

$$\begin{aligned}\Omega_{J,3}(x) &= c_1 e^{\sqrt[3]{2}x} + c_2 e^{-\frac{1}{2}\sqrt[3]{2}(1+\sqrt{3}i)x} + c_3 e^{-\frac{1}{2}\sqrt[3]{2}(1-\sqrt{3}i)x} \\ &\quad + \bar{c}_1 e^{-x} + \bar{c}_2 e^{\frac{1}{2}(1+\sqrt{3}i)x} + \bar{c}_3 e^{\frac{1}{2}(1-\sqrt{3}i)x}.\end{aligned}$$

Also, if ϕ and σ are the golden ratio and silver ratio, respectively, then the general solution to a Fibonacci-like and Pell-like homogeneous differential equation are given by

$$\Omega_{F,k}(x) = \sum_{j=1}^{k} c_j \exp\left(\phi^{1/k}\Theta_{2j}x\right) + \sum_{j=1}^{k} \bar{c}_j \exp\left((\phi-1)^{1/k}\Theta_{2j+1}x\right)$$

and

$$\Omega_{P,k}(x) = \sum_{j=1}^{k} c_j \exp\left(\sigma^{1/k}\Theta_{2j}x\right) + \sum_{j=1}^{k} \bar{c}_j \exp\left((2-\sigma)^{1/k}\Theta_{2j+1}x\right),$$

where $\Theta_m = \cos(m\pi/k) + i\sin(m\pi/k)$ and $c_{j's}, \bar{c}_{j's} \in \mathbb{R}$, for all $x \in \mathbb{R}$, respectively.

In the rest of our discussion, we investigate the quotient of solutions of a second-order linear recurrent homogeneous differential equation with period k.

2.23. Theorem. *Let $p, q \in \mathbb{R}^+$ and $k \in \mathbb{N}$ be the period of a SOLRHDE-k defined in (2.1) and let $w(x)$ be its general solution. Then, the limit $\lim_{n\to\infty}\frac{w^{((n+1)k)}(x)}{w^{(n)}(x)}$ exists and is given by*

$$\lim_{n\to\infty}\frac{w^{((n+1)k)}(x)}{w^{(n)}(x)} = \alpha \text{ (resp. } \beta), \quad \text{as} \quad x \to \infty \text{ (resp. } x \to -\infty), \tag{2.16}$$

where α and β are the roots of the quadratic equation $x^2 - px - q = 0$. Particularly, if $f(x), j(x)$, and $p(x)$ are solutions to a Fibonacci-like, Jacobsthal-like, and Pell-like homogeneous differential equation with period k, respectively, then

$$\lim_{n\to\infty}\frac{f^{((n+1)k)}(x)}{f^{(n)}(x)} = \phi \text{ (resp. } 1-\phi), \quad \text{as} \quad x \to \infty \text{ (resp. } x \to -\infty) \tag{2.17}$$

$$\lim_{n\to\infty}\frac{j^{((n+1)k)}(x)}{j^{(n)}(x)} = 2 \text{ (resp. } -1), \quad \text{as} \quad x \to \infty \text{ (resp. } x \to -\infty) \tag{2.18}$$

$$\lim_{n\to\infty}\frac{p^{((n+1)k)}(x)}{p^{(n)}(x)} = \sigma \text{ (resp. } 1-\sigma), \quad \text{as} \quad x \to \infty \text{ (resp. } x \to -\infty). \tag{2.19}$$

Proof. Let $k, n \in \mathbb{N}$, $p, q \in \mathbb{R}^+$, and consider the quotient $Q(x) := \frac{\omega^{(k)}(x)}{\omega(x)}$, where $\omega(x) = w^{(nk)}(x)$ satisfies a SOLRHDE-k. We suppose $x \to \infty$. The case when $x \to -\infty$ can be proven in a similar fashion.

We consider two cases: (i) $Q(x) < 0$, and (ii) $Q(x) > 0$.

CASE 1. Suppose that $Q(x) < 0$. Hence, we can assume without loss of generality (WLOG) that $\omega(x) > 0$ and $\omega^{(k)}(x) < 0$. By assumption, $w(x)$ satisifes (2.1), so we have

$$\begin{aligned}
w^{(2k)}(x) &= -pw^{(k)}(x) + qw(x),\\
w^{(3k)}(x) &= pw^{(2k)}(x) - qw^{(k)}(x) = p(-pw^{(k)}(x) + qw(x)) - qw^{(k)}(x)\\
&= -(p^2+q)w^{(k)}(x) + pqw(x),\\
w^{(4k)}(x) &= pw^{(3k)}(x) + qw^{(2k)}(x)\\
&= p(-(p^2+q)w^{(k)}(x) + pqw(x)) + q(-pw^{(k)}(x) + qw(x))\\
&= -(p^3+2pq)w^{(2k)}(x) + q(p^2+q)w^{(k)}(x),\\
&\ \ \vdots\\
w^{(nk)}(x) &= -W_n w^{(k)}(x) + qW_{n-1}w(x), \quad \forall n \in \mathbb{N},
\end{aligned}$$

where W_n is the number sequence satisfying equation (2.3). We let $\omega(x) = w^{(nk)}(x)$. Hence, by Proposition (2.3), $\omega(x)$ is also a solution to (2.1). It follows that

$$\begin{aligned}
\frac{\omega^{(k)}(x)}{\omega(x)} &= \frac{1}{w^{(nk)}(x)}\frac{d^k}{dx^k}\left(w^{(nk)}(x)\right) = \frac{-W_{n+1}w^{(k)}(x) + qW_n w(x)}{-W_n w^{(k)}(x) + qW_{n-1}w(x)}\\
&= \frac{-w^{(k)}(x)\frac{W_{n+1}}{W_n} + qw(x)}{-w^{(k)}(x) + qw(x)\frac{W_{n-1}}{W_n}}.
\end{aligned}$$

So we have

$$\begin{aligned}
\lim_{n\to\infty}\frac{\omega^{(k)}(x)}{\omega(x)} &= \lim_{n\to\infty}\frac{-w^{(k)}(x)\frac{W_{n+1}}{W_n} + qw(x)}{-w^{(k)}(x) + qw(x)\frac{W_{n-1}}{W_n}}\\
&= \frac{-w^{(k)}(x)\left(\lim_{n\to\infty}\frac{W_{n+1}}{W_n}\right) + qw(x)}{-w^{(k)}(x) + qw(x)\left(\lim_{n\to\infty}\frac{W_{n-1}}{W_n}\right)}.
\end{aligned}$$

Since $\beta = (p - \sqrt{p^2+4q})/2 \in (-1, 0)$, then $\lim_{n\to\infty}\beta^n = 0$. Thus,

$$\lim_{n\to\infty}\frac{\omega^{(k)}(x)}{\omega(x)} = \frac{-\alpha w^{(k)}(x) + qw(x)}{-w^{(k)}(x) + \alpha^{-1}qw(x)} = \alpha < \infty,$$

because $\lim_{n\to\infty}\frac{W_{n+1}}{W_n} = \lim_{n\to\infty}\frac{\alpha^{n+1}-\beta^{n+1}}{\alpha^n-\beta^n} = \alpha$ and $\alpha > \beta$.

CASE 2. Suppose (WLOG) that $\omega(x)$ and $\omega^{(k)}(x)$ are both positive. By Proposition (2.3), $\omega(x) = w^{(nk)}(x)$ is also a solution to (2.1). Hence,

$$\begin{aligned}
\lim_{n\to\infty}\frac{\omega^{(k)}(x)}{\omega(x)} &= \lim_{n\to\infty}\frac{w^{((n+1)k)}(x)}{w^{(nk)}(x)} = \lim_{n\to\infty}\frac{W_{n+1}w^{(k)}(x) + qW_n w(x)}{W_n w^{(k)}(x) + qW_{n-1}w(x)}\\
&= \lim_{n\to\infty}\frac{w^{(k)}(x)\frac{W_{n+1}}{W_n} + qw(x)}{w^{(k)}(x) + qw(x)\frac{W_{n-1}}{W_n}}\\
&= \frac{w^{(k)}(x)\left(\lim_{n\to\infty}\frac{W_{n+1}}{W_n}\right) + qw(x)}{w^{(k)}(x) + qw(x)\left(\lim_{n\to\infty}\frac{W_{n-1}}{W_n}\right)}\\
&= \alpha.
\end{aligned}$$

By letting $(p, q) = (1, 1), (1, 2), (2, 1)$, we obtain equations (2.17), (2.18), and (2.19), respectively. This completes the proof of the theorem. □

We also have the following theorem for oSOLRHDE-k.

2.24. Theorem. *Let $p, q \in \mathbb{R}^+$ and $k \in \mathbb{N}$ be the period of an oSOLRHDE-k defined by* (2.11) *and let $w(x)$ be its solutions. Then, the limit $\lim_{n\to\infty} \frac{w^{((n+1)k)}(x)}{w^{(n)}(x)}$ exists and is given by*

$$\lim_{n\to\infty} \frac{w^{((n+1)k)}(x)}{w^{(n)}(x)} = -\beta \; (\textit{resp.} \; -\alpha), \quad \textit{as} \quad x \to \infty \; (\textit{resp.} \; x \to -\infty), \tag{2.20}$$

where α and β are the roots of the quadratic equation $x^2 - px - q = 0$. Particularly, if $f(x), j(x)$, and $p(x)$ are solutions to an odd Fibonacci-like, odd Jacobsthal-like, and odd Pell-like homogeneous differential equation with period k, respectively, then

$$\lim_{n\to\infty} \frac{f^{((n+1)k)}(x)}{f^{(n)}(x)} = -(1-\phi) \; (\textit{resp.} \; -\phi), \quad \textit{as} \quad x \to \infty \; (\textit{resp.} \; x \to -\infty)$$
$$\lim_{n\to\infty} \frac{j^{((n+1)k)}(x)}{j^{(n)}(x)} = 1 \; (\textit{resp.} \; -2), \quad \textit{as} \quad x \to \infty \; (\textit{resp.} \; x \to -\infty)$$
$$\lim_{n\to\infty} \frac{p^{((n+1)k)}(x)}{p^{(n)}(x)} = -(1-\sigma) \; (\textit{resp.} \; -\sigma), \quad \textit{as} \quad x \to \infty \; (\textit{resp.} \; x \to -\infty).$$

The proof of the above theorem follows the same argument as in the proof of Theorem (2.23), so we omit it.

References

[1] A. F. Horadam, Basic properties of certain generalized sequence of numbers, *Fibonacci Quarterly,* **3**, 1965, 161–176.

[2] J. S. Han, H. S. Kim and J. Neggers, The Fibonacci norm of a positive integer n-observations and conjectures, *Int. J. Number Theory,* **6,** 2010, 371–385.

[3] J. S. Han, H. S. Kim and J. Neggers, Fibonacci sequences in groupoids, *Adv. Differ. Equ.,* **19,** 2012.

[4] J. S. Han, H. S. Kim and J. Neggers, On Fibonacci functions with Fibonacci numbers, *Adv. Differ. Equ.,* **126,** 2012.

[5] B. Sroysang, On Fibonacci functions with period k, *Discrete Dynamics in Nature and Society,* Volume 2013, Article ID 418123, 4 pages.

[6] J. F. T. Rabago, On second-order linear recurrent functions with period k and proofs to two conjectures of Sroysang, submitted.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 935 – 942

Hermite-Hadamard type inequalities for harmonically convex functions

İmdat İşcan*

Abstract

The author introduces the concept of harmonically convex functions and establishes some Hermite-Hadamard type inequalities of these classes of functions.

2000 AMS Classification: Primary 26D15; Secondary 26A51

Keywords: Harmonically convex function, Hermite-Hadamard type inequality.

Received 17 : 07 : 2013 : Accepted 07 : 10 : 2013 Doi : 10.15672/HJMS.2014437519

1. Introduction

Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a convex function defined on the interval I of real numbers and $a, b \in I$ with $a < b$. The following inequality

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2} \tag{1.1}$$

holds. This double inequality is known in the literature as Hermite-Hadamard integral inequality for convex functions. Note that some of the classical inequalities for means can be derived from (1.1) for appropriate particular selections of the mapping f. Both inequalities hold in the reversed direction if f is concave. For some results which generalize, improve and extend the inequalities (1.1) we refer the reader to the recent papers (see [1, 2, 3, 4, 6, 5, 7]).

The main purpose of this paper is to introduce the concept of harmonically convex functions and establish some results connected with the right-hand side of new inequalities similar to the inequality (1.1) for these classes of functions. Some applications to special means of positive real numbers are also given.

*Department of Mathematics, Faculty of Arts and Sciences, Giresun University, 28100, Giresun, Turkey.
Email: imdati@yahoo.com, imdat.iscan@giresun.edu.tr

2. Main Results

2.1. Definition. Let $I \subset \mathbb{R} \setminus \{0\}$ be a real interval. A function $f : I \to \mathbb{R}$ is said to be harmonically convex, if

$$f\left(\frac{xy}{tx + (1-t)y}\right) \le tf(y) + (1-t)f(x) \tag{2.1}$$

for all $x, y \in I$ and $t \in [0,1]$. If the inequality in (1.1) is reversed, then f is said to be harmonically concave.

2.2. Example. Let $f : (0,\infty) \to \mathbb{R}$, $f(x) = x$, and $g : (-\infty, 0) \to \mathbb{R}$, $g(x) = x$, then f is a harmonically convex function and g is a harmonically concave function.

The following proposition is obvious from this example:

2.3. Proposition. *Let $I \subset \mathbb{R} \setminus \{0\}$ be a real interval and $f : I \to \mathbb{R}$ is a function, then ;*

- *if $I \subset (0,\infty)$ and f is convex and nondecreasing function then f is harmonically convex.*
- *if $I \subset (0,\infty)$ and f is harmonically convex and nonincreasing function then f is convex.*
- *if $I \subset (-\infty, 0)$ and f is harmonically convex and nondecreasing function then f is convex.*
- *if $I \subset (-\infty, 0)$ and f is convex and nonincreasing function then f is a harmonically convex.*

The following result of the Hermite-Hadamard type holds.

2.4. Theorem. *Let $f : I \subset \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be a harmonically convex function and $a, b \in I$ with $a < b$. If $f \in L[a,b]$ then the following inequalities hold*

$$f\left(\frac{2ab}{a+b}\right) \le \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \le \frac{f(a)+f(b)}{2}. \tag{2.2}$$

The above inequalities are sharp.

Proof. Since $f : I \to \mathbb{R}$ is a harmonically convex function, we have, for all $x, y \in I$ (with $t = \frac{1}{2}$ in the inequality (2.1))

$$f\left(\frac{2xy}{x+y}\right) \le \frac{f(y)+f(x)}{2}.$$

Choosing $x = \frac{ab}{ta+(1-t)b}$, $y = \frac{ab}{tb+(1-t)a}$, we get

$$f\left(\frac{2ab}{a+b}\right) \le \frac{f\left(\frac{ab}{tb+(1-t)a}\right) + f\left(\frac{ab}{ta+(1-t)b}\right)}{2}.$$

Further, integrating for $t \in [0,1]$, we have

$$\begin{aligned} & f\left(\frac{2ab}{a+b}\right) \\ \le \; & \frac{1}{2}\left[\int_0^1 f\left(\frac{ab}{tb+(1-t)a}\right)dt + \int_0^1 f\left(\frac{ab}{ta+(1-t)b}\right)dt\right]. \end{aligned} \tag{2.3}$$

Since each of the integrals is equal to $\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx$, we obtain the left-hand side of the inequality (2.2) from (2.3).

The proof of the second inequality follows by using (2.1) with $x = a$ and $y = b$ and integrating with respect to t over $[0, 1]$.

Now, consider the function $f : (0, \infty) \to \mathbb{R}$, $f(x) = 1$. thus

$$\begin{aligned} 1 &= f\left(\frac{xy}{tx + (1-t)y}\right) \\ &= tf(y) + (1-t)f(x) = 1 \end{aligned}$$

for all $x, y \in (0, \infty)$ and $t \in [0, 1]$. Therefore f is harmonically convex on $(0, \infty)$. We also have

$$f\left(\frac{2ab}{a+b}\right) = 1, \ \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx = 1,$$

and

$$\frac{f(a) + f(b)}{2} = 1$$

which shows us the inequalities (2.2) are sharp. ∎

For finding some new inequalities of Hermite-Hadamard type for functions whose derivatives are harmonically convex, we need a simple lemma below.

2.5. Lemma. *Let $f : I \subset \mathbb{R}\backslash\{0\} \to \mathbb{R}$ be a differentiable function on I° and $a, b \in I$ with $a < b$. If $f' \in L[a, b]$ then*

$$\begin{aligned} &\frac{f(a) + f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \\ &= \frac{ab(b-a)}{2}\int_0^1 \frac{1-2t}{(tb + (1-t)a)^2} f'\left(\frac{ab}{tb + (1-t)a}\right) dt. \end{aligned} \tag{2.4}$$

Proof. Let

$$I^* = \frac{ab(b-a)}{2}\int_0^1 \frac{1-2t}{(tb + (1-t)a)^2} f'\left(\frac{ab}{tb + (1-t)a}\right) dt.$$

By integrating by part, we have

$$I^* = \frac{(2t-1)}{2} f\left(\frac{ab}{tb + (1-t)a}\right)\Bigg|_0^1 - \int_0^1 f\left(\frac{ab}{tb + (1-t)a}\right) dt.$$

Setting $x = \frac{ab}{tb+(1-t)a}$, $dx = \frac{-ab(b-a)}{(tb+(1-t)a)^2}dt = \frac{-x^2(b-a)}{ab}dt$, we obtain

$$I^* = \frac{f(a) + f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx$$

which gives the desired representation (2.4). ∎

2.6. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b \in I$ with $a < b$, and $f' \in L[a,b]$. If $|f'|^q$ is harmonically convex on $[a,b]$ for $q \geq 1$, then*

$$\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \leq \frac{ab\,(b-a)}{2}\lambda_1^{1-\frac{1}{q}}\left[\lambda_2 \left|f'(a)\right|^q + \lambda_3 \left|f'(b)\right|^q\right]^{\frac{1}{q}}, \tag{2.5}$$

where

$$\begin{aligned}
\lambda_1 &= \frac{1}{ab} - \frac{2}{(b-a)^2}\ln\left(\frac{(a+b)^2}{4ab}\right),\\
\lambda_2 &= \frac{-1}{b\,(b-a)} + \frac{3a+b}{(b-a)^3}\ln\left(\frac{(a+b)^2}{4ab}\right),\\
\lambda_3 &= \frac{1}{a\,(b-a)} - \frac{3b+a}{(b-a)^3}\ln\left(\frac{(a+b)^2}{4ab}\right)\\
&= \lambda_1 - \lambda_2.
\end{aligned}$$

Proof. From Lemma 2.5 and using the Hölder inequality, we have

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right|\\
\leq\;& \frac{ab\,(b-a)}{2}\int_0^1 \left|\frac{1-2t}{(tb+(1-t)a)^2}\right|\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|dt\\
\leq\;& \frac{ab\,(b-a)}{2}\left(\int_0^1 \left|\frac{1-2t}{(tb+(1-t)a)^2}\right|dt\right)^{1-\frac{1}{q}}\\
&\times\left(\int_0^1 \left|\frac{1-2t}{(tb+(1-t)a)^2}\right|\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|^q dt\right)^{\frac{1}{q}}.
\end{aligned}$$

Hence, by harmonically convexity of $|f'|^q$ on $[a,b]$, we have

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right|\\
\leq\;& \frac{ab\,(b-a)}{2}\left(\int_0^1 \frac{|1-2t|}{(tb+(1-t)a)^2}dt\right)^{1-\frac{1}{q}}\\
&\times\left(\int_0^1 \frac{|1-2t|\left[t\left|f'(a)\right|^q + (1-t)\left|f'(b)\right|^q\right]}{(tb+(1-t)a)^2}dt\right)^{\frac{1}{q}}\\
\leq\;& \frac{ab\,(b-a)}{2}\lambda_1^{1-\frac{1}{q}}\left[\lambda_2\left|f'(a)\right|^q + \lambda_3\left|f'(b)\right|^q\right]^{\frac{1}{q}}.
\end{aligned}$$

It is easily check that

$$\begin{aligned}&\int_0^1 \frac{|1-2t|}{(tb+(1-t)a)^2}dt \\ &= \frac{1}{ab}-\frac{2}{(b-a)^2}\ln\left(\frac{(a+b)^2}{4ab}\right),\end{aligned}$$

$$\begin{aligned}&\int_0^1 \frac{|1-2t|\,(1-t)}{(tb+(1-t)a)^2}dt \\ &= \frac{1}{a\,(b-a)}-\frac{3b+a}{(b-a)^3}\ln\left(\frac{(a+b)^2}{4ab}\right),\end{aligned}$$

$$\begin{aligned}&\int_0^1 \frac{|1-2t|\,t}{(tb+(1-t)a)^2}dt \\ &= \frac{-1}{b\,(b-a)}+\frac{3a+b}{(b-a)^3}\ln\left(\frac{(a+b)^2}{4ab}\right).\end{aligned}$$

∎

2.7. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a,b \in I$ with $a<b$, and $f' \in L[a,b]$. If $|f'|^q$ is harmonically convex on $[a,b]$ for $q>1$, $\frac{1}{p}+\frac{1}{q}=1$, then*

(2.6)
$$\begin{aligned}&\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \\ \leq\ & \frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\left(\mu_1\left|f'(a)\right|^q+\mu_2\left|f'(b)\right|^q\right)^{\frac{1}{q}},\end{aligned}$$

where

$$\begin{aligned}\mu_1 &= \frac{\left[a^{2-2q}+b^{1-2q}\left[(b-a)(1-2q)-a\right]\right]}{2\,(b-a)^2\,(1-q)\,(1-2q)},\\ \mu_2 &= \frac{\left[b^{2-2q}-a^{1-2q}\left[(b-a)(1-2q)+b\right]\right]}{2\,(b-a)^2\,(1-q)\,(1-2q)}.\end{aligned}$$

Proof. From Lemma 2.5, Hölder's inequality and the harmonically convexity of $|f'|^q$ on $[a,b]$,we have,

$$\begin{aligned}
&\left| \frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \right| \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\int_0^1 |1-2t|^p\,dt\right)^{\frac{1}{p}} \\
&\times\left(\int_0^1 \frac{1}{(tb+(1-t)a)^{2q}}\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|^q dt\right)^{\frac{1}{q}} \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \\
&\times\left(\int_0^1 \frac{t\,|f'\,(a)|^q+(1-t)\,|f'\,(b)|^q}{(tb+(1-t)a)^{2q}}dt\right)^{\frac{1}{q}},
\end{aligned}$$

where an easy calculation gives

$$\begin{aligned}
&\int_0^1 \frac{t}{(tb+(1-t)a)^{2q}}dt \\
=\ & \frac{\left[a^{2-2q}+b^{1-2q}\left[(b-a)(1-2q)-a\right]\right]}{2\,(b-a)^2\,(1-q)\,(1-2q)}
\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}
&\int_0^1 \frac{1-t}{(tb+(1-t)a)^{2q}}dt \\
=\ & \frac{\left[b^{2-2q}-a^{1-2q}\left[(b-a)(1-2q)+b\right]\right]}{2\,(b-a)^2\,(1-q)\,(1-2q)}.
\end{aligned} \tag{2.8}$$

Substituting equations (2.7) and (2.8) into the above inequality results in the inequality (2.6), which completes the proof. ∎

3. Some applications for special means

Let us recall the following special means of two nonnegative number a,b with $b>a$:

(1) The arithmetic mean

$$A = A\,(a,b) := \frac{a+b}{2}.$$

(2) The geometric mean

$$G = G\,(a,b) := \sqrt{ab}.$$

(3) The harmonic mean

$$H = H\,(a,b) := \frac{2ab}{a+b}.$$

(4) The Logarithmic mean

$$L = L(a,b) := \frac{b-a}{\ln b - \ln a}.$$

(5) The p-Logarithmic mean

$$L_p = L_p(a,b) := \left(\frac{b^{p+1} - a^{p+1}}{(p+1)(b-a)}\right)^{\frac{1}{p}}, \quad p \in \mathbb{R} \setminus \{-1, 0\}.$$

(6) the Identric mean

$$I = I(a,b) = \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}}.$$

These means are often used in numerical approximation and in other areas. However, the following simple relationships are known in the literature:

$$H \leq G \leq L \leq I \leq A.$$

It is also known that L_p is monotonically increasing over $p \in \mathbb{R}$, denoting $L_0 = I$ and $L_{-1} = L$.

3.1. Proposition. *Let $0 < a < b$. Then we have the following inequality*

$$H \leq \frac{G^2}{L} \leq A.$$

Proof. The assertion follows from the inequality (2.2) in Theorem 2.4, for $f : (0, \infty) \to \mathbb{R}$, $f(x) = x$. ∎

3.2. Proposition. *Let $0 < a < b$. Then we have the following inequality*

$$H^2 \leq G^2 \leq A(a^2, b^2).$$

Proof. The assertion follows from the inequality (2.2) in Theorem 2.4, for $f : (0, \infty) \to \mathbb{R}$, $f(x) = x^2$. ∎

3.3. Proposition. *Let $0 < a < b$ and $p \in (-1, \infty) \setminus \{0\}$. Then we have the following inequality*

$$H^{p+2} \leq G^2 . L_p^p \leq A(a^{p+2}, b^{p+2}).$$

Proof. The assertion follows from the inequality (2.2) in Theorem 2.4, for $f : (0, \infty) \to \mathbb{R}$, $f(x) = x^{p+2}$, $p\,(-1, \infty) \setminus \{0\}$. ∎

3.4. Proposition. *Let $0 < a < b$. Then we have the following inequality*

$$H^2 \ln H \leq G^2 \ln I \leq A\left(a^2 \ln a, b^2 \ln b\right).$$

Proof. The assertion follows from the inequality (2.2) in Theorem 2.4, for $f : (0, \infty) \to \mathbb{R}$, $f(x) = x^2 \ln x$. ∎

References

[1] Dragomir, S.S. and Pearce, C.E.M., *Selected Topics on Hermite-Hadamard Inequalities and Applications*, RGMIA Monographs, Victoria University, 2000.

[2] İşcan, İ., *A new generalization of some integral inequalities for* (α, m)*-convex functions*, Mathematical Sciences, **7** (22), 1-8, 2013. doi:10.1186/2251-7456-7-22.

[3] Kavurmacı, H., Özdemir, M.E. and Avcı, M., *New Ostrowski type inequalities for* m*-convex functions and applications*, Hacettepe Journal of Mathematics and Statistics, **40** (2), 135 – 145, 2011.

[4] Park, J., *New integral inequalities for products of similar* s*-convex functions in the first sense*, International Journal of Pure and Applied Mathematics, **80 (**4), 585-596, 2012.

[5] Set, E., Ozdemir, M.E. and Dragomir, S.S., *On Hadamard-type inequalities involving several kinds of convexity*, Journal of Inequalities and Applications, **2010**, Article ID 286845, 12 pages, 2010. doi:10.1155/2010/ 286845.

[6] Sulaiman, W.T., *Refinements to Hadamard's inequality for log-convex functions*, Applied Mathematics, **2,** 899-903, 2011.

[7] Zhang, T.-Y., Ji, A.-P. and Qi, F., *On integral inequalities of Hermite Hadamard type for* s*-gometrically convex functions*, Abstract and Applied Analysis, **2012**, Article ID 560586, 14 pages, 2012. doi:10.1155/2012/560586.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 943 – 951

A combinatorial discussion on finite dimensional Leavitt path algebras

A. Koç* , S. Esin† , İ. Güloğlu‡ and M. Kanuni§

Abstract

Any finite dimensional semisimple algebra A over a field K is isomorphic to a direct sum of finite dimensional full matrix rings over suitable division rings. We shall consider the direct sum of finite dimensional full matrix rings over a field K. All such finite dimensional semisimple algebras arise as finite dimensional Leavitt path algebras. For this specific finite dimensional semisimple algebra A over a field K, we define a uniquely determined specific graph - called a truncated tree associated with A - whose Leavitt path algebra is isomorphic to A. We define an algebraic invariant $\kappa(A)$ for A and count the number of isomorphism classes of Leavitt path algebras with the same fixed value of $\kappa(A)$.
Moreover, we find the maximum and the minimum K-dimensions of the Leavitt path algebras of possible trees with a given number of vertices and we also determine the number of distinct Leavitt path algebras of line graphs with a given number of vertices.

2000 AMS Classification: 16S99, 05C05.

Keywords: Finite dimensional semisimple algebra, Leavitt path algebra, Truncated trees, Line graphs.

Received 22 : 02 : 2012 : Accepted 05 : 11 : 2013 Doi : 10.15672/HJMS.2014437524

*Department of Mathematics and Computer Sciences, İstanbul Kültür University, Ataköy, Istanbul, TURKEY
†Tüccarbaşı Sok. Kaşe Apt. No:10A/25 Erenköy, Istanbul, TURKEY
‡Department of Mathematics, Doğuş University, Acıbadem, Istanbul, TURKEY
§Department of Mathematics, Düzce University, Konuralp Düzce, TURKEY

1. Introduction

By the well-known Wedderburn-Artin Theorem [4], any finite dimensional semisimple algebra A over a field K is isomorphic to a direct sum of finite dimensional full matrix rings over suitable division rings. We shall consider the direct sum of finite dimensional full matrix rings over a field K. All such finite dimensional semisimple algebras arise as finite dimensional Leavitt path algebras as studied in [2]. The Leavitt path algebras are introduced independently by Abrams-Aranda Pino in [1] and by Ara-Moreno-Pardo in [3] via different approaches.

In general, the Leavitt path algebra $L_K(E_1)$ can be isomorphic to the Leavitt path algebra $L_K(E_2)$ for non-isomorphic graphs E_1 and E_2. In this paper, we introduce a class of specific graphs which we call the class of truncated trees, denoted by $\mathfrak{T}$, and prove that for any finite acyclic graph E there exists a unique element F in $\mathfrak{T}$ such that $L_K(E)$ is isomorphic to $L_K(F)$. Furthermore, for any two acyclic graphs E_1 and E_2 and their corresponding truncated trees F_1 and F_2 we have

$$L_K(E_1) \cong L_K(E_2) \text{ if and only if } F_1 \cong F_2.$$

For a given finite dimensional Leavitt path algebra $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$ with $2 \leq n_1 \leq n_2 \leq \ldots \leq n_s = N$, the number s is the number of minimal ideals of A and N^2 is the maximum of the dimensions of the minimal ideals. Therefore, the integer $s + N - 1$ is an algebraic invariant of A which we denote by $\kappa(A)$.

Then, we prove that the number of isomorphism classes of finite dimensional Leavitt path algebras A, with the invariant $\kappa(A) > 1$, having no ideals isomorphic to K is equal to the number of distinct truncated trees with $\kappa(A)$ vertices. The number of distinct truncated trees with m vertices is computed in Proposition 3.4.

We also compute the best upper and lower bounds of the K-dimension of possible trees on m vertices, as a function of m and the number of sinks.

In the last section, we calculated the number of isomorphism classes of Leavitt path algebras of line graphs with m vertices as a function of m.

2. Preliminaries

We start by recalling the definitions of a path algebra and a Leavitt path algebra. For a more detailed discussion see [1]. A *directed graph* $E = (E^0, E^1, r, s)$ consists of two countable sets E^0, E^1 and functions $r, s : E^1 \to E^0$. The elements E^0 and E^1 are called *vertices* and *edges*, respectively. For each $e \in E^0$, $s(e)$ is the source of e and $r(e)$ is the range of e. If $s(e) = v$ and $r(e) = w$, then v is said to emit e and w is said to receive e. A vertex which does not receive any edges is called a *source,* and a vertex which emits no edges is called a *sink*. An *isolated* vertex is both a sink and a source. A graph is *row-finite* if $s^{-1}(v)$ is a finite set for each vertex v. A row-finite graph is *finite* if E^0 is a finite set.

A *path* in a graph E is a sequence of edges $\mu = e_1 \ldots e_n$ such that $r(e_i) = s(e_{i+1})$ for $i = 1, \ldots, n-1$. The *source* of μ and the *range* of μ are defined as $s(\mu) = s(e_1)$ and $r(\mu) = r(e_n)$ respectively. The number of edges in a path μ is called the *length* of μ, denoted by $l(\mu)$. If $s(\mu) = r(\mu)$ and $s(e_i) \neq s(e_j)$ for every $i \neq j$, then μ is called a *cycle*. A graph E is called *acyclic* if E does not have any cycles.

The *total-degree* of the vertex v is the number of edges that either have v as its source or as its range, that is, $totdeg(v) = \left|s^{-1}(v) \cup r^{-1}(v)\right|$. A finite graph E is a *line graph* if it is connected, acyclic and $totdeg(v) \leq 2$ for every $v \in E^0$. A line graph E is called an *m-line graph* if E has m vertices.

For $n \geq 2$, define E^n to be the set of paths of length n, and $E^* = \bigcup_{n\geq 0} E^n$ the set of all paths. Given a vertex v in a graph, the number of all paths ending at v is denoted by $n(v)$.

The path K-algebra over E, KE, is defined as the free K-algebra $K[E^0 \cup E^1]$ with the relations:

(1) $v_i v_j = \delta_{ij} v_i$ for every $v_i, v_j \in E^0$,
(2) $e_i = e_i r(e_i) = s(e_i) e_i$ for every $e_i \in E^1$.

Given a graph E, define the extended graph of E as the new graph $\widehat{E} = (E^0, E^1 \cup (E^1)^*, r', s')$ where $(E^1)^* = \{e_i^* \mid e_i \in E^1\}$ is a set with the same cardinality as E and disjoint from E so that the map assigning e* to e is a one-to-one correspondence; and the functions r' and s' are defined as

$$r'|_{E^1} = r, \quad s'|_{E^1} = s, \quad r'(e_i^*) = s(e_i) \quad \text{and} \quad s'(e_i^*) = r(e_i).$$

The Leavitt path algebra of E, $L_K(E)$, with coefficients in K is defined as the path algebra over the extended graph $\widehat{E}$, which satisfies the additional relations:

(CK1) $e_i^* e_j = \delta_{ij} r(e_j)$ for every $e_j \in E^1$ and $e_i^* \in (E^1)^*$,
(CK2) $v_i = \sum_{\{e_j \in E^1 \mid s(e_j) = v_i\}} e_j e_j^*$ for every $v_i \in E^0$ which is not a sink, and emits only finitely many edges.

The conditions (CK1) and (CK2) are called the Cuntz-Krieger relations. Note that the condition of row-finiteness is needed in order to define the equation (CK2).

Finite dimensional Leavitt path algebras are studied in [2] by Abrams, Aranda Pino and Siles Molina. The authors characterize the structure theorems for finite dimensional Leavitt path algebras. Their results are summarized in the following proposition:

2.1. Proposition. (1) *The Leavitt path algebra $L_K(E)$ is a finite-dimensional K-algebra if and only if E is a finite and acyclic graph.*

(2) *If $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$, then $A \cong L_K(E)$ for a graph E having s connected components each of which is an oriented line graph with n_i vertices, $i = 1, 2, \cdots, s$.*

(3) *A finite dimensional K-algebra A arises as a $L_K(E)$ for a graph E if and only if $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$.*

(4) *If $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$ and $A \cong L_K(E)$ for a finite, acyclic graph E, then the number of sinks of E is equal to s, and each sink v_i $(i = 1, 2, \cdots, s)$ has $n(v_i) = n_i$ with a suitable indexing of the sinks.*

3. Truncated Trees

For a finite dimensional Leavitt path algebra $L_K(E)$ of a graph E, we construct a distinguished graph F having the Leavitt path algebra isomorphic to $L_K(E)$ as follows:

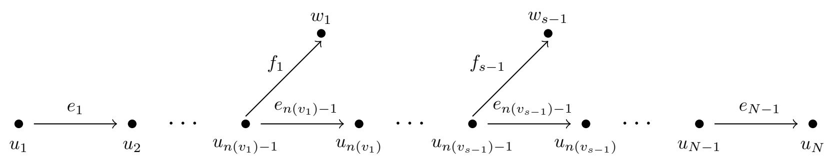
3.1. Theorem. *Let E be a finite, acyclic graph with no isolated vertices. Let $s = |S(E)|$ where $S(E)$ is the set of sinks of E and $N = \max\{n(v) \mid v \in S(E)\}$. Then there exists a unique (up to isomorphism) tree F with exactly one source and $s + N - 1$ vertices such that $L_K(E) \cong L_K(F)$.*

Proof. Let the sinks $v_1, v_2, \ldots, v_s$ of E be indexed such that

$$2 \leq n(v_1) \leq n(v_2) \leq \ldots \leq n(v_s) = N.$$

Define a graph $F = (F^0, F^1, r, s)$ as follows:

$$\begin{array}{rcll} F^0 & = & \{u_1, u_2, \ldots, u_N, w_1, w_2, \ldots w_{s-1}\} & \\ F^1 & = & \{e_1, e_2, \ldots, e_{N-1}, f_1, f_2, \ldots, f_{s-1}\} & \\ s(e_i) & = & u_i \quad \text{and} \quad r(e_i) = u_{i+1} & i = 1, \ldots, N-1 \\ s(f_i) & = & u_{n(v_i)-1} \quad \text{and} \quad r(f_i) = w_i & i = 1, \ldots, s-1. \end{array}$$



Clearly, F is a directed tree with unique source u_1 and $s + N - 1$ vertices. The graph F has exactly s sinks, namely $u_N, w_1, w_2, \ldots w_{s-1}$ with $n(u_N) = N$, $n(w_i) = n(v_i)$, $i = 1, \ldots, s-1$. Therefore, $L_K(E) \cong L_K(F)$ by Proposition 2.1.

For the uniqueness part, take a tree T with exactly one source and $s+N-1$ vertices such that $L_K(E) \cong L_K(T)$. Now $N = \max\{n(v) \mid v \in S(E)\}$ is equal to the square root of the maximum of the K-dimensions of the minimal ideals of $L_K(E)$ and also of $L_K(T)$. So there exists a sink v in T with $|\{\mu_i \in T^* \mid r(\mu_i) = v\}| = N$. Since, any vertex in T is connected to the unique source by a uniquely determined path, the unique path joining v to the source must contain exactly N vertices, say $a_1, ..., a_{N-1}, v$ where a_1 is the unique source and the length of the path joining a_k to a_1 being equal to $k-1$ for any $k = 1, 2, ..., N-1$. As $L_K(E) = \bigoplus_{i=1}^{s} M_{n_i}(K)$ with s summands, all the remaining $s-1$ vertices, say $b_1, ..., b_{s-1}$, must be sinks by Proposition 2.1(4). For any vertex a different from the unique source, clearly $n(a) > 1$. Also, there exists an edge g_i with $r(g_i) = b_i$ for each $i = 1, \ldots, s-1$. Since $s(g_i)$ is not a sink, it follows that $s(g_i) \in \{a_1, a_2, ..., a_{N-1}\}$, more precisely $s(g_i) = a_{n(b_i)-1}$ for $i = 1, 2, ..., s-1$. Thus T is isomorphic to F. □

We name the graph F constructed in Theorem 3.1 as the *truncated tree associated with E*.

3.2. Proposition. *With the above definition of F, there is no tree T with $|T^0| < |F^0|$ such that $L_K(T) \cong L_K(F)$.*

Proof. Notice that since T is a tree, any vertex contributing to a sink represents a unique path ending at that sink.

Assume on the contrary there exists a tree T with n vertices and $L_K(T) \cong A = \bigoplus_{i=1}^{s} M_{n_i}(K)$ such that $n < s + N - 1$. Since N is the maximum of n_i's there exists a sink v with $n(v) = N$. But in T the number $n - s$ of vertices which are not sinks is less than $N - 1$. Hence the maximum contribution to any sink can be at most $n - s + 1$ which is strictly less than N. This is the desired contradiction. □

Remark that the above proposition does not state that it is impossible to find a graph G with smaller number of vertices having $L_K(G)$ isomorphic to $L_K(E)$. The next example illustrates this point.

3.3. Example. Consider the graphs G and F.

Both $L_K(G) \cong M_3(K) \cong L_K(F)$ and $|G^0| = 2$ where as $|F^0| = 3$.

G F

Given any graphs G_1 and G_2, $L_K(G_1) \cong L_K(G_2)$ does not necessarily imply $G_1 \cong G_2$. However, for truncated trees F_1, F_2 we have $F_1 \cong F_2$ if and only if $L_K(F_1) \cong L_K(F_2)$. So there is a one-to-one correspondence between the Leavitt path algebras and the truncated trees.

Consider a finite dimensional Leavitt path algebra $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$ with $2 \leq n_1 \leq n_2 \leq \ldots \leq n_s = N$. Here, the number s is the number of minimal ideals of A and N^2 is the maximum of the dimensions of the minimal ideals. Therefore, the integer $s + N - 1$ is an algebraic invariant of A which is denoted by $\kappa(A)$. Notice that the number of isomorphism classes of finite dimensional Leavitt path algebras A, with the invariant $\kappa(A) > 1$, having no ideals isomorphic to K is equal to the number of distinct truncated trees with $\kappa(A)$ vertices by the previous paragraph. The next proposition computes this number.

3.4. Proposition. *The number of distinct truncated trees with m vertices is 2^{m-2}.*

Proof. In a truncated tree, $n(v_1) \neq n(v_2)$ for any two distinct non-sinks v_1 and v_2. For every sink v, there is a unique non-sink w so that there exists an edge e with $s(e) = w$ and $r(e) = v$. Namely the non-sink w is with $n(w) = n(v) - 1$. This w is denoted by $b(v)$.

Now, define $d(u) = |\{v : n(v) \leq n(u)\}|$ for any $u \in E^0$. Clearly, $d(u)$ is equal to the sum of $n(u)$ and the number of sinks v with $n(b(v)) < n(u)$ for any $u \in E^0$. Assign an m-tuple $\alpha(E) = (\alpha_1, \alpha_2, ..., \alpha_m) \in \{0,1\}^m$ to a truncated tree E with m vertices by letting $\alpha_j = 1$ if and only if $j = d(v)$ for some vertex v which is not a sink. Clearly, there is just one vertex v with $n(v) = 1$, namely the unique source of E and that vertex is not a sink, so $\alpha_1 = 1$. Since there cannot be any non-sink v with $d(v) = m$, it follows that $\alpha_m = 0$.

Conversely, for $\beta = (\beta_1, \beta_2, ..., \beta_m) \in \{0,1\}^m$ with $\beta_1 = 1$ and $\beta_m = 0$ there exists a unique truncated tree E with m vertices such that $\alpha(E) = \beta$: If $\beta_i = 1$, then assign a non-sink v to E with $n(v) = |\{k : 1 \leq k < i \text{ and } \beta_k = 1\}|$. If $\beta_i = 0$ and $j = |\{k : 1 \leq k < i \text{ and } \beta_k = 1\}|$ then construct a sink which is joined to the non-sink v with $n(v) = j$. Clearly, the graph E is a truncated tree with m vertices and $\alpha(E) = \beta$.

Hence the number of distinct truncated trees with m vertices is equal to 2^{m-2} which is the number of all elements of $\{0,1\}^m$ with the first component 1 and the last component 0. □

Hence, we have the following corollary.

3.5. Corollary. *Given $n \geq 2$, the number of isomorphism classes of finite dimensional Leavitt path algebras A with $\kappa(A) = n$ and which do not have any ideals isomorphic to K is 2^{n-2}.*

4. Bounds on the K-Dimension of finite dimensional Leavitt Path Algebras

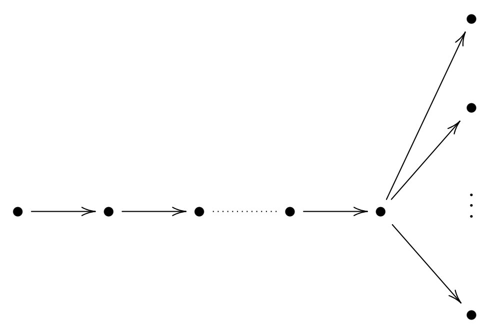
For a tree F with m vertices, the K-dimension of $L_K(F)$ is not uniquely determined by the number of vertices only. However, we can compute the maximum and the minimum K-dimensions of $L_K(F)$ where F ranges over all possible trees with m vertices.

4.1. Lemma. *The maximum K-dimension of $L_K(E)$ where E ranges over all possible trees with m vertices and s sinks is attained at a tree in which $n(v) = m - s + 1$ for each sink v. In this case, the value of the dimension is $s(m - s + 1)^2$.*

Proof. Assume E is a tree with m vertices. Then $L_K(E) \cong \bigoplus_{i=1}^{s} M_{n_i}(K)$, by Proposition 2.1 (3) where s is the number of sinks in E and $n_i \leq m - s + 1$ for all $i = 1, \ldots s$. Hence

$$\dim L_K(E) = \sum_{i=1}^{s} n_i^2 \leq s(m - s + 1)^2.$$

Notice that there exists a tree E as sketched below



with m vertices and s sinks such that $\dim L_K(E) = s(m - s + 1)^2$. □

4.2. Theorem. *The maximum K-dimension of $L_K(E)$ where E ranges over all possible trees with m vertices is given by $f(m)$ where*

$$f(m) = \begin{cases} \dfrac{m(2m+3)^2}{27} & if \quad m \equiv 0 (mod\ 3) \\[2ex] \dfrac{1}{27}\,(m+2)\,(2m+1)^2 & if \quad m \equiv 1 (mod\ 3) \\[2ex] \dfrac{4}{27}(m+1)^3 & if \quad m \equiv 2 (mod\ 3) \end{cases}$$

Proof. Assume E is a tree with m vertices. Then $L_K(E) \cong \bigoplus_{i=1}^{s} M_{n_i}$ where s is the number of sinks in E. Now, to find the maximum dimension of $L_K(E)$, determine the maximum value of the function $f(s) = s(m - s + 1)^2$ for $s = 1, 2, \ldots, m - 1$. Extending the domain of $f(s)$ to real numbers $1 \leq s \leq m - 1$ f becomes a continuous function, hence its maximum value can be computed.

$$f(s) = s(m - s + 1)^2 \Rightarrow \frac{d}{ds}\left(s(m - s + 1)^2\right) = (m - 3s + 1)\,(m - s + 1)$$

Then $s = \dfrac{m+1}{3}$ is the only critical point in the interval $[1, m-1]$ and since $\dfrac{d^2 f}{ds^2}(\dfrac{m+1}{3}) < 0$, it is a local maximum. In particular f is increasing on the interval $\left[1, \dfrac{m+1}{3}\right]$ and decreasing on $\left[\dfrac{m+1}{3}, m-1\right]$. There are three cases:

Case 1: $m \equiv 2 \pmod 3$. In this case $s = \dfrac{m+1}{3}$ is an integer and maximum K-dimension of $L_K(E)$ is $f\left(\dfrac{m+1}{3}\right) = \dfrac{4}{27}(m+1)^3$ and $n_i = \dfrac{2(m+1)}{3}$, for each $i = 1, 2, \ldots, s$.

Case 2: $m \equiv 0 \pmod 3$. Then: $\dfrac{m}{3} = t < t + \dfrac{1}{3} = s < t+1$ and

$$f\left(\frac{m}{3}\right) = \frac{(2m+3)^2 m}{27} = \alpha_1 \text{ and } f\left(\frac{m}{3}+1\right) = \frac{4m^2(m+3)}{27} = \alpha_2.$$

Note that, $\alpha_1 > \alpha_2$. So α_1 is maximum K -dimension of $L_K(E)$ and $n_i = \dfrac{2}{3}m + 1$, for each $i = 1, 2, \ldots, s$.

Case 3: $m \equiv 1 \pmod 3$. Then $\dfrac{m-1}{3} = t < t + \dfrac{2}{3} = s < t+1$ and

$$f\left(\frac{m-1}{3}\right) = \frac{4}{27}(m+2)^2(m-1) = \beta_1$$

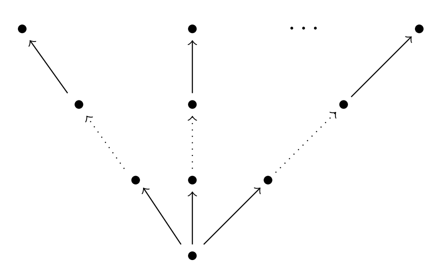
and

$$f\left(\frac{m+2}{3}\right) = \frac{1}{27}(2m+1)^2(m+2) = \beta_2.$$

In this case $\beta_2 > \beta_1$ and so β_2 gives the maximum K-dimension of $L_K(E)$ and $n_i = \dfrac{2m+1}{3}$, for each $i = 1, 2, \ldots, s$. □

4.3. Theorem. *The minimum K-dimension of $L_K(E)$ where E ranges over all possible trees with m vertices and s sinks is equal to $r(q+2)^2 + (s-r)(q+1)^2$, where $m-1 = qs + r$, $0 \le r < s$.*

Proof. We call a graph a *bunch tree* if it is obtained by identifying the unique sources of the finitely many disjoint oriented finite line graphs as seen in the figure.



Let $\mathcal{E}(m, s)$ be the set of all bunch trees with m vertices and s sinks. Every element of $\mathcal{E}(m, s)$ can be uniquely represented by an s -tuple $(t_1, t_2, ..., t_s)$ where each t_i is the number of vertices different from the source contributing to the i^{th} sink, with $1 \leq t_1 \leq t_2 \leq ... \leq t_s$ and $t_1 + t_2 + ... + t_s = m - 1$.

Let $E \in \mathcal{E}(m,s)$ with $t_s - t_1 \leq 1$. This E is represented by the s-tuple $(q, \ldots, q, q+1, \ldots, q+1)$ where $m - 1 = sq + r$, $0 \leq r < s$.

Now, claim that the dimension of E is the minimum of the set

$$\{\dim L_K(F) : F \text{ tree with } s \text{ sinks and } m \text{ vertices}\}.$$

If we represent $U \in \mathcal{E}(m,s)$ by the s-tuple $(u_1, u_2, ..., u_s)$ then $E \neq U$ implies that $u_s - u_1 \geq 2$.

Consider the s-tuple $(t_1, t_2, ..., t_s)$ where $(t_1, t_2, ..., t_s)$ is obtained from $(u_1 + 1, u_2, ..., u_{s-1}, u_s - 1)$ by reordering the components in increasing order.

In this case, the dimension d_U of U is

$$d_U = (u_1 + 1)^2 + \ldots + (u_s + 1)^2.$$

Similarly, the dimension d_T of the bunch graph T represented by the s-tuple $(t_1, t_2, ..., t_s)$, is

$$d_T = (t_1 + 1)^2 + \ldots + (t_s + 1)^2 = (u_1 + 2)^2 + \ldots + (u_{s-1} + 1)^2 + u_s^2.$$

Hence

$$d_U - d_T = 2(u_s - u_1) - 2 > 0.$$

Repeating this process sufficiently many times, the process has to end at the exceptional bunch tree E showing that its dimension is the smallest among the dimensions of all elements of $\mathcal{E}(m,s)$.

Now let F be an arbitrary tree with m vertices and s sinks. As above assign to F the s-tuple $(n_1, n_2, ..., n_s)$ with $n_i = n(v_i) - 1$ where the sinks $v_i,\ i = 1, 2, \ldots, s$ are indexed in such a way that $n_i \leq n_{i+1},\ i = 1, \ldots, s-1$. Observe that $n_1 + n_2 + \cdots + n_s \geq m - 1$. Let $\beta = \sum_{i=1}^{s} n_i - (m-1)$. Since $s \leq m - 1$, $\beta \leq \sum_{i=1}^{s}(n_i - 1)$. Either $n_1 - 1 \geq \beta$ or there exists a unique $k \in \{2, \ldots, s\}$ such that $\sum_{i=1}^{k-1}(n_i - 1) < \beta \leq \sum_{i=1}^{k}(n_i - 1)$. If $n_1 - 1 \geq \beta$, then let

$$m_i = \begin{cases} n_1 - \beta & , \quad i = 1 \\ n_i & , \quad i > 1 \end{cases}.$$

Otherwise, let

$$m_i = \begin{cases} 1 & , \quad i \leq k-1 \\ n_k - \left(\beta - \sum_{i=1}^{k-1}(n_i - 1)\right) & , \quad i = k \\ n_i & , \quad i \geq k+1 \end{cases}.$$

In both cases, the s-tuple $(m_1, m_2, \ldots, m_s)$ that satisfies $1 \leq m_i \leq n_i$, $m_1 \leq m_2 \leq \cdots \leq m_s$ and $m_1 + m_2 + \cdots + m_s = m - 1$ is obtained. So, there exists a bunch tree M namely the one corresponding uniquely to $(m_1, m_2, \ldots, m_s)$ which has dimension $d_M \leq d_F$. This implies that $d_F \geq d_E$.

Hence the result follows. □

4.4. Lemma. *The minimum K-dimension of $L_K(E)$ where E ranges over all possible trees with m vertices occurs when the number of sinks is $m-1$ and is equal to $4(m-1)$.*

Proof. By the previous theorem observe that

$$\dim L_K(E) \geq r(q+2)^2 + (s-r)(q+1)^2$$

where $m - 1 = qs + r,\quad 0 \leq r < s$. Then

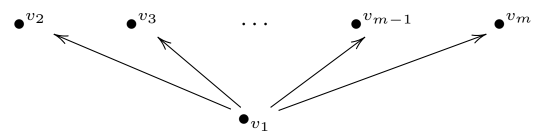
$$r(q+2)^2 + (s-r)(q+1)^2 = (m-1)(q+2) + qr + r + s.$$

Thus

$$(m-1)(q+2)+qr+r+s-4(m-1) = (m-1)(q-2)+qr+r+s \geq 0 \quad if \quad q \geq 2.$$

If $q = 1$, then $-(m-1)+2r+s = -(m-1)+r+(m-1) = r \geq 0$. Hence $\dim L_K(E) \geq 4(m-1)$.

Notice that there exists a truncated tree E with m vertices and $\dim L_K(E) = 4(m-1)$ as sketched below :



□

5. Line Graphs

In [2], the Proposition 5.7 shows that a semisimple finite dimensional algebra $A = \bigoplus_{i=1}^{s} M_{n_i}(K)$ over the field K can be described as a Leavitt path algebra $L_K(E)$ defined by a line graph E, if and only if A has no ideals of K-dimension 1 and the number of minimal ideals of A of K-dimension 2^2 is at most 2. On the other hand, if $A \cong L_K(E)$ for some m-line graph E then $m-1 = \sum_{i=1}^{s}(n_i - 1)$, that is, m is an algebraic invariant of A.

Therefore the following proposition answers a reasonable question.

5.1. Proposition. *The number A_m of isomorphism classes of Leavitt path algebras defined by line graphs having exactly m vertices is*

$$A_m = P(m-1) - P(m-4)$$

where $P(t)$ is the number of partitions of the natural number t.

Proof. Any m-line graph has $m-1$ edges. In a line graph, for any edge e there exists a unique sink v so that there exists a path from $s(e)$ to v. In this case we say that e is directed towards v. The number of edges directed towards v is clearly equal to $n(v)-1$. Let E and F be two m -line graphs. Then $L_K(E) \cong L_K(F)$ if and only if there exists a bijection $\phi : S(E) \to S(F)$ such that for each v in $S(E)$, $n(v) = n(\phi(v))$. Therefore the number of isomorphism classes of Leavitt path algebras determined by m-line graphs is the number of partitions of $m-1$ edges in which the number of parts having exactly one edge is at most two. Since the number of partitions of k objects having at least three parts each of which containing exactly one element is $P(k-3)$, the result $A_m = P(m-1) - P(m-4)$ follows. □

References

[1] G. Abrams, G. Aranda Pino, The Leavitt path algebra of a graph, J. Algebra 293 (2) (2005), 319 - 334.

[2] G. Abrams, G. Aranda Pino, M. Siles Molina, Finite-dimensional Leavitt path algebras, J. Pure Appl. Algebra 209 (2007) 753 - 762.

[3] P. Ara, M.A. Moreno, and E. Pardo, Nonstable K-theory for graph algebras, Alg. Rep. Thy. 10 (2007), 157–178.

[4] T.Y. Lam, A First Course In Noncommutative Rings, Springer-Verlag 2001.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 953 – 961

On partially τ-quasinormal subgroups of finite groups

Changwen Li* , Xuemei Zhang† and Xiaolan Yi‡

Abstract

Let H be a subgroup of a group G. We say that: (1) H is τ-quasinormal in G if H permutes with every Sylow subgroup Q of G such that $(|H|, |Q|) = 1$ and $(|H|, |Q^G|) \neq 1$; (2) H is partially τ-quasinormal in G if G has a normal subgroup T such that HT is S-quasinormal in G and $H \cap T \leq H_{\tau G}$, where $H_{\tau G}$ is the subgroup generated by all those subgroups of H which are τ-quasinormal in G. In this paper, we find a condition under which every chief factor of G below a normal subgroup E of G is cyclic by using the partial τ-quasinormality of some subgroups.

2000 AMS Classification: 20D10, 20D20.

Keywords: S-quasinormal, partially τ-quasinormal, cyclic.

Received 05 : 06 : 2013 : Accepted 03 : 10 : 2013 Doi : 10.15672/HJMS.2104437528

1. Introduction

All groups considered in the paper are finite. The notations and terminology in this paper are standard, as in [4] and [6]. G always denotes a finite group, $\pi(G)$ denotes the set of all prime dividing $|G|$ and $F^*(G)$ is the generalized Fitting subgroup of G, i.e., the product of all normal quasinilpotent subgroups of G.

Normal subgroup plays an important role in the study of the structure of groups. Many authors are interested to extend the concept of normal subgroup. For example, a subgroup H of G is said to be S-quasinormal [7] in G if H permutes with every Sylow subgroup of G. As a generalization of S-quasinormality, a subgroup H of G is said to be τ-quasinormal [11] in G if H permutes with every Sylow subgroup Q of G such that $(|H|, |Q|) = 1$ and $(|H|, |Q^G|) \neq 1$. On the other hand, Wang [17] extended normality as

*School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, 221116, China Email: lcw2000@126.com

†Department of Basic Sciences, Yancheng Institute of Technology, Yancheng, 224051, China

‡School of Science, Zhejiang Sci-Tech University, Hangzhou, 310018, China

The project is supported by the Natural Science Foundation of China (No:11401264) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

follows: a subgroup H of G is said to be c-normal in G if there exists a normal subgroup K of G such that $HK = G$ and $H \cap K \leq H_G$, where H_G is the maximal normal subgroup of G contained in H. In the literature, many people have studied the influence of the τ-quasinormality and c-normality on the structure of finite groups and obtained many interesting results (see [2, 5, 8, 11, 12, 17, 19]). As a development, we now introduce a new concept:

1.1. Definition. A subgroup H of a group G is said to be partially τ-quasinormal in G if there exists a normal subgroup T of G such that HT is S-quasinormal in G and $H \cap T \leq H_{\tau G}$, where $H_{\tau G}$ is the subgroup generated by all those subgroups of H which are τ-quasinormal in G.

Clearly, partially τ-quasinormal subgroup covers both the concepts of τ-quasinormal subgroup and c-normal subgroup. However, the following examples show that the converse is not true.

1.2. Example. Let $G = S_4$ be the symmetric group of degree 4.

(1) Let H be a Sylow 3-subgroup of G and N the normal abelian 2-subgroup of G of order 4. Then $HN = A_4 \trianglelefteq G$ and $H \cap N$=1. Hence H is a partially τ-quasinormal subgroup of G. But, obviously, H is not c-normal in G.

(2) Let H=$\langle(14)\rangle$. Obviously, HA_4=G and $H \cap A_4$=1. Hence H is partially τ-quasinormal in G. But, obviously, H is not τ-quasinormal in G.

A normal subgroup E of a group G is said to be hypercyclically embedded in G if every chief factor of G below E is cyclic. The product of all normal hypercyclically embedded subgroups of G is denoted by $Z_{\mathscr{U}}(G)$. In [15] and [16], Skiba gave some characterizations of normal hypercyclically embedded subgroups related to S-quasinormal subgroups. The main purpose of this paper is to give a new characterization by using partially τ-quasinormal property of maximal subgroups of some Sylow subgroups. We obtain the following result.

Main Theorem. *Let E be a normal subgroup of G. Suppose that there exists a normal subgroup X of G such that $F^*(E) \leq X \leq E$ and X satisfies the following properties: for every non-cyclic Sylow p-subgroup P of X, every maximal subgroup of P not having a supersoluble supplement in G is partially τ-quasinormal in G. Then E is hypercyclically embedded in G.*

The following theorems are the main stages in the proof of Main Theorem.

1.3. Theorem. Let P be a Sylow p-subgroup of a group G, where p is a prime divisor of $|G|$ with $(|G|, p-1) = 1$. If every maximal subgroup of P not having a p-nilpotent supplement in G is partially τ-quasinormal in G, then G is soluble.

1.4. Theorem. Let P be a Sylow p-subgroup of a group G, where p is a prime divisor of $|G|$ with $(|G|, p-1) = 1$. Then G is p-nilpotent if and only if every maximal subgroup of P not having a p-nilpotent supplement in G is partially τ-quasinormal in G.

1.5. Theorem. Let E be a normal subgroup in G and let P be a Sylow p-subgroup of E, where p is a prime divisor of $|E|$ with $(|E|, p-1) = 1$. Suppose that every maximal subgroup of P not having a p-supersoluble supplement in G is partially τ-quasinormal in G. Then each chief factor of G between E and $O_{p'}(E)$ is cyclic.

1.6. Theorem. Let E be a normal subgroup of a group G. Suppose that for each $p \in \pi(E)$, every maximal subgroup of non-cyclic Sylow p-subgroup P of E not having a p-supersoluble supplement in G is partially τ-quasinormal in G. Then every chief factor of G below E is cyclic.

2. Preliminaries

2.1. Lemma ([3] and [7])**.** Suppose that H is a subgroup of G and H is S-quasinormal in G. Then

(1) If $H \leq K \leq G$, then H is S-quasinormal in K.

(2) If N is a normal subgroup of G, then HN is S-quasinormal in G and HN/N is S-quasinormal in G/N.

(3) If $K \leq G$, then $H \cap K$ is S-quasinormal in K.

(4) H is subnormal in G.

(5) If $K \leq G$ and K is S-quasinormal in G, then $H \cap K$ is S-quasinormal in G.

2.2. Lemma ([11, Lemmas 2.2 and 2.3])**.** Let G be a group and $H \leq K \leq G$.

(1) If H is τ-quasinormal in G, then H is τ-quasinormal in K.

(2) Suppose that H is normal in G and $\pi(K/H) = \pi(K)$. If K is τ-quasinormal in G, then K/H is τ-quasinormal in G/H.

(3) Suppose that H is normal in G. Then EH/H is τ-quasinormal in G/H for every τ-quasinormal subgroup E in G satisfying $(|H|, |E|) = 1$.

(4) If H is τ-quasinormal in G and $H \leq O_p(G)$ for some prime p, then H is S-quasinormal in G.

(5) $H_{\tau G} \leq H_{\tau K}$.

(6) Suppose that K is a p-group and H is normal in G. Then $K_{\tau G}/H \leq (K/H)_{\tau(G/H)}$.

(7) Suppose that H is normal in G. Then $E_{\tau G}H/H \leq (EH/H)_{\tau(G/H)}$ for every p-subgroup E of G satisfying $(|H|, |E|) = 1$.

2.3. Lemma. Let G be a group and $H \leq G$. Then

(1) If H is partially τ-quasinormal in G and $H \leq K \leq G$, then H is partially τ-quasinormal in K.

(2) Suppose that $N \trianglelefteq G$ and $N \leq H$. If H is a p-group and H is partially τ-quasinormal in G, then H/N is partially τ-quasinormal in G/N.

(3) Suppose that H is a p-subgroup of G and N is a normal p'-subgroup of G. If H is partially τ-quasinormal in G, then HN/N is partially τ-quasinormal in G/N.

(4) If H is partially τ-quasinormal in G and $H \leq K \trianglelefteq G$, then there exists $T \trianglelefteq G$ such that HT is S-quasinormal in G, $H \cap T \leq H_{\tau G}$ and $HT \leq K$.

Proof. (1) Let N be a normal subgroup of G such that HN is S-quasinormal in G and $H \cap N \leq H_{\tau G}$. Then $K \cap N \trianglelefteq K$, $H(K \cap N) = HN \cap K$ is S-quasinormal in K by Lemma 2.1(3) and $H \cap (K \cap N) = H \cap N \leq H_{\tau G} \leq H_{\tau K}$ by Lemma 2.2(5). Hence H is partially τ-quasinormal in K.

(2) Suppose that H is partially τ-quasinormal in G. Then there exists $K \trianglelefteq G$ such that HK is S-quasinormal in G and $H \cap K \leq H_{\tau G}$. This implies that $KN/N \trianglelefteq G/N$ and $(H/N)(KN/N) = HK/N$ is S-quasinormal in G/N by Lemma 2.1(2). In view of Lemma 2.2(6), $H/N \cap KN/N = (H \cap K)N/N \leq H_{\tau G}N/N = H_{\tau G}/N \leq (H/N)_{\tau(G/N)}$. Thus H/N is partially τ-quasinormal in G/N.

(3) Suppose that H is partially τ-quasinormal in G. Then there exists $K \trianglelefteq G$ such that HK is S-quasinormal in G and $H \cap K \leq H_{\tau G}$. Clearly, $KN/N \trianglelefteq G$ and $(HN/N)(KN/N) = HKN/N$ is S-quasinormal in G/N by Lemma 2.1(2). On the other hand, since $(|HN : H|, |HN : N|)$=1, $HN/N \cap KN/N = (HN \cap K)N/N = (H \cap K)(N \cap K)N/N = (H \cap K)N/N \leq H_{\tau G}N/N$. In view of Lemma 2.2(7), we have $H_{\tau G}N/N \leq (HN/N)_{\tau(G/N)}$. Hence HN/N is partially τ-quasinormal in G/N.

(4) Suppose that H is partially τ-quasinormal in G. Then there exists $N \trianglelefteq G$ such that HN is S-quasinormal in G and $H \cap N \leq H_{\tau G}$. Let $T = N \cap K$. Then $T \trianglelefteq G$, $HT = H(N \cap K) = HN \cap K$ is S-quasinormal in G by Lemma 2.1(5), $HT \leq K$ and $H \cap T = H \cap N \cap K = H \cap N \leq H_{\tau G}$. □

2.4. Lemma. Let G be a group and p a prime dividing $|G|$ with $(|G|, p-1) = 1$.
(1) If N is normal in G of order p, then N lies in $Z(G)$.
(2) If G has cyclic Sylow p-subgroups, then G is p-nilpotent.
(3) If $M \leq G$ and $|G : M| = p$, then $M \trianglelefteq G$.
(4) If G is p-supersoluble, then G is p-nilpotent.

Proof. (1), (2) and (3) can be found in [18, Theorem 2.8]. Now we only prove (4). Let A/B be an arbitrary chief factor of G. If G is p-supersolvable, then A/B is either a cyclic group with order p or a p'-group. If $|A/B| = p$, then $|\mathrm{Aut}(A/B)| = p - 1$. Since $G/C_G(A/B)$ is isomorphic to a subgroup of $\mathrm{Aut}((A/B)$, the order of $G/C_G(A/B)$ must divide $(|G|, p-1) = 1$, which shows that $G = C_G(A/B)$. Therefore, we have G is p-nilpotent. □

2.5. Lemma ([10, Lemma 2.12])**.** Let P be a Sylow p-subgroup of a group G, where p is a prime divisor of $|G|$ with $(|G|, p-1) = 1$. If every maximal subgroup of P has a p-nilpotent supplement in G, then G is p-nilpotent.

2.6. Lemma ([13, Theorem A])**.** If P is an S-quasinormal p-subgroup of a group G for some prime p, then $N_G(P) \geq O^p(G)$.

2.7. Lemma ([6, VI, 4.10])**.** Assume that A and B are two subgroups of a group G and $G \neq AB$. If $AB^g = B^g A$ holds for any $g \in G$, then either A or B is contained in a nontrivial normal subgroup of G.

2.8. Lemma ([20, Chap.1, Theorem 7.19])**.** Let H be a normal subgroup of G. Then $H \leq Z_{\mathscr{U}}(G)$ if and only if $H/\Phi(H) \leq Z_{\mathscr{U}}(G/\Phi(H))$.

2.9. Lemma ([14, Lemma 2.11])**.** Let N be an elementary abelian normal subgroup of a group G. Assume that N has a subgroup D such that $1 < |D| < |N|$ and every subgroup H of N satisfying $|H| = |D|$ is S-quasinormal in G. Then some maximal subgroup of N is normal in G.

2.10. Lemma. Let N be a non-identity normal p-subgroup of a group G. If N is elementary and every maximal subgroup of N is partially τ-quasinormal in G, then some maximal subgroup of N is normal in G.

Proof. If $|N| = p$, then it is clear. Let L be a non-identity minimal normal p-subgroup of G contained in N. First we assume that $N \neq L$. By Lemma 2.3(2), the hypothesis still holds on G/L. Then by induction some maximal subgroup M/L of N/L is normal in G/L. Clearly, M is a maximal subgroup of N and M is normal in G. Consequently the lemma follows. Now suppose that $L = N$. Let M be any maximal subgroup of N. Then by the hypothesis, there exists $T \trianglelefteq G$ such that MT is S-quasinormal in G and $M \cap T \leq M_{\tau G}$. Suppose that $M \neq M_{\tau G}$. Then $MT \neq M$ and $T \neq 1$. If $N \leq MT$, then $N = N \cap MT = M(N \cap T)$. Hence $N \leq T$, which implies that $M = M \cap T = M_{\tau G}$, a contradiction. If $N \nleq MT$, then $M = M(T \cap N) = MT \cap N$ is S-quasinormal in G by Lemma 2.1(5), a contradiction again. Hence $M = M_{\tau G}$. In view of Lemma 2.2(4), M is S-quasinormal in G. By Lemma 2.9, some maximal subgroup of N is normal in G. Thus the lemma holds. □

2.11. Lemma ([15, Theorem B])**.** Let $\mathscr{F}$ be any formation and G a group. If $H \triangleleft G$ and $F^*(H) \leq Z_{\mathscr{F}}(G)$, then $H \leq Z_{\mathscr{F}}(G)$.

3. Proofs of Theorems

Proof of Theorem 1.3. Assume that this theorem is false and let G be a counterexample with minimal order. We proceed the proof via the following steps.

(1) $O_p(G) = 1$.

Assume that $L = O_p(G) \neq 1$. Clearly, P/L is a Sylow p-subgroup of G/L. Let M/L be a maximal subgroup of P/L. Then M is a maximal subgroup of P. If M has a p-nilpotent supplement D in G, then M/L has a p-nilpotent supplement DL/L in G/L. If M is partially τ-quasinormal in G, then M/L is partially τ-quasinormal in G/L by Lemma 2.3(2). Hence G/L satisfies the hypothesis of the theorem. The minimal choice of G implies that G/L is soluble. Consequently, G is soluble. This contradiction shows that step (1) holds.

(2) $O_{p'}(G) = 1$.

Assume that $R = O_{p'}(G) \neq 1$. Then, obviously, PR/R is a Sylow p-subgroup of G/R. Suppose that M/R is a maximal subgroup of PR/R. Then there exists a maximal subgroup P_1 of P such that $M = P_1R$. If P_1 has a p-nilpotent supplement D in G, then M/R has a p-nilpotent supplement DR/R in G/R. If P_1 is partially τ-quasinormal in G, then M/R is partially τ-quasinormal in G/R by Lemma 2.3(3). The minimal choice of G implies that G/R is soluble. By the well known Feit-Thompson's theorem, we know that R is soluble. It follows that G is soluble, a contradiction.

(3) P is not cyclic.

If P is cyclic, then G is p-nilpotent by Lemma 2.4, and so G is soluble, a contradiction.

(4) If N is a minimal normal subgroup of G, then N is not soluble. Moreover, $G = PN$.

If N is p-soluble, then $O_p(N) \neq 1$ or $O_{p'}(N) \neq 1$. Since $O_p(N)$ char $N \trianglelefteq G$, $O_p(N) \leq O_p(G)$. Analogously $O_{p'}(N) \leq O_{p'}(G)$. Hence $O_p(G) \neq 1$ or $O_{p'}(G) \neq 1$, which contradicts step (1) or step (2). Therefore N is not soluble. Assume that $PN < G$. By Lemma 2.3(1), every maximal subgroup of P not having a p-nilpotent supplement in PN is partially τ-quasinormal in PN. Thus PN satisfies the hypothesis. By the minimal choice of G, PN is soluble and so N is soluble. This contradiction shows that $G = PN$.

(5) G has a unique minimal normal subgroup N.

By step (4), we see that $G = PN$ for every normal subgroup N of G. It follows that G/N is soluble. Since the class of all soluble groups is closed under subdirect product, G has a unique minimal normal subgroup, say N.

(6) The final contradiction.

If every maximal subgroup of P has a p-nilpotent supplement in G, then, in view of Lemma 2.5, G is p-nilpotent and so G is soluble. This contradiction shows that we may choose a maximal subgroup P_1 of P such that P_1 is partially τ-quasinormal in G. Then there exists a normal subgroup T of G such that P_1T is S-quasinormal in G and $P_1 \cap T \leq (P_1)_{\tau G}$. If $T = 1$, then P_1 is S-quasinormal in G. In view of Lemma 2.6, $P_1 \trianglelefteq PO^p(G) = G$. By step (5), $P_1 = 1$ or $N \leq P_1$. Since N is not soluble by step (4), we have that $P_1 = 1$. Consequently, P is cyclic, which contradicts step (3). Hence $T \neq 1$ and $N \leq T$. It follows that $P_1 \cap N = (P_1)_{\tau G} \cap N$. For any Sylow q-subgroup N_q of N with $q \neq p$, N_q is also a Sylow q-subgroup of G by step (4). From step (2) it is easy to see that $(P_1)_{\tau G}N_q = N_q(P_1)_{\tau G}$. Then $(P_1)_{\tau G}N_q \cap N = N_q((P_1)_{\tau G} \cap N) = N_q(P_1 \cap N)$, i.e., $P_1 \cap N$ is τ-quasinormal in N. Since N is a direct product of some isomorphic non-abelian simple groups, we may assume that $N \cong N_1 \times \cdots \times N_k$. By Lemma 2.2(1), $P_1 \cap N$ is τ-quasinormal in $(P_1 \cap N)N_1$. Thus $(P_1 \cap N)(N_{1q})^{n_1} \cap N_1 = (N_{1q})^{n_1}(P_1 \cap N \cap N_1) = (N_{1q})^{n_1}(P_1 \cap N_1)$ for any $n_1 \in N_1$, where N_{1q} is a Sylow q-subgroup of N_1 with $q \neq p$. Since $(N_{1q})^{n_1}(P_1 \cap N_1) \neq N_1$, we have N_1 is not simple by Lemma 2.7, a contradiction.

Proof of Theorem 1.4. If G is p-nilpotent, then G has a normal Hall p'-subgroup $G_{p'}$. Let P_1 be any maximal subgroup of P. Then $|G : P_1G_{p'}| = p$. In view of Lemma 2.4(3), $P_1G_{p'} \trianglelefteq G$. Obviously, $P_1 \cap G_{p'} = 1$. Hence P_1 is partially τ-quasinormal in G.

Now we prove the sufficient part. Assume that the assertion is false and let G be a counterexample with minimal order.

(1) G is soluble.

It follows directly from Theorem 1.3.

(2) G has a unique minimal normal subgroup N such that G/N is p-nilpotent. Moreover, $\Phi(G) = 1$.

Let N be a minimal normal subgroup of G. Since G is solvable by step (1), N is an elementary abelian subgroup. It is easy to see that G/N satisfies the hypothesis of our theorem by Lemma 2.3. By the minimal choice of G, G/N is p-nilpotent. Since the class of all p-nilpotent groups is a saturated formation, N is a unique minimal normal subgroup of G and $\Phi(G) = 1$.

(3) P is not cyclic.

If P is cyclic, G is p-nilpotent by Lemma 2.4(2), a contradiction.

(4) $O_{p'}(G) = 1$.

(5) Every maximal subgroup of P has a p-nilpotent supplement in G.

It is clear that $N \leq O_p(G)$. By $\Phi(G) = 1$, we may choose a maximal subgroup M of G such that $G = NM$ and $G/N \cong M$. Let P_1 be an arbitrary maximal subgroup of P. We will show P_1 has a p-nilpotent supplement in G. Since N has the p-nilpotent supplement M in G, we only need to prove $N \leq P_1$ when P_1 is partially τ-quasinormal in G. Let T be a normal subgroup of G such that P_1T is S-quasinormal in G and $P_1 \cap T \leq (P_1)_{\tau G}$. First, we assume that $T = 1$, i.e., P_1 is S-quasinormal in G. In view of Lemma 2.6, $P_1 \trianglelefteq PO^p(G) = G$. By virtue of Lemma 2.4(2) and step (3), $P_1 \neq 1$. Hence $N \leq P_1$ by step (2). Now, assume that $T \neq 1$. Then $N \leq T$. It follows that $P_1 \cap N = (P_1)_{\tau G} \cap N$. For any Sylow q-subgroup G_q of G ($p \neq q$), $(P_1)_{\tau G}G_q = G_q(P_1)_{\tau G}$ in view of step (4). Then $(P_1)_{\tau G} \cap N = (P_1)_{\tau G}G_q \cap N \trianglelefteq (P_1)_{\tau G}G_q$. Obviously, $P_1 \cap N \trianglelefteq P$. Therefore $P_1 \cap N$ is normal in G. By the minimality of N, we have $P_1 \cap N = N$ or $P_1 \cap N = 1$. If the later holds, then the order of N is p since $P_1 \cap N$ is a maximal subgroup of N. Consequently, G is p-nilpotent by step (2) and Lemma 2.4(1). This contradiction shows that $P_1 \cap N = N$ and so $N \leq P_1$.

(6) The final contradiction.

Since every maximal subgroup of P has a p-nilpotent supplement in G by step (5), we have G is p-nilpotent by Lemma 2.5, a contradiction.

Proof of Theorem 1.5. Assume that this theorem is false and and consider a counterexample (G, E) for which $|G||E|$ is minimal.

(1) E is p-nilpotent.

Let P_1 be a maximal subgroup of P. If P_1 has a p-supersolvable supplement T in G, then P_1 has a p-supersolvable supplement $T \cap E$ in E. Since $(|E|, p-1) = 1$, $T \cap E$ is also p-nilpotent by Lemma 2.4(4). If P_1 is partially τ-quasinormal in G, then P_1 is also partially τ-quasinormal in E by Lemma 2.3(1). Hence every maximal subgroup of P not having a p-nilpotent supplement in E is partially τ-quasinormal in E. In view of Theorem 1.4, E is p-nilpotent.

(2) $P = E$.

By step (1), $O_{p'}(E)$ is the normal Hall p'-subgroup of E. Suppose that $O_{p'}(E) \neq 1$. It is easy to see that the hypothesis of the theorem holds for $(G/O_{p'}(E), E/O_{p'}(E))$. By induction, every chief factor of $G/O_{p'}(E)$ between $E/O_{p'}(E)$ and 1 is cyclic. Consequently, each chief factor of G between E and $O_{p'}(E)$ is cyclic. This condition shows that $O_{p'}(E) = 1$ and so $P = E$.

(3) $\Phi(P) = 1$.

Suppose that $\Phi(P) \neq 1$. By Lemma 2.3(2), it is easy to see that the hypothesis of the theorem holds for $(G/\Phi(P), P/\Phi(P))$. By the choice of (G, E), every chief factor of $G/\Phi(P)$ below $P/\Phi(P)$ is cyclic. In view of Lemma 2.8, every chief factor of G below P is cyclic, a contradiction.

(4) Every maximal subgroup of P is partially τ-quasinormal in G.

Suppose that there is some maximal subgroup V of P such that V has a p-supersolvable supplement B in G, then $G = PB$ and $P \cap B \neq 1$. Since $P \cap B \trianglelefteq B$, we may assume that B has a minimal normal subgroup N contained in $P \cap B$. It is clear that $|N| = p$. Since P is elementary abelian and $G = PB$, we have that N is also normal in G. It is easy to see that the hypothesis is still true for $(G/N, P/N)$. Hence every chief factor of G/N below P/N is cyclic by virtue of the choice of (G, E). It follows that every chief factor of G below P is cyclic. This contradiction shows that all maximal subgroups of P are partially τ-quasinormal in G.

(5) P is not a minimal normal subgroup of G.

Suppose that P is a minimal normal subgroup of G, then some maximal subgroup of P is normal in G by Lemma 2.10, which contradicts the minimality of P.

(6) If N is a minimal normal subgroup of G contained in P, then $P/N \leq Z_{\mathscr{U}}(G/N)$, N is the only minimal normal subgroup of G contained in P and $|N| > p$.

Indeed, by Lemma 2.3(2), the hypothesis holds on $(G/N, P/N)$ for any minimal normal subgroup N of G contained in P. Hence every chief factor of G/N below P/N is cyclic by the choice of $(G, E) = (G, P)$. If $|N| = p$, every chief factor of G below P is cyclic, a contradiction. If G has two minimal normal subgroups R and N contained in P, then $NR/R \leq P/R$ and from the G-isomorphism $NR/R \cong N$ we have $|N| = p$, a contradiction. Hence, (6) holds.

(7) The final contradiction.

Let N be a minimal normal subgroup of G contained in P and N_1 any maximal subgroup of N. We show that N_1 is S-quasinormal in G. Since P is an elementary abelian p-group, we may assume that D is a complement of N in P. Let $V = N_1D$. Obviously, V is a maximal subgroup of P. By step (4), V is partially τ-quasinormal in G. By Lemma 2.3(4), there exist a normal subgroup T of G such that VT is S-quasinormal in G, $V \cap T \leq V_{\tau G}$ and $VT \leq P$. In view of Lemma 2.2(4), $V_{\tau G}$ is an S-quasinormal subgroup of G. If $T = P$, then $V = V_{\tau G}$ is S-quasinormal in G and hence $V \cap N = N_1D \cap N = N_1(D \cap N) = N_1$ is S-quasinormal in G by Lemma 2.1(5). If $T = 1$, then $V = VT$ is S-quasinormal in G. Consequently, we have also N_1 is S-quasinormal in G. Now we assume that $1 < T < P$. Hence $N \leq T$ by step (6). Then, $N_1 = V \cap N = V_{\tau G} \cap N$ is S-quasinormal in G by virtue of Lemma 2.1(5). Hence some maximal subgroup of N is normal in G by Lemma 2.9. Consequently, $|N| = p$. This contradicts step (6).

Proof of Theorem 1.6. Let q be the smallest prime dividing $|E|$. In view of step (1) of the proof of Theorem 1.5, E is q-nilpotent. Let $E_{q'}$ be the normal Hall q'-subgroup of E. If $E_{q'} = 1$, then every chief factor of G below E is cyclic by Theorem 1.5. Hence we may assume that $E_{q'} \neq 1$. Since $E_{q'}$ char $E \trianglelefteq G$, we see that $E_{q'} \trianglelefteq G$. By Lemma 2.3(3), the hypothesis of the theorem holds for $(G/E_{q'}, E/E_{q'})$. By induction, every chief factor of $G/E_{q'}$ below $E/E_{q'}$ is cyclic. On the other hand, $(G, E_{q'})$ also satisfies the hypothesis of the theorem in view of Lemma 2.3(1). By induction again, we have also every chief factor of G below $E_{q'}$ is cyclic. Hence it follows that every chief factor of G below E is cyclic.

Proof of Main Theorem. Applying Theorem 1.6, X is hypercyclically embedded in G. Since $F^*(E) \leq X$, we have that $F^*(E)$ is also hypercyclically embedded in G. By virtue of Lemma 2.11, E is also hypercyclically embedded in G.

4. Some Applications

4.1. Theorem. Let $\mathscr{F}$ be a saturated formation containing $\mathscr{U}$ and E a normal subgroup of a group G such that $G/E \in \mathscr{F}$. Suppose that for every non-cyclic Sylow subgroup P of E, every maximal subgroup of P not having a supersoluble supplement in G is partially τ-quasinormal in G. Then $G \in \mathscr{F}$.

Proof. Applying our Main Theorem, every chief factor of G below E is cyclic. Since $\mathscr{F}$ contains $\mathscr{U}$, we know E is contained in the $\mathscr{F}$-hypercentre of G. From $G/E \in \mathscr{F}$, it follows that $G \in \mathscr{F}$. □

4.2. Theorem. Let $\mathscr{F}$ be a saturated formation containing $\mathscr{U}$ and E a normal subgroup of a group G such that $G/E \in \mathscr{F}$. Suppose that for every non-cyclic Sylow subgroup P of $F^*(E)$, every maximal subgroup of P not having a supersoluble supplement in G is partially τ-quasinormal in G. Then $G \in \mathscr{F}$.

Proof. The proof is similar to that of Theorem 4.1. □

4.3. Corollary ([9, Theorem 3.4])**.** Let $\mathscr{F}$ be a saturated formation containing $\mathscr{U}$ and E a normal subgroup of a group G such that $G/E \in \mathscr{F}$. If every maximal subgroup of any Sylow subgroup of $F^*(E)$ is S-quasinormal in G, then $G \in \mathscr{F}$.

4.4. Corollary ([19, Theorem 3.4])**.** Let $\mathscr{F}$ be a saturated formation containing $\mathscr{U}$ and E a normal subgroup of a group G such that $G/E \in \mathscr{F}$. If every maximal subgroup of any Sylow subgroup of $F^*(E)$ is c-normal in G, then $G \in \mathscr{F}$.

4.5. Corollary ([1, Theorem 1.4])**.** Let $\mathscr{F}$ be a saturated formation containing $\mathscr{U}$ and E a soluble normal subgroup of a group G such that $G/E \in \mathscr{F}$. If every maximal subgroup of any Sylow subgroup of $F(E)$ is S-quasinormal in G, then $G \in \mathscr{F}$.

4.6. Corollary ([8, Theorem 2])**.** Let G be a group and E a soluble normal subgroup of G such that G/E is supersolvable. If all maximal subgroups of the Sylow subgroups of $F(E)$ are c-normal in G, then G is supersolvable.

References

[1] M. Asaad, On maximal subgroups of Sylow subgroups of finite groups, Comm. Algebra, 26(1998), 3647-3652.

[2] A. Ballester-Bolinches and Y. Wang, Finite groups with some c-normal minimal subgroups, J. Pure Appl. Algebra, 153 (2000), 121–127.

[3] W. E. Deskins, On quasinormal subgroups of finite groups, Math. Z. 82 (1963), 125-132.

[4] W. Guo, The Theory of Class of Groups, Science Press-Kluwer Academic Publishers, Beijing-New York-Dordrecht-Boston-London, 2000.

[5] X. Guo and K. P. Shum, On c-normal maximal and minimal subgroups of Sylow p-subgroups of finite groups, Arch. Math. 80 (2003), 561–569.

[6] B. Huppert, Endliche Gruppen I, Springer-Verlag, Berlin-Heidelberg-New York, 1967.

[7] O. H. Kegel, Sylow Gruppen und subnormalteiler endlicher Gruppen, Math. Z. 78 (1962), 205–221.

[8] D. Li and X. Guo, The influence of c-normality of subgroups on the structure of finite groups II, Comm. Algebra, 26 (1998), 1913–1922.

[9] Y. Li, H. Wei and Y. Wang, The influence of π-quasinormality of some subgroups of a finite griup, Arch. Math. 81 (2003), 245–252.

[10] C. Li, Finite groups with some primary subgroups SS-quasinormally embedded, Indian J. Pure Appl. Math. 42 (2011) 291–306.

[11] V. O. Lukyanenko and A. N. Skiba, On weakly τ-quasinormal subgroups of finite groups, Acta Math. Hungar, 125 (2009), 237–248.

[12] V. O. Lukyanenko and A. N. Skiba, Finite groups in which τ-quasinormality is a transitive relation, Rend Sem Mat Univ Padova, 124 (2010), 231–246.

[13] P. Schmidt, Subgroups permutable with all Sylow subgroups, J. Algebra, 207 (1998), 285–293.

[14] A. N. Skiba, On weakly S-permutable subgroups of finite groups, J. Algebra, 315 (2007), 192–209.

[15] A. N. Skiba, On two questions of L. A. Shemetkov concerning hypercyclically embedded subgroups of finite groups, J. Group Theory, 13 (2010), 841–850.

[16] A. N. Skiba, Cyclicity conditions for G-chief factors of normal subgroups of a group G, Siberian Math. J. 52 (2011), 127–130.

[17] Y. Wang, C-normality and groups and its properties, J. Algebra, 180 (1996), 954–965.

[18] H. Wei and Y. Wang, On c^*-normality and its properties, J. Group Theory, 10 (2007), 211-223.

[19] H. Wei, Y. Wang and Y. Li, On c-normal maximal and minimal subgroups of Sylow subgroups of finite groups II, Comm. Algebra, 31 (2003), 4807–4816.

[20] M. Weinstein, Between Nilpotent and Solvable, Polygonal Publishing House, Passaic, 1982.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 963 – 969

Common fixed and coincidence point theorems for maps in Menger space with Hadzic type t - norm

K.P.R.Rao* , K.Bhanu Lakshmi† and N.Srinivasa Rao‡

Abstract

In this paper, we obtain a unique common fixed point theorem for two weakly compatible mappings in a Menger space and also obtain a common coincidence point theorem for two hybrid pairs of mappings.

2000 AMS Classification: 47H10, 54H25.

Keywords: Menger space , Hadzic type t-norm, weakly compatible mappings.

Received 30 : 01 : 2013 : Accepted 18 : 11 : 2013 Doi : 10.15672/HJMS.2014437530

1. Introduction and preliminaries

In 1942, Menger [6] introduced the notion of a statistical metric space as a generalization of a metric space (M, d) in which the distance $d(x, y)$, $(x, y \in M)$ between x and y is replaced by a distribution function $F_{x,y}$. Schweizer and Sklar [9] studied this concept and established some fundamental results on this space . First , we give some known preliminaries.

1.1. Definition. . A mapping $F : R \to [0, 1]$ is said to be a distribution function if

(i) F is non-decreasing,
(ii) F is left continuous,
(iii) $\inf_{x\in R} F(x) = 0$ and $\sup_{x\in R} F(x) = 1$.

We denote the set of all distribution functions by $\mathbb{D}$.

*Department of Mathematics, Acharya Nagarjuna University, Nagarjuna Nagar -521 510, A.P., INDIA.
Email: kprrao2004@yahoo.com
†Department of Science and Humanities, Lakireddy Balireddy College of Engineering, Mylavaram -522 230, A.P., INDIA.
Email:bhanulaks@gmail.com
‡Department of Science and Humanities, Vignan University,Vadlamudi-522 213, Guntur Dt., A.P., INDIA.
Email:srinivasunimmala@yahoo.co.in

1.2. Definition. ([9]). A probabilistic metric space is an ordered pair (M, F), where M is a non empty set and F is a function defined on $M \times M$ to $\mathbb{D}$ which satisfies the following conditions: For $x, y, z \in M$,

(i) $F_{x,\ y}(0) = 0$,
(ii) $F_{x,\ y}(s) = 1$ for all $s > 0$ if and only if $x = y$,
(iii) $F_{x,\ y}(s) = F_{y,\ x}(s)$ for all $s \in R$ and
(iv) $F_{x,\ y}(s_1) = 1$ and $F_{y,\ z}(s_2) = 1$ for all $s_1, s_2 > 0$ imply $F_{x,\ z}(s_1 + s_2) = 1$.

1.3. Definition. ([9]). A function $t : [0, 1] \times [0, 1] \to [0, 1]$ is said to be a triangular norm or t - norm if it satisfies the following conditions: For $a, b, c, d \in [0, 1]$,

(i) $t(a, 1) = a$,
(ii) $t(a, b) = t(b, a)$,
(iii) $t(c, d) \geq t(a, b)$ if $c \geq a$ and $d \geq b$,
(iv) $t(t(a, b), c) = t(a, t(b, c))$.

1.4. Definition. ([9]). Let M be a nonempty set, ' t ' is a t - norm and $F : M \times M \to \mathbb{D}$ satisfy:

(i) $F_{x,y}(0) = 0$ for all $x, y \in M$,
(ii) $F_{x,y}(s) = 1$ for all $s > 0$ if and only if $x = y$,
(iii) $F_{x,y}(s) = F_{y,x}(s)$ for all $s \in R$ and
(iv) $F_{x,y}(u + v) \geq t(F_{x,z}(u), F_{z,y}(v))$ for all $u, v \geq 0$ and $x, y, z \in M$.

Then the triplet (M, F, t) is called a Menger space.

1.5. Remark. If (M, d) is a metric space then ‘ d ’ induces a mapping $F : M \times M \to \mathbb{D}$, where F is defined by $F_{p,q}(x) = H(x - d(p, q))$, where $H(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0 \end{cases}$ is the Heaviside function.

Further, if $t : [0, 1] \times [0, 1] \to [0, 1]$ is defined by $t(a, b) = \min\{a, b\}$, then (M, F, t) is a Menger space. It is complete if the metric space (M, d) is complete.

1.6. Definition. ([9]). Let (M, F, t) be a Menger space. Let $x \in M$. For $\epsilon > 0$ and $0 < \lambda < 1$, the (ϵ, λ) - neighbourhood of x is defined as
$N_x(\epsilon, \lambda) = \{y \in M : F_{x,y}(\epsilon) > 1 - \lambda\}$.

The topology induced by the family $\{N_p(\epsilon, \lambda) : p \in M, \epsilon > 0, 0 < \lambda < 1\}$ is known as the (ϵ, λ)- topology.

1.7. Proposition. ([9]). If t is continuous then (ϵ, λ)- topology is a Hausdorff topology on M.

1.8. Definition. ([9]). Let (M, F, t) be a Menger space. A sequence $\{x_n\}$ in M converges to $x \in M$, if for any $\epsilon > 0$ and $0 < \lambda < 1$, there exists a positive integer $N = N(\epsilon, \lambda)$ such that $F_{x_n,x}(\epsilon) > 1 - \lambda$ for all $n \geq N$. A sequence $\{x_n\}$ in (M, F, t) is said to be Cauchy sequence in M if for $\epsilon > 0$ and $0 < \lambda < 1$, there exists a positive integer $N = N(\epsilon, \lambda)$ such that $F_{x_n,x_m}(\epsilon) > 1 - \lambda$ for all $n, m \geq N$. A Menger space (M, F, t), where t is continuous, is said to be complete if every Cauchy sequence in M is convergent in (ϵ, λ) - topology.

In 1972, Sehgal and Reid [10] introduced the notion of contraction mapping on a probabilistic metric space and proved fixed point theorems for such mappings.

1.9. Definition. ([10]). Let (M, F, t) be a Menger space. A map $T : M \to M$ is said to be a contraction mapping if there exists a constant $0 < p < 1$ such that $F_{Tx,Ty}(s) \geq F_{x,y}(\frac{s}{p})$ for each $x, y \in M$ and for all $s > 0$.

1.10. Theorem. ([10]). Let (M, F, t) be a complete Menger space, where 't' is a continuous function satisfying $t(x, x) \geq x$ for each $x \in [0, 1]$. If $T : M \to M$ is a contraction mapping then there is a unique $p \in M$ such that $Tp = p$. Moreover $T^n q \to p$ for each $q \in M$.

In 1978, Hadzic [4] introduced a class $\mathcal{F}$ of t- norms $t \neq t_{\min}$, for which every contraction in a complete Menger space (M, F, t) has a fixed point.

1.11. Definition. ([4]). We say that the t - norm t is of Hadzic - type and we write $t \in \mathcal{F}$ if the family $\{t^n\}_{n \in N}$ of it's iterates defined, for each $x \in [0, 1]$ by $t^0(x) = 1$ and $t^{n+1}(x) = t(t^n(x), x)$ for all $n \geq 0$ is equicontinuous at $x = 1$.

i.e., for each $\epsilon \in (0, 1)$, there exists $\delta \in (0, 1)$ such that $x > 1 - \delta$ implies $t^n x > 1 - \epsilon$ for all $n \geq 1$.

1.12. Theorem. ([4]). Let (M, F, t) be a complete Menger space, where 't' is a continuous t - norm of Hadzic type. If $T : M \to M$ is a contraction mapping then there is a unique $p \in M$ such that $Tp = p$. Moreover $T^n q \to p$ for each $q \in M$.

Recently Choudhury and Das [1], proved the following

1.13. Theorem. ([1]). Let (M, F, t_M) be a complete Menger space with continuous t-norm t_M given by $t_M(a, b) = \min\{a, b\}$ and $f : M \to M$ be satisfying $F_{fx,fy}(\varphi(s)) \geq F_{x,y}(\varphi(\frac{s}{c}))$ for all $x, y \in M$ and for $s \geq 0$, where $0 < c < 1$ and $\varphi : R \to R^+$ satisfies

(i) $\varphi(t) = 0$ iff $t = 0$,
(ii) $\varphi(t)$ is increasing and $\varphi(t) \to \infty$ as $t \to \infty$,
(iii) φ is left continuous on $(0, \infty)$,
(iv) φ is continuous at 0 .

Then f has a unique fixed point in M.

Later several authors obtained fixed point theorems in Menger spaces using an altering distance function,for example refer [2],[3],[7]etc.
Sastry et.al. [8] , defined altering function of type (S) as follows :

1.14. Definition. ([8]) A function $\varphi : R^+ \to R^+$ is said to be an altering distance function of type (S) if it satisfies

(i) $\varphi(t) = 0$ iff $t = 0$,
(ii) $\varphi(t) \to \infty$ as $t \to \infty$,
(iii) φ is continuous at 0 .

1.15. Lemma. ([8]) Let (M, F, t) be a Menger space with a continuous Hadzic type t - norm , $0 < c < 1$ and φ be an altering distance function of type (S). Suppose $\{x_n\}_{n=0}^{\infty}$ is a sequence in M such that for any $r > 0$, $F_{x_n,x_{n+1}}(\varphi(r)) \geq F_{x_0,x_1}(\varphi(\frac{r}{c^n}))$. Then $\{x_n\}$ is a Cauchy sequence.

1.16. Theorem. ([8]) Let (M, F, t) be a complete Menger space with a continuous Hadzic type t - norm 't' and φ be an altering distance function of type (S), $P : M \to M$ be satisfying $F_{Px,Py}(\varphi(s)) \geq F_{x,y}(\varphi(\frac{s}{c}))$ for all $x, y \in M$ and for $s > 0$ and $0 < c < 1$. Then P has a unique fixed point $z \in M$. Moreover, $P^n x \to z$ for each $x \in M$.

1.17. Definition. ([5]) A pair of self mappings is called weakly compatible if they commute at their coincidence points.

In this paper, we extend Theorem 1.16 for two pairs of weakly compatible mappings.

2. Main results

2.1. Theorem. Let (M, F, t) be a Menger space with continuous Hadzic type t-norm 't' and φ be an altering distance function of type (S). Let $P, Q, f, g : M \to M$ be maps such that

(2.1.1) $F_{Px,Qy}(\varphi(s)) \geq F_{fx,gy}(\varphi(\frac{s}{c}))$ for all $x, y \in M$ and for $s > 0$ and $0 < c < 1$.
(2.1.2) $P(M) \subseteq g(M)$, $Q(M) \subseteq f(M)$,
(2.1.3) either $f(M)$ or $g(M)$ is complete,
(2.1.4) the pairs (f, P) and (g, Q) are weakly compatible .

Then f, g, P and Q have a unique common fixed point in M .

Proof. Let $x_0 \in M$.
Since $P(M) \subseteq g(M)$, there exists $x_1 \in M$ such that $y_1 = gx_1 = Px_0$.
Since $Q(M) \subseteq f(M)$, there exists $x_2 \in M$ such that $y_2 = fx_2 = Qx_1$.
Continuing in this way, we get sequences $\{x_n\}$ and $\{y_n\}$ in M such that $y_{2n+1} = gx_{2n+1} = Px_{2n}$ and $y_{2n+2} = fx_{2n+2} = Qx_{2n+1}$, $n = 0, 1, 2 \cdots$
Since φ is continuous at 0 and vanishes only at 0, it follows that for given $s > 0$ there exists $r > 0$ such that $\frac{s}{2} > \varphi(r)$. Now

$$\begin{aligned} F_{y_{2n+1},\ y_{2n+2}}(s) &\geq F_{y_{2n+1},\ y_{2n+2}}(\varphi(r)), \\ &= F_{Px_{2n},\ Qx_{2n+1}}(\varphi(r)), \\ &\geq F_{fx_{2n},\ gx_{2n+1}}(\varphi(\tfrac{r}{c})), \\ &= F_{y_{2n},\ y_{2n+1}}(\varphi(\tfrac{r}{c})) \end{aligned}$$

Similarly,

$$F_{y_{2n+1},\ y_{2n}}(s) \geq F_{y_{2n},\ y_{2n-1}}(\varphi(\frac{r}{c})).$$

Thus

$$F_{y_{n+1},\ y_n}(s) \geq F_{y_n,\ y_{n-1}}(\varphi(\frac{r}{c})) \geq \cdots \geq F_{y_1,\ y_0}(\varphi(\frac{r}{c^n})).$$

Hence from Lemma 1.15 , $\{y_n\}$ is Cauchy.
Suppose $g(M)$ is complete.
Then there exist $z, v \in M$ such that $y_{2n+1} = gx_{2n+1} \to z = gv$.
Since $\{y_n\}$ is Cauchy , we have $y_n \to z$.
Again for given $s > 0$ there exists $r > 0$ such that $\frac{s}{2} > \varphi(r)$. Now,

$$\begin{aligned} F_{z,\ Qv}(s) &\geq t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{y_{2n+1},\ Qv}(s - \varphi(r))), \\ &\geq t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{Px_{2n},\ Qv}(\varphi(r))), \\ &\geq t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{fx_{2n},\ gv}(\varphi(\tfrac{r}{c}))), \quad \text{from} \quad (2.1.1) \\ &= t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{y_{2n},\ gv}(\varphi(\tfrac{r}{c}))), \\ &\to 1 \quad \text{as} \quad n \to \infty. \end{aligned}$$

since $y_n \to z$ and t is a continuous Hadzic type t-norm .
Thus $z = Qv$. Hence

(2.1) $\quad gv = z = Qv.$

Since $z = Qv \in Q(M) \subseteq f(M)$, there exists $u \in M$ such that

(2.2) $\quad z = fu.$

Now

$$F_{Pu,\ z}(\varphi(s)) = F_{Pu,\ Qv}(\varphi(s)) \geq F_{fu,\ gv}(\varphi(\frac{s}{c})) = F_{z,\ z}(\varphi(\frac{s}{c})) = 1.$$

Thus $Pu = z$. Hence

(2.3) $\quad Pu = z = fu.$

Since (f, P) is weakly compatible and from (2.3), we have $Pz = fz$.
Now from (2.1.1), we have
$F_{Pz,\ z}(\varphi(s)) = F_{Pz,\ Qv}(\varphi(s)) \geq F_{fz,\ gv}(\varphi(\frac{s}{c})) = F_{Pz,\ z}(\varphi(\frac{s}{c}))$

$$\geq F_{Pz,\ z}(\varphi(\frac{s}{c^2})) \cdots \geq F_{Pz,\ z}(\varphi(\frac{s}{c^n})) \to 1 \quad \text{as} \quad n \to \infty.$$

Thus $Pz = z$. Hence

(2.4) $\quad z = Pz = fz.$

Since (g, Q) is weakly compatible, from(2.1), we have $gz = Qz$.
From (2.1.1), we have
$F_{z,\ Qz}(\varphi(s)) = F_{Pu,\ Qz}(\varphi(s)) \geq F_{fu,\ gz}(\varphi(\frac{s}{c})) = F_{z,\ Qz}(\varphi(\frac{s}{c}))$

$$\geq F_{z,\ Qz}(\varphi(\frac{s}{c^2})) \cdots \geq F_{z,\ Qz}(\varphi(\frac{s}{c^n})) \to 1 \quad \text{as} \quad n \to \infty.$$

Thus $Qz = z$. Hence

(2.5) $\quad z = Qz = gz.$

From (2.4) and (2.5), z is a common fixed point of P, Q, f and g.
Suppose z' is another common fixed point of P, Q, f and g. Then
From (2.1.1), we have
$F_{z,\ z'}(\varphi(s)) = F_{Pz,\ Qz'}(\varphi(s)) \geq F_{fz,\ gz'}(\varphi(\frac{s}{c})) = F_{z,\ z'}(\varphi(\frac{s}{c}))$

$$\geq F_{z,\ z'}(\varphi(\frac{s}{c^2})) \cdots \geq F_{z,\ z'}(\varphi(\frac{s}{c^n})) \to 1 \quad \text{as} \quad n \to \infty.$$

Thus $z' = z$. Hence z is the unique common fixed point of P, Q, f and g.
Similarly the theorem holds when $f(M)$ is complete. □

Recently, Sastry et.al. [8] proved the following theorem for a multivalued map in a complete Menger space with Hadzic type t-norm.

2.2. Theorem. ([8]). Let (M, F, t) be a complete Menger space with a continuous Hadzic type t - norm 't' , φ be an altering distance function of type (S) and P be a multivalued map of M into the class of nonempty subsets of M. Suppose that there exists $0 < c < 1$ such that for any $x, y \in M$, $F_{u,\ v}(\varphi(s)) \geq F_{x,\ y}(\varphi(\frac{s}{c}))$ for all $s > 0$, whenever $u \in Px$, $v \in Py$.
Then P has a unique fixed point $z \in M$ and $Pz = \{z\}$.

Now we extend this theorem for two pairs of hybrid mappings.

2.3. Definition. Let (M, F, t) be a Menger space and $f : M \to M$, P be a multi valued map of M into the class of nonempty subsets of M. Then f is said to be P- weakly commuting at $x \in M$ if $f^2x \in Pfx$.

2.4. Theorem. Let (M, F, t) be a Menger space with a continuous Hadzic type t - norm 't' and φ be an altering distance function of type (S). Let P and Q be multivalued maps of M into the class of nonempty subsets of M and f and g be self maps on M. Suppose that there exists $0 < c < 1$ such that for any $x, y \in M$,

(2.4.1) $F_{u,\ v}(\varphi(s)) \geq F_{fx,\ gy}(\varphi(\frac{s}{c}))$ for all $s > 0$, whenever $u \in Px$, $v \in Qy$.
(2.4.2) $P(M) \subseteq g(M), \quad Q(M) \subseteq f(M)$,
(2.4.3) either $f(M)$ or $g(M)$ is complete,
(2.4.4) f is P-weakly commuting and g is Q-weakly commuting at their coincidence points.

Then the pairs (f, P) and (g, Q) have a common coincidence point in M.

Proof. Let $x_0 \in M$.
Since $P(x_0) \subseteq g(M)$, there exists $x_1 \in M$ such that $y_1 = gx_1 \in Px_0$.
Since $Q(x_1) \subseteq f(M)$, there exists $x_2 \in M$ such that $y_2 = fx_2 \in Qx_1$.
Continuing in this way, we get sequences $\{x_n\}$ and $\{y_n\}$ in M such that $y_{2n+1} = gx_{2n+1} \in Px_{2n}$ and $y_{2n+2} = fx_{2n+2} \in Qx_{2n+1}$, $n = 0, 1, 2 \cdots$

$$\begin{aligned} F_{y_{2n+1},\ y_{2n+2}}(\varphi(s)) &\geq F_{fx_{2n},\ gx_{2n+1}}(\varphi(\tfrac{s}{c})), \\ &= F_{y_{2n},\ y_{2n+1}}(\varphi(\tfrac{s}{c})) \end{aligned}$$

Similarly,

$$F_{y_{2n},\ y_{2n+1}}(\varphi(s)) \geq F_{y_{2n-1},\ y_{2n}}(\varphi(\frac{s}{c})).$$

Thus

$$F_{y_n,\ y_{n+1}}(\varphi(s)) \geq F_{y_{n-1},\ y_n}(\varphi(\frac{s}{c}))$$

Since φ is continuous at 0 and vanishes only at 0, it follows that for given $s > 0$ there exists $r > 0$ such that $\frac{s}{2} > \varphi(r)$. Now

$$F_{y_n,\ y_{n+1}}(s) \geq F_{y_n,\ y_{n+1}}(\varphi(r)) \geq F_{y_{n-1},\ y_n}(\varphi(\frac{r}{c})) \geq \cdots \geq F_{y_0,\ y_1}(\varphi(\frac{r}{c^n})).$$

Hence from Lemma 1.15 , $\{y_n\}$ is Cauchy sequence in M.
Suppose $f(M)$ is complete.
Then there exist $z, p \in M$ such that $y_n \to z = fp$.
Let $z_1 \in Pp$. Since $y_{2n+2} = fx_{2n+2} \in Qx_{2n+1}$, from (2.4.1),we have

$$\begin{aligned} F_{fp,\ z_1}(s) &\geq t(F_{fp,\ fx_{2n+2}}(\varphi(r)), F_{fx_{2n+2},\ z_1}(s - \varphi(r))), \\ &\geq t(F_{z,\ y_{2n+2}}(\varphi(r)), F_{fx_{2n+2},\ z_1}(\varphi(r))), \\ &\geq t(F_{z,\ y_{2n+2}}(\varphi(r)), F_{fp,\ gx_{2n+1}}(\varphi(\tfrac{r}{c}))), \\ &= t(F_{z,\ y_{2n+2}}(\varphi(r)), F_{z,\ y_{2n+1}}(\varphi(\tfrac{r}{c}))), \\ &\to 1 \quad \text{as} \quad n \to \infty. \end{aligned}$$

since $y_n \to z$ and t is a continuous Hadzic type t-norm .
Thus $F_{fp,\ z_1}(s) = 1$ for $s > 0$ so that $fp = z_1$. Thus

(2.6) $\quad fp \in Pp.$

Since $z = fp \in Pp \subseteq g(M)$, there exists $q \in M$ such that $z = fp = gq$.
Let $z_2 \in Qq$. Since $y_{2n+1} = gx_{2n+1} \in Px_{2n}$, from (2.4.1), we have

$$\begin{aligned} F_{gq,\ z_2}(s) &\geq t(F_{gq,\ gx_{2n+1}}(\varphi(r)), F_{gx_{2n+1},\ z_2}(s - \varphi(r))), \\ &\geq t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{gx_{2n+1},\ z_2}(\varphi(r))), \\ &\geq t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{fx_{2n},\ gq}(\varphi(\tfrac{r}{c}))), \\ &= t(F_{z,\ y_{2n+1}}(\varphi(r)), F_{y_{2n},\ z}(\varphi(\tfrac{r}{c}))), \\ &\to 1 \quad \text{as} \quad n \to \infty. \end{aligned}$$

since $y_n \to z$ and t is a continuous Hadzic type t-norm .
Thus $F_{gq,\ z_2}(s) = 1$ for $s > 0$ so that $gq = z_2$. Thus

(2.7) $\quad gq \in Qq.$

From (2.6) and (2.7), p is a coincidence point of f and P ; q is a concidence point of g and Q.
From (2.4.4), $fz \in Pz$ and $gz \in Qz$. Thus z is a common coincidence point of the hybrid pairs (f, P) and (g, Q).

□

Acknowledgement . The authors are thankful to the referee for his valuable suggestions.

References

[1] Choudhury, B.S. and Das, Krishnapada. A new contraction principle in Menger spaces , Acta Mathematica Sinica , English Series, 24(8),(2008), 1379-1386.

[2] Choudhury,B.S.,Das,K. and Dutta,P.N.,A fixed point result in Menger spaces using a real function, Acta Math.Hungarica,122(3),(2009),203-216.

[3] Dutta,P.N.,Choudhury,B.S. and Das,K.,Some fixed point results in Menger spaces using a control function , Surveys in Mathematics and its applications,4,(2009),41-52.

[4] Hadzic, O. On $(\epsilon - \lambda)$ - topology in probabilistic locally convex spaces, Glasnik Matem 13 (33), (1978), 193 - 297.

[5] Jungck,G and Rhoades,B.E. , Fixed points for set valued functions without continuity condition , Indian.J.Pure.Appl.Math., 29(3),(1998),227-238.

[6] Menger, K. Statistical metrics, Proc.of the National Academy of Sciences of the United States of America , 28,(1942),535-537.

[7] Mihet,D,Altering distances in probabilistic Menger spaces, Nonlinear Anal.,71,(2009),2734-2738.

[8] Sastry, K.P.R., Babu, G.V.R. and Sandhya, M.L. Fixed point theorems in Menger spaces for a contractive map under the influence of an altering distace function of type(S)(Communicated).

[9] Schweizer, B. and Sklar, A. Statistical metric spaces, Pacific J.Math., 10 (1960), 313 - 334.

[10] Sehgal, V.M. and Bharucha - Reid, A.T. Fixed points of contraction mappings on probabilistic metric spaces, Math. Systems theory, 6 (1972), 97 - 102.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 971 – 984

An investigation on homomorphisms and subhyperalgebras of Σ-hyperalgebras

S. Rasouli* , D. Heidari† and B. Davvaz‡

Abstract

In this paper, we introduce the notion of Σ-hyperalgebras for an arbitrary signature Σ and provide some examples. Then we extend the notions of several kinds of homomorphisms and study their properties. Also, we study subhyperalgebras of a Σ-hyperalgebra $\mathfrak{A}$, $Sub(\mathfrak{A})$, under algebraic closure operators S, H and I. Finally, we introduce the notions of closed, invertible, ultraclosed and conjugable subhyperalgebras and investigate their connections to each other.

2000 AMS Classification: 20N20, 08A05, 08A30, 03C05.

Keywords: universal algebra, hyperalgebra, homomorphism, subhyperalgebra.

Received 13 : 02 : 2012 : Accepted 13 : 10 : 2013 Doi : 10.15672/HJMS.2014437521

1. Introduction

The theory of hypergroups was introduced by F. Marty in 1934,(see [22]). Hyperalgebras (or multialgebras) are particular relational systems which generalize the concept of universal algebras. The hyperstructure theory and its applications so far, have been investigated by many mathematicians in various fields, for example in graphs and hypergraphs theory [4], in categories theory [12, 25] and in n-ary hyperalgebras [8, 21]. Recent book [7] contains wealthy applications. There are also applications to the following subjects: geometry, hypergraphs, binary relations, combinatorics, codes, cryptography, probability, etc.

There are several types of homomorphisms have been considered since the first papers on hypergroups (for instance, by M. Dresher, O. Ore [11], M. Krasner [18], J. Kuntzmann [19]) and later by M. Koskas [17]. However, the first explicit construction of hypergroup

*Department of Mathematics, Persian Gulf University, Bushehr, 75169, IRAN
Email: srasouli@pgu.ac.ir
†Department of Mathematics, Najafabad Branch Islamic Azad University Najafabad, Isfahan, Iran
Email:dheidari82@gmail.com
‡Department of Mathematics, Yazd University, Yazd, IRAN
Email:davvaz@yazduni.ac.ir

homomorphisms was given by P. Corsini [2]. A unified theory of various types of homomorphisms was given by J. Jantosciak [16]. Some other types of homomorphisms and connections among them were studied by V. Leoreanu [20]. There are more than 10 types of hypergroup homomorphisms. A detailed presentation of all these homomorphisms, connections between them and various examples can be found in [3].

Also, in the hypergroup theory there are several kinds of subhypergroups. Among the mathematicians who studied this topic, we mention F. Marty [22], M. Dresher, O. Ore [11], M. Krasner [18] who analyzed closed and invertible subhypergroups. Later Y. Sureau [32] has studied ultraclosed, invertible and conjugable subhypergroups.

On the other hand, in 1962 G. Grätzer began to study of multialgebras [12] and it follows by H. E. Pickett [26] who investigated homomorphisms and subalgebras of multialgebras. Then H. Höft and P. E. Howard [14] and G. E. Hansoul [15] studied multialgebras and T. Vougiouklis [33] discussed about the representations of hyperstructures. Furthermore, C. Pelea studied the fundamental relation of multialgebras [23] and investigated them on the universal algebras point of view in [24, 25]. Thus it seems adequate to undertake a study of multialgebras in the spirit of universal algebras.

In this article, we introduce the notion of Σ-hyperalgebra, $\mathfrak{A}$, for an arbitrary signature Σ. Then we generalize notions of various types of homomorphisms and subhyperalgebras of $\mathfrak{A}$ and we study their properties.

2. A brief excursion into hyperalgebras

This section contains a survey of the basic elements of universal algebras and hyperalgebras which will be used in the next sections. In fact, we explain what is meant by a hyperalgebra and then give several examples of familiar hyperalgebras. These examples show that different hyperalgebras may have several common properties. This observation provides a motivation for the study of Σ-hyperalgebras.

In the sequel A is a fixed nonempty set, $\mathcal{P}^*(A)$ is the family of all nonempty subsets of A, w is the set of positive integers and for $n \in w$ we denote the set of n-tuples over A, By A^n. Also, by $B \subseteq_w A$ we mean that B is a finite subset of A.

Let A be a nonempty set. A family $\mathcal{C}$ of subsets of A that is closed under the intersection of arbitrary subfamilies is called a closed-set system over A. If $\mathcal{C}$ is also closed under the union of subfamilies that are directed under inclusion, then it is called an algebraic closed-set system. An algebraic closed set system $\mathcal{C}$ forms an algebraic lattice $(\mathcal{C}, \cap, \vee)$ under set-theoretic inclusion. The closure operator associated with a given closed set system $\mathcal{C}$ is denoted by $Cl_{\mathcal{C}}$.

An n-ary hyperoperation (or function) on A is a function σ from A^n to $\mathcal{P}^*(A)$; n is the arity (or rank) of σ. A finitary hyperoperation is an n-ary hyperoperation, for some n. The image of $(a_1, \cdots, a_n)$ under an n-ary hyperoperation σ is denoted by $\sigma(a_1, \cdots, a_n)$. A hyperoperation σ on A is called a nullary hyperoperation (or constant) if its arity is zero; in fact, A nullary hyperoperation on A is just an element of $\mathcal{P}^*(A)$; i.e. a nonvoid subset of A. A hyperoperation σ on A is unary, binary or ternary if its arity is $1, 2$ or 3, respectively.

A signature or language type is a set Σ together with a mapping $\rho : \Sigma \longrightarrow w$. The elements of Σ are called hyperoperation symbols. For each $\sigma \in \Sigma$, $\rho(\sigma)$ is called the arity or rank of σ. In the sequel, for each $n \in w$, $\Sigma_n = \{\sigma | \rho(\sigma) = n\}$.

Let Σ be a signature. A Σ-hyperalgebraic structure is an ordered couple $\mathfrak{A} = (A, (\sigma^{\mathfrak{A}} : \sigma \in \Sigma))$, where A is a nonempty set and $\sigma^{\mathfrak{A}}$ is a function from $A^{\rho(\sigma)}$ to $\mathcal{P}^*(A)$, for all $\sigma \in \Sigma$. The set A is called the universe (or underlying set) of $\mathfrak{A}$ and the $\sigma^{\mathfrak{A}}$'s are called the fundamental hyperoperations of $\mathfrak{A}$. In the following, we prefer to write just σ for $\sigma^{\mathfrak{A}}$ if this convention creates an ambiguity which seldom causes a problem.

In this paper we shall use the following abbreviated notations: the sequence $x_i, \cdots, x_j$ will be denoted by x_i^j. For $j < i$ is the empty symbol. In this convention

$$\sigma^{\mathfrak{A}}(x_1, \cdots, x_i, y_{i+1}, \cdots, y_j, z_{j+1}, \cdots, z_{\rho(\sigma)})$$

will be written as $\sigma^{\mathfrak{A}}(x_1^i, y_{i+1}^j, z_{j+1}^{\rho(\sigma)})$. In this case when $y_{i+1} = \cdots = y_j = y$ the last expression will be written in the form $\sigma^{\mathfrak{A}}(x_1^i, \overset{(j-i)}{y}, z_{j+1}^{\rho(\sigma)})$. Similarly, for subsets $B_1^{\rho(\sigma)}$ of A we define

$$\sigma^{\mathfrak{A}}(B_1^{\rho(\sigma)}) = \cup\{\sigma^{\mathfrak{A}}(b_1^{\rho(\sigma)}) | b_i \in B_i,\ \forall i \in I_{\rho(\sigma)}\}.$$

Also, for each $i \in I_{\rho(\sigma)}$ and $a_1^{\rho(\sigma)} \in A$ we define

$$\sigma_i^{\mathfrak{A}}(a_1^{\rho(\sigma)}) = \{z \in A : a_i \in \sigma^{\mathfrak{A}}(a_1^{i-1}, z, a_{i+1}^{\rho(\sigma)})\}$$

and for each $i \in I_{\rho(\sigma)}$ and $B_1^{\rho(\sigma)} \subseteq A$ we define

$$\sigma_i^{\mathfrak{A}}(B_1^{\rho(\sigma)}) = \cup\{\sigma_i^{\mathfrak{A}}(b_1^{\rho(\sigma)}) | b_i \in B_i\}.$$

A hyperalgebra $\mathfrak{A}$ is unary if all hyperoperations are unary and it is mono-unary if it has just one unary hyperoperation. $\mathfrak{A}$ is a hypergroupoid if it has just one binary hyperoperation σ. According to [9], Σ-hyperalgebra $\mathfrak{A}$ is called an n-ary hypergroupoid if $\Sigma = \{\sigma\}$ and $\rho(\sigma) = n$. If σ is a binary hyperoperation, we write $a\sigma b$ for the image of (a, b) under σ. A hyperalgebra $\mathfrak{A}$ is finite if $|A|$ is finite. Let $\mathfrak{A}$ be a Σ-hyperalgebra. Hyperoperation $\sigma^{\mathfrak{A}}$ is called trivial hyperoperation if for any $(a_1^{\rho(\sigma)}) \in A^{\rho(\sigma)}$, we have $|\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})| = 1$. $\mathfrak{A}$ is called a trivial Σ-hyperalgebra if for any $\sigma \in \Sigma$, $\sigma^{\mathfrak{A}}$ is trivial.

Let $\mathfrak{A} = (A, (\sigma^{\mathfrak{A}} : \sigma \in \Sigma))$ be a Σ-hyperalgebra and $\sigma \in \Sigma$. Now, we extend an n-ary hyperoperation $\sigma^{\mathfrak{A}}$ to an n-ary operation $\sigma^{\mathfrak{P}_{\mathfrak{A}}}$ on $\mathcal{P}^*(A)$ by setting for all $(A_1^{\rho(\sigma)}) \in \mathcal{P}^*(A)^{\rho(\sigma)}$

$$\sigma^{\mathfrak{P}_{\mathfrak{A}}}(A_1^{\rho(\sigma)}) = \bigcup\{\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) | a_i \in A_i, i \in I_{\rho(\sigma)})\}.$$

It is easy to see that $\mathfrak{P}_{\mathfrak{A}} = (\mathcal{P}^*(A), < \sigma^{\mathcal{P}^*(A)} : \sigma \in \Sigma >)$ is a Σ-algebra.

2.1. Definition. Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and $j < i < \rho(\sigma)$. We say that σ is weakly (i, j)-associative if for each $a_1^{2\rho(\sigma)-1} \in A$ we have

$$\sigma^{\mathfrak{A}}(a_1^{i-1}, \sigma^{\mathfrak{A}}(a_i^{\rho(\sigma)+i-1}), a_{\rho(\sigma)+i}^{2\rho(\sigma)-1}) \cap \sigma^{\mathfrak{A}}(a_1^{j-1}, \sigma^{\mathfrak{A}}(a_j^{\rho(\sigma)+j-1}), a_{\rho(\sigma)+j}^{2\rho(\sigma)-1}) \neq \emptyset,$$

and (i, j)-associative if

$$\sigma^{\mathfrak{A}}(a_1^{i-1}, \sigma^{\mathfrak{A}}(a_i^{\rho(\sigma)+i-1}), a_{\rho(\sigma)+i}^{2\rho(\sigma)-1}) = \sigma^{\mathfrak{A}}(a_1^{j-1}, \sigma^{\mathfrak{A}}(a_j^{\rho(\sigma)+j-1}), a_{\rho(\sigma)+j}^{2\rho(\sigma)-1}).$$

We say that σ is (weakly) associative if it is (weakly) (i, j)-associative for each $i, j \in I_{\rho(\sigma)}$.

Definition 2.1 is a generalization of associative binary hyperoperation. An n-ary hypergroupoid $\mathfrak{A}$ is called an n-ary semihypergroup (n-ary H_v semigroup) if σ is (weakly) associative.

2.2. Examples. (i) [9] Let $A = \{a, b, c\}$ with the hyperoperation $\circ$ defined by the following table

$\circ$	a	b	c
a	a	b	c
b	b	$\{a, c\}$	$\{a, b\}$
c	c	$\{a, b\}$	$\{a, b\}$

Now, we define the ternary hyperoperation σ on A by $\sigma(x, y, z) = x \circ y \circ z$, for each $x, y, z \in A$. One can see that $\mathfrak{A} = (A, \sigma)$ is a 3-ary semihypergroup.

(ii) Let $A = \{a, b, c\}$ with the hyperoperation $\circ$ defined by the following table

$\circ$	a	b	c
a	a	b	c
b	b	c	$\{a, b\}$
c	c	b	$\{a, b\}$

It is easy to check that $\mathfrak{A} = (A, \circ)$ is a H_v semigroup but it is not a semihypergroup.

2.3. Definition. [8, 9] Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and $B \subseteq A$. We say that B has the reproduction axiom with respect to σ if for each $b_1^{i-1}, b, b_{i+1}^{\rho(\sigma)} \in B$ and $1 \leq i \leq \rho(\sigma)$ the relation

(2.1) $$b \in \sigma_i^{\mathfrak{A}}(b_1^{i-1}, x, b_{i+1}^{\rho(\sigma)})$$

has a solution $x \in B$. Observe that condition (1) can be reformulated as follows

(2.2) $$B = \sigma^{\mathfrak{A}}(b_1^{i-1}, B, b_{i+1}^{\rho(\sigma)-i}).$$

An n-ary semihypergroup (H_v semigroup) $\mathfrak{A}$ is called an n-ary hypergroup (H_v group) if A satisfies the reproduction axiom. An n-ary hypergropouid which satisfies the reproduction axiom is called n-ary hyperquasigroup.

2.4. Examples. (1) [8] Let $A = \{a, b, c\}$ be a set with a 3-ary hyperoperation σ as follows:

$\sigma(a, a, a) = a$	$\sigma(b, b, a) = \{a, c\}$	$\sigma(c, a, a) = c$
$\sigma(a, a, b) = b$	$\sigma(b, b, b) = \{b, c\}$	$\sigma(c, a, b) = \{b, c\}$
$\sigma(a, a, c) = c$	$\sigma(b, b, c) = A$	$\sigma(c, a, c) = \{a, b\}$
$\sigma(a, b, a) = b$	$\sigma(b, a, a) = b$	$\sigma(c, b, a) = \{b, c\}$
$\sigma(a, b, b) = \{a, c\}$	$\sigma(b, a, b) = \{a, c\}$	$\sigma(c, b, b) = A$
$\sigma(a, b, c) = \{b, c\}$	$\sigma(b, a, c) = \{b, c\}$	$\sigma(c, b, c) = A$
$\sigma(a, c, a) = c$	$\sigma(b, c, a) = \{b, c\}$	$\sigma(c, c, a) = \{a, b\}$
$\sigma(a, c, b) = \{b, c\}$	$\sigma(b, c, b) = A$	$\sigma(c, c, b) = A$
$\sigma(a, c, c) = \{a, b\}$	$\sigma(b, c, c) = A$	$\sigma(c, c, c) = \{b, c\}$

One can see that $\mathfrak{A} = \{A, \sigma\}$ is a 3-ary hypergroup.

(2) [9] Let $B = \{a, b, c\}$. We define hyperoperation σ as follows

$$\sigma(x, y, z) = \begin{cases} x & x = y = z \\ b & x \neq y \neq z \\ z & x = y, x \neq z, x \neq b \\ \{a, c\} & x = y = b, z \neq b \end{cases}$$

One can see that $\mathfrak{B} = \{B, \sigma\}$ is a 3-ary hypergroup.

2.5. Definition. Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and $\{i, j\} \subseteq I_{\rho(\sigma)}$. We say that σ is $\{i, j\}$-commutative if for each $a_1^{\rho(\sigma)} \in A$ we have

$$\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) = \sigma^{\mathfrak{A}}(a_1^{t-1}, a_s, a_{t+1}^{s-1}, a_t, a_{s+1}^{\rho(\sigma)})$$

where $t = \min\{i, j\}$ and $s = \max\{i, j\}$ We say that σ is commutative if it is $\{i, j\}$-commutative for each $i, j \in I_{\rho(\sigma)}$.

Obviously, the 3-ary hypergroup $\mathfrak{A} = (A, \sigma)$ in Example 2.4 is commutative, too.

2.6. Definition. [9] Let $\mathfrak{A}$ be a Σ-hyperalgebra and $\sigma \in \Sigma$. We say that $\mathfrak{A}$ has a weak neutral element with respect to σ if there exists an element $e \in A$ such that

$$x \in \sigma^{\mathfrak{A}}(\overset{(i-1)}{e}, x, \overset{(\rho(\sigma)-i)}{e})$$

holds for all $x \in A$ and $i \in I_{\rho(\sigma)}$. If for all $x \in A$ and $i \in I_{\rho(\sigma)}$, we have

$$x = \sigma^{\mathfrak{A}}(\overset{(i-1)}{e}, x, \overset{(\rho(\sigma)-i)}{e})$$

then "e" is called a neutral element. The set of weak neutral elements is denoted by $WN_{\mathfrak{A}}$ and the set of neutral elements is denoted by $N_{\mathfrak{A}}$. Clearly, $N_{\mathfrak{A}} \subseteq WN_{\mathfrak{A}}$.

An n-ary semihypergroup $\mathfrak{A}$ is called an n-ary W-hypermonoid if $WN_{\mathfrak{A}} \neq \emptyset$ and it is called an n-ary hypermonoid if $N_{\mathfrak{A}} \neq \emptyset$. Obviously, if $\mathfrak{A}$ is a 2-ary hypermonoid, then $|N_{\mathfrak{A}}| = 1$.

2.7. Definition. Let $\mathfrak{A}$ be a Σ-hyperalgebra and $\sigma \in \Sigma$. We say that $\sigma^{\mathfrak{A}}$ satisfies the weak idempotent property if $x \in \sigma(\overset{\rho(\sigma)}{x})$, for each $x \in A$. Also, we say that $\sigma^{\mathfrak{A}}$ has the idempotent property if $x = \sigma(\overset{\rho(\sigma)}{x})$.

An n-ary commutative semihypergroup $\mathfrak{A} = (A, \sigma)$ is called an n-ary hypersemilattice if "$\circ$" satisfies the weak idempotent property and $\mathfrak{A}$ is called an n-ary semihyperlattice if "$\circ$" satisfies the idempotent property.

2.8. Examples. 1) Let $A = \{a, b, c, d\}$ with the hyperoperation $\circ$ defined by the following table

$\circ$	a	b	c	d
a	$\{a\}$	$\{b\}$	$\{c\}$	$\{d\}$
b	$\{b\}$	$\{a, b\}$	$\{d\}$	$\{c, d\}$
c	$\{c\}$	$\{d\}$	$\{a, c\}$	$\{b, d\}$
d	$\{d\}$	$\{c, d\}$	$\{b, d\}$	A

One can see that $\mathfrak{A} = (A, \circ)$ is a 2-ary hypersemilattice. Also, if we define the ternary hyperoperation $\circ'$ on A by $\circ'(x, y, z) = x \circ y \circ z$ for all $x, y, z \in A$, then one can see easily that $\mathfrak{A}' = (A, \circ')$ is a 3-ary hypersemilattice.

2) Let $B = \{a, b, c, d\}$ with the hyperoperation $\circ$ defined by the following table

$\circ$	a	b	c	d
a	$\{a\}$	$\{b, c, d\}$	$\{c, d\}$	$\{d\}$
b	$\{b, c, d\}$	$\{b\}$	$\{c, d\}$	$\{d\}$
c	$\{c, d\}$	$\{c, d\}$	$\{c\}$	$\{d\}$
d	$\{d\}$	$\{d\}$	$\{d\}$	$\{d\}$

One can see that $\mathfrak{B} = (B, \circ)$ is a 2-ary semihyperlattice. Also, if we define the ternary hyperoperation $\circ'$ on B by $\circ'(x, y, z) = x \circ y \circ z$ for all $x, y, z \in B$, then one can see easily that $\mathfrak{B}' = (B, \circ')$ is a 3-ary semihyperlattice.

2.9. Example. Consider Examples 2.8. One can see that $\mathfrak{A}'$ is a 3-ary W-hypermonoid and $\mathfrak{B}'$ is a 3-ary hypermonoid.

2.10. Examples. i) A commutative hypermonoid $\mathfrak{A} = (A, \circ, e)$ is called canonical hypergroup if

- every element has a unique inverse, which means that for all $x \in A$, there exists a unique $x^{-1} \in A$, such that $e \in x \circ x^{-1}$,
- it is reversible, which means that if $x \in y \circ z$, then $y \in x \circ z^{-1}$ and $z \in y^{-1} \circ x$.

ii) A Krasner hyperring is a hyperalgebraic structure $\mathfrak{R} = (R, +, .)$ which $(R, +, 0)$ is a canonical hypergroup, $(R, .)$ is a semiring and the multiplication, ., is distributive with respect to the hyperoperation $+$.

One can find definitions of hyperlattices [27, 28], hyper-MV algebras [29, 31], hyper-BCK algebras [30] and etc.

3. Homomorphisms

In the universal algebra theory, the concepts of congruence, quotient algebra and homomorphism are closely related. In this section, we give some ideas about homomorphisms of hyperalgebras and we state connection between them.

3.1. Definition. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-algebras. A map $f : A \longrightarrow B$ is a homomorphism (*resp.* dual homomorphism), in symbols $f : \mathfrak{A} \longrightarrow \mathfrak{B}$, if, for all $\sigma \in \Sigma$ and all $a_1^{\rho(\sigma)} \in A$,

$$f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})) \subseteq \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$$

$$(\textit{resp. } f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})) \supseteq \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$$

Also, f is called a good homomorphism if

$$f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})) = \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$$

$Hom(\mathfrak{A}, \mathfrak{B})$ will denote the set of all homomorphisms and $Hom_G(\mathfrak{A}, \mathfrak{B})$ will denote the set of all good homomorphisms from $\mathfrak{A}$ to $\mathfrak{B}$. $f : A \longrightarrow B$ is called an isomorphism if $f \in Hom(\mathfrak{A}, \mathfrak{B})$ and $f^{-1} \in Hom(\mathfrak{B}, \mathfrak{A})$, too.

3.2. Example. Let $M = \{0, a, b, 1\}$. Consider the following tables:

$\oplus$	0	a	b	1
0	$\{0\}$	$\{a\}$	$\{b\}$	$\{1\}$
a	$\{a\}$	$\{b\}$	$\{1\}$	$\{1\}$
b	$\{b\}$	$\{1\}$	$\{1\}$	$\{1\}$
1	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$

*	0	a	b	1
	1	b	a	0

It is easy to see that $\mathfrak{M} = (M, \oplus, ^*, 0)$ is a trivial hyper-MV algebra which is totally ordered as an MV-algebra. Also, let $HS_3 = (S_3, \oplus, ^*, 0)$ is the hyper-MV algebra where $S_3 = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, $x \oplus y := [0, \min(1, x + y)] \cap S_3$ and $x^* = 1 - x$. Now, let $f : \mathfrak{M} \longrightarrow HS_3$ be a map that $f(0) = 0$, $f(a) = \frac{1}{3}$, $f(b) = \frac{2}{3}$ and $f(1) = 1$. It is easy to check that f is a hyper-MV algebra homomorphism. For example $f(a \oplus a) = f(\{b\}) = \{\frac{2}{3}\}$ and $f(a) \oplus f(a) = \frac{1}{3} \oplus_{H3} \frac{1}{3} = \{0, \frac{1}{3}, \frac{2}{3}\}$ so $f(a \oplus a) \subsetneq f(a) \oplus f(a)$. Also, if we define $g : HS_3 \longrightarrow \mathfrak{M}$ such that $g(0) = 0$, $g(\frac{1}{3}) = a$, $g(\frac{2}{3}) = b$ and $g(1) = 1$, then g is a dual homomorphism that is not a homomorphism, since $g(\frac{1}{3} \oplus_{H3} \frac{1}{3}) = \{0, a, b\}$ but $g(\frac{1}{3}) \oplus g(\frac{1}{3}) = \{b\}$.

3.3. Theorem. *A bijective homomorphism of Σ-hyperalgebras is an isomorphism if and only if it is good.*

Proof. Let $\mathfrak{A}, \mathfrak{B}$ be Σ-hyperalgebras, $f : A \longrightarrow B$ is a bijection, $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in A$. First, suppose that f is an isomorphism and $b \in \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$. So $f^{-1}(b) \in f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$ and since $f^{-1} \in Hom(\mathfrak{B}, \mathfrak{A})$ we obtain that $b \in f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}))$ and it implies that $f \in Hom_G(\mathfrak{A}, \mathfrak{B})$. Conversely, suppose that f is a bijective good homomorphism. Let $\sigma \in \Sigma$ and $b_1^{\rho(\sigma)} \in B$. Since, f is onto, there are $a_1^{\rho(\sigma)} \in A$, such that $f(a_i) = b_i$, for each $i \in I_{\rho(\sigma)}$. Now, we have

$$\begin{aligned} f^{-1}(\sigma^{\mathfrak{B}}(b_1^{\rho(\sigma)})) &= f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))) \\ &= f^{-1}(f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}))) \\ &= \sigma^{\mathfrak{A}}(f^{-1}(b_1), \cdots, f^{-1}(b_{\rho(\sigma)}))). \end{aligned}$$

□

If there is a good isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$, then we write $\mathfrak{A} \cong \mathfrak{B}$. Clearly, $\cong$ is an equivalence relation on the set of all Σ-hyperalgebras.

3.4. Definition. Let $\mathfrak{A}, \mathfrak{B}$ be two Σ-hyperalgebras. A map $f : A \longrightarrow B$, is called

- a *very good homomorphism* if it is good and for each $\sigma \in \Sigma$, $a_1^{\rho(\sigma)} \in A$ and $i \in I_{\rho(\sigma)}$, we have
  $f(\sigma_i^{\mathfrak{A}}(a_1^{\rho(\sigma)})) = \sigma_i^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$.
- a *2-homomorphism* if for each $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in A$, we have
  $f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))) = f^{-1}(f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})))$.
- an *almost strong homomorphism* if for each $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in A$, we have
  $f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))) = \sigma^{\mathfrak{A}}(f^{-1}(f(a_1)), \cdots, f^{-1}(f(a_{\rho(\sigma)})))$.

3.5. Theorem. *Let $\mathfrak{A}, \mathfrak{B}$ be two Σ-hyperalgebras. If $f : \mathfrak{A} \longrightarrow \mathfrak{B}$ is a good homomorphism, then f is a 2-homomorphism. Furthermore, if f is a very good homomorphism, then f is an almost strong homomorphism.*

Proof. Let $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in A$. First let $f \in Hom_G(\mathfrak{A}, \mathfrak{B})$. Let $a \in f^{-1}(f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})))$, so $f(a) \in f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}))$ and it implies that there is $a' \in \sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})$ such that $f(a) = f(a')$. Since f is good we obtain that $a \in f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$. Conversely, assume that $a \in f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$, so $f(a) \in \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$ and it shows that f is a 2-homomorphism.

Now, suppose that f is a very good homomorphism. Since $f \in Hom(\mathfrak{A}, \mathfrak{B})$ we obtain that $f(\sigma^{\mathfrak{A}}(f^{-1}(f(a_1)), \cdots, f^{-1}(f(a_{\rho(\sigma)})))) \subseteq \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$. It shows that

$$\sigma^{\mathfrak{A}}(f^{-1}(f(a_1)), \cdots, f^{-1}(f(a_{\rho(\sigma)}))) \subseteq f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))).$$

Conversely, assume that $a \in f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$. Hence, $f(a) \in \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$, so for each $i \in I_{\rho(\sigma)}$ we have

$$f(a_i) \in \sigma_i^{\mathfrak{B}}(f(a_1), \cdots, f(a_{i-1}), f(a), f(a_{i+1}), \cdots, f(a_{\rho(\sigma)})).$$

Since, f is very good we obtain that

$$f(a_i) \in f(\sigma_i^{\mathfrak{A}}(a_1^{i-1}, a, a_{i+1}^{\rho(\sigma)})).$$

Hence, there is $a'_i \in \sigma_i^{\mathfrak{A}}(a_1^{i-1}, a, a_{i+1}^{\rho(\sigma)})$ such that $f(a_i) = f(a'_i)$. It implies that $a \in \sigma^{\mathfrak{A}}(a_1^{i-1}, f^{-1}(f(a_i)), a_{i+1}^{\rho(\sigma)})$. Since i is arbitrary, continuing the above method shows that f is an almost strong homomorphism. □

3.6. Definition. Let $\mathfrak{A}, \mathfrak{B}$ be two Σ-hyperalgebras, $f \in Hom(\mathfrak{A}, \mathfrak{B})$, $\sigma \in \Sigma$ and $i \in I_{\rho(\sigma)}$. f is called *strong on the i-th component* respect to σ if for all $a_1^{\rho(\sigma)} \in A$, $f(a) \in \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$ implies that there exists $a'_i \in A$ such that $f(a_i) = f(a'_i)$ and $a \in \sigma^{\mathfrak{A}}(a_1^{i-1}, a'_i, a_{i+1}^{\rho(\sigma)})$. f is called *strong respect to* σ if for all $i \in I_{\rho(\sigma)}$, f is strong on the i-th component. f is called *strong* if for each $\sigma \in \Sigma$, f is strong respect to σ.

3.7. Theorem. *Any strong homomorphism is almost strong.*

Proof. Let $f \in Hom(\mathfrak{A}, \mathfrak{B})$ be a strong homomorphism, $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in A$. Since $f \in Hom(\mathfrak{A}, \mathfrak{B})$, similar to Theorem 3.5, we can obtain that $\sigma^{\mathfrak{A}}(f^{-1}(f(a_1)), \cdots, f^{-1}(f(a_{\rho(\sigma)}))) \subseteq f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$. Now, assume that $a \in f^{-1}(\sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})))$. So $f(a) \in \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)}))$. Since f is strong respect to σ, for each $i \in I_{\rho(\sigma)}$, there exists $a'_i \in A$ such that $f(a_i) = f(a'_i)$ and $a \in \sigma^{\mathfrak{A}}(a_1^{i-1}, a'_i, a_{i+1}^{\rho(\sigma)})$, whence $a \in \sigma^{\mathfrak{A}}(a_1^{i-1}, f^{-1}(f(a_i)), a_{i+1}^{\rho(\sigma)}) \subseteq \sigma^{\mathfrak{A}}(f^{-1}(f(a_1)), \cdots, f^{-1}(f(a_{\rho(\sigma)})))$. □

4. Subhyperalgebras

There are several important methods of constructing new hyperalgebras from given ones. Three of the most fundamental are the formation of subhyperalgebras, homomorphic images, and direct products. In this section, subhyperalgebras will occupy us.

4.1. Definition. Let $\mathfrak{A}$ be a Σ-hyperalgebra and $B \subseteq A$. Then B is called a subhyperuniverse of $\mathfrak{A}$, if for all $\sigma \in \Sigma$ and $b_1^{\rho(\sigma)} \in B$ we have $\sigma^{\mathfrak{A}}(b_1^{\rho(\sigma)}) \subseteq B$. The set of all subhyperuniverses of $\mathfrak{A}$ will be denoted by $Sub(\mathfrak{A})$.

Note that this implies $\sigma^{\mathfrak{A}} \subseteq B$ for every $\sigma \in \Sigma_0$, and that the empty set is a subhyperuniverse of $\mathfrak{A}$ if and only if $\Sigma_0 = \emptyset$, i.e., $\mathfrak{A}$ has no distinguished constants.

4.2. Example. Consider hypersemilattice $\mathfrak{A}$ in Examples 2.8. It is easy to see that $B = \{a, b\}$ is a subhyperuniverse of $\mathfrak{A}$.

4.3. Theorem. *$(A, Sub(\mathfrak{A}))$ is an algebraic closed set system for every Σ-hyperalgebra $\mathfrak{A}$.*

Proof. Let $\mathcal{K} \subseteq Sub(\mathfrak{A})$. Let $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in \cap\mathcal{K}$. Then for every $K \in \mathcal{K}$, we have $a_1^{\rho(\sigma)} \in K$ and hence $\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) \subseteq K$. So $\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) \subseteq \cap\mathcal{K}$. Therefore, $\cap\mathcal{K} \in Sub(\mathfrak{A})$.

Now, assume that $\mathcal{K}$ is a directed subset of $Sub(\mathfrak{A})$. Let $a_1^{\rho(\sigma)} \in \cup\mathcal{K}$. Since there is only a finite number of a_i and $\mathcal{K}$ is directed, they are all contained in a single $K \in \mathcal{K}$. Hence $\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) \subseteq K$. Therefore, $\cup\mathcal{K} \in Sub(\mathfrak{A})$. □

The closure operator associated with the closed set system $(A, Sub(\mathfrak{A}))$ is denoted by $Sg^{\mathfrak{A}}$. Thus $Sg^{\mathfrak{A}} : \mathcal{P}(A) \longrightarrow Sub(\mathfrak{A})$ and $Sg^{\mathfrak{A}}(X) = \cap\{K \in Sub(\mathfrak{A}) : X \subseteq K\}$; this is called the subhyperuniverse generated by X. Hyperalgebra $\mathfrak{A}$ is called finitely generated if there exists a finite subset X of A such that $Sg^{\mathfrak{A}}(X) = A$.

B is a maximal proper subhyperuniverse of $\mathfrak{A}$ if $B \neq A$ and there does not exist a $C \in Sub(\mathfrak{A})$ such that$B \subsetneq C \subsetneq A$.

4.4. Theorem. *Let $\mathfrak{A}$ be a finitely generated Σ-algebra. Then every proper subhyperuniverse of $\mathfrak{A}$ is included in a maximal proper one.*

Proof. Let $A = Sg^{\mathfrak{A}}(X)$, for $X \subseteq_w A$. Assume that B is a proper subhyperuniverse of $\mathfrak{A}$. Let $\mathcal{K} = \{K \in Sub(\mathfrak{A}) : B \subseteq K \subsetneq A\}$. Since $\mathcal{K}$ contains B, it is nonempty. Suppose that $\mathcal{C} \subseteq \mathcal{K}$ be a chain. Clearly, $\mathcal{C}$ is directed so by Theorem 4.3, $\cup\mathcal{C} \in Sub(\mathfrak{A})$. Because X is finite and $\mathcal{K}$ is directed, one can see that $\cup\mathcal{C}$ is a proper subhyperuniverse of $\mathfrak{A}$. hence, by Zorn's lemma $\mathcal{K}$ has a maximal element. □

4.5. Theorem. *(Principle of Structural Induction) Let $\mathfrak{A}$ be a Σ-hyperalgebra generated by X. To prove that a property $\mathcal{P}$ holds for each element of A, it suffices to show that*

induction basis. $\mathcal{P}$ holds for each element of X.

induction step. If $\sigma \in \Sigma$ and $\mathcal{P}$ holds for each of elements $a_1^{\rho(\sigma)} \in A$, then $\mathcal{P}$ holds for each elements of $\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})$.

Proof. Let $P = \{x \in A : \mathcal{P} \text{ holds for } x\}$. $X \subseteq P$ and P is closed under the hyperoperations of $\mathfrak{A}$. Hence, $P \in Sub(\mathfrak{A})$ and it implies that $A = Sg^{\mathfrak{A}}(X) \subseteq P$. □

Let $\mathfrak{A}$ be a Σ-hyperalgebra and $X \subseteq A$. We define

$$E(X) = X \cup \{t \in \sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) : \forall \sigma \in \Sigma,\ a_1^{\rho(\sigma)} \in X\}.$$

Now, we define $E^n(X)$ for $n \geq 0$ by

$$E^0(X) = X,$$
$$E^n(X) = E(E^{n-1}(X)).$$

4.6. Theorem. *Let $\mathfrak{A}$ be a Σ-hyperalgebra and $X \subseteq A$. Then*

$$Sg^{\mathfrak{A}}(X) = \cup_{i=0}^{\infty} E^i(X).$$

Proof. Let $P = \{x \in Sg^{\mathfrak{A}}(X) : x \in I = \cup_{i=0}^{\infty} E^i(X)\}$. Clearly, $X \subseteq P$. Now, let $\sigma \in \Sigma$ and $x_1, \cdots, x_{\rho(\sigma)} \in P$. One can see that $E^i(X) \subseteq E^{i+1}(X)$ for each $i \in w$, so it implies that there is a $n \in w$ such that $x_1, \cdots, x_{\rho(\sigma)} \in E^n(X)$. Thus $\sigma^{\mathfrak{A}}(x_1, \cdots, x_{\rho(\sigma)}) \subseteq E^{n+1}(X) \subseteq I$. Therefore by Theorem 4.5, we get the result. □

4.7. Definition. Let $\mathfrak{A}$ and $\mathfrak{B}$ be two Σ-hyperalgebras. Then $\mathfrak{B}$ is a subhyperalgebra of $\mathfrak{A}$ if $B \subseteq A$ and every fundamental operation of $\mathfrak{B}$ is the restriction of the corresponding operation of $\mathfrak{A}$, i.e., for each function symbol $\sigma \in \Sigma$, $\sigma^{\mathfrak{B}}$ is $\sigma^{\mathfrak{A}}$ restricted to B; we write simply $\mathfrak{B} \leq \mathfrak{A}$. If $\mathfrak{B} \leq \mathfrak{A}$, then $B \in Sub(\mathfrak{A})$. Conversely, if $B \in Sub(\mathfrak{A})$ and $B \neq \emptyset$, then there is a unique $\mathfrak{B} \leq \mathfrak{A}$ such that B is the universe of $\mathfrak{B}$.

4.8. Theorem. *Let $\mathfrak{A}, \mathfrak{B}$ be Σ-hyperalgebras, $f \in Hom_G(\mathfrak{A}, \mathfrak{B})$ and $h \in Hom(\mathfrak{A}, \mathfrak{B})$. Then*

(i) *For each $K \in Sub(\mathfrak{A})$, $f(K) \in Sub(\mathfrak{B})$.*
(ii) *For each $L \in Sub(\mathfrak{B})$, $h^{-1}(L) \in Sub(\mathfrak{A})$, if $h^{-1}(L) \neq \emptyset$.*
(iii) *For each $X \subseteq A$, $f(Sg^{\mathfrak{A}}(X)) \in Sg^{\mathfrak{A}}(f(X))$.*

Proof. (i) Let $\sigma \in \Sigma$ and $b_1^{\rho(\sigma)} \in f(K)$. Choose $a_1^{\rho(\sigma)} \in K$ such that $f(a_1) = b_1, \cdots, f(a_{\rho(\sigma)}) = b_{\rho(\sigma)}$. Then $\sigma^{\mathfrak{B}}(b_1^{\rho(\sigma)}) = \sigma^{\mathfrak{B}}(f(a_1), \cdots, f(a_{\rho(\sigma)})) = f(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})) \subseteq f(K)$.

(ii) Assume that $h^{-1}(L) \neq \emptyset$. Let $\sigma \in \Sigma$ and $a_1^{\rho(\sigma)} \in h^{-1}(L)$. So $f(a_1), \cdots, f(a_{\rho(\sigma)}) \in L$. Then $h(\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)})) \subseteq \sigma^{\mathfrak{B}}(h(a_1), \cdots, h(a_{\rho(\sigma)})) \subseteq L$. Hence, $\sigma^{\mathfrak{A}}(a_1^{\rho(\sigma)}) \subseteq h^{-1}(L)$.

(iii) $f(X) \subseteq f(Sg^{\mathfrak{A}}(X)) \in Sub(\mathfrak{B})$, by part (i). Therefore, $Sub^{\mathfrak{B}}(f(X)) \subseteq f(Sg^{\mathfrak{A}}(X))$. For the reverse inclusion, $X \subseteq f^{-1}(f(X)) \subseteq f^{-1}(Sg^{\mathfrak{A}}(f(X)) \in Sub(\mathfrak{A})$, by part (ii). Hence, $Sg^{\mathfrak{A}}(X) \subseteq f^{-1}(Sg^{\mathfrak{A}}(f(X)))$.

□

Let $Alg_H(\Sigma)$ be the class of all Σ-hyperalgebras. $\leq$ is a partial ordering of $Alg_H(\Sigma)$. Clearly, it is reflexive and antisymmetric. If $\mathfrak{C} \leq \mathfrak{B}$ and $\mathfrak{B} \leq \mathfrak{A}$, then $C \subseteq A$ and for each σ and $c_1^{\rho(\sigma)} \in C$, we have $\sigma^{\mathfrak{C}}(c_1^{\rho(\sigma)}) = \sigma^{\mathfrak{B}}(c_1^{\rho(\sigma)}) = \sigma^{\mathfrak{A}}(c_1^{\rho(\sigma)})$. Hence, $\leq$ is transitive.

For any class K of $Alg_H(\Sigma)$ we define

$$\mathsf{S}(\mathsf{K}) = \{\mathfrak{B} \in \mathsf{Alg_H}(\Sigma) : \exists\, \mathfrak{A} \in \mathsf{K}(\mathfrak{B} \leq \mathfrak{A})\}.$$

4.9. Theorem. *S is an algebraic closure operator on $Alg_H(\Sigma)$.*

Proof. Clearly, $\mathsf{K} \subseteq \mathsf{S}(\mathsf{K})$ by the reflexivity of $\leq$ and by transitivity of $\leq$, we obtain that $\mathsf{SS}(\mathsf{K}) = (\mathsf{K})$. Also, $\mathsf{K} \subseteq \mathsf{L}$ implies $\mathsf{S}(\mathsf{K}) \subseteq \mathsf{S}(\mathsf{L})$. Furthermore, $\mathsf{S}(\mathsf{K}) = \cup\{\mathsf{S}(\mathfrak{A}) : \mathfrak{A} \in \mathsf{S}(\mathsf{K})\}$ and it implies that S is algebraic closed. □

Define the binary relation $\preccurlyeq$ on $Alg_H(\Sigma)$ by $\mathfrak{B} \preccurlyeq \mathfrak{A}$ if there is an onto good homomorphism $f : \mathfrak{A} \longrightarrow \mathfrak{B}$. Clearly, $\preccurlyeq$ is reflexive and transitive.

For any class K of $Alg_H(\Sigma)$ we define

$$\mathsf{H}(\mathsf{K}) = \{\mathfrak{B} \in \mathsf{Alg_H}(\Sigma) : \exists\, \mathfrak{A} \in \mathsf{K}(\mathfrak{B} \preccurlyeq \mathfrak{A})\},$$

$$\mathsf{I}(\mathsf{K}) = \{\mathfrak{B} \in \mathsf{Alg_H}(\Sigma) : \exists\, \mathfrak{A} \in \mathsf{K}(\mathfrak{B} \cong \mathfrak{A})\},$$

the classes respectively of homomorphic and isomorphic images of algebras of K. Similar to Theorem 4.9, we can show that H and I are algebraic closure operators on $Alg_H(\Sigma)$.

4.10. Theorem. *For any class K of Σ-hyperalgebras, we have*

(i) $\mathsf{SH}(\mathsf{K}) \subseteq \mathsf{HS}(\mathsf{K})$,
(ii) HS *is an algebraic closure operator on* $Alg_H(\Sigma)$.

Proof. (i) Let $\mathfrak{C} \in \mathsf{SH}(\mathsf{K})$ so for some $\mathfrak{A} \in \mathsf{K}$ and onto good homomorphism $f : \mathfrak{A} \longrightarrow \mathfrak{B}$ we have $\mathfrak{C} \leq \mathfrak{B}$. By Theorem 4.8, we have $f^{-1}(\mathfrak{C}) \leq \mathfrak{A}$. Since f is onto we have $\mathfrak{C} \in \mathsf{HS}(\mathsf{K})$.

(ii) Obviously, HS is extensive. By part (i), we have $\mathsf{HSHS}(\mathsf{K}) \subseteq \mathsf{HHSS}(\mathsf{K}) = \mathsf{HS}(\mathsf{K})$ so we get that HS is idempotent. Finally, monotonicity of S and H imply that HS is monotonic. Also, $\mathfrak{B} \in \mathsf{HS}(\mathsf{K})$ if and only if there is a $\mathfrak{A} \in \mathsf{K}$ such that $\mathfrak{B} \in \mathsf{HS}(\mathfrak{A})$ so we have $\mathsf{HS}(\mathsf{K}) \subseteq \cup\{\mathsf{HS}(\mathfrak{A}) : \mathfrak{A} \in \mathsf{K}\}$ and it implies that HS is algebraic.

□

5. Some type of subhyperalgebras

There are several kinds of subhyperalgebras. In what follows, we introduce closed, invertible, ultraclosed and conjugable subhyperalgebras and some connections among them. Let us present now the definition of these types of subhyperalgebras.

5.1. Definition. Let $\mathfrak{A}$ be a Σ-hyperalgebra, $B \in Sub(\mathfrak{A})$, $\sigma \in \Sigma$ and $1 \leq i \leq \rho(\sigma)$. We say that B is

(1) i-closed with respect to σ if for each $x \in A$ and $b_1^{\rho(\sigma)} \in B$, from $x \in \sigma_i^A(b_1^{\rho(\sigma)})$ it follows that $x \in B$.

(2) i-invertible with respect to σ if for each $x, y \in A$, from $x \in \sigma^{\mathfrak{A}}(\overset{(i-1)}{B}, y, \overset{(\rho(\sigma)-i)}{B})$ it follows that $y \in \sigma^{\mathfrak{A}}(\overset{(i-1)}{B}, x, \overset{(\rho(\sigma)-i)}{B})$.

(3) ultraclosed on the right (on the left) with respect to σ if for each $x \in A$ we have $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cap \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c) = \emptyset$ $(\sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, x) \cap \sigma^{\mathfrak{A}}(B^c, \overset{(\rho(\sigma)-2)}{B}, x) = \emptyset)$.

(4) i-conjugable with respect to σ if it is i-closed with respect to σ and for each $x \in A$ there exist $x_1^{\rho(\sigma)-1} \in A$ such that $\sigma^{\mathfrak{A}}(x_1^{i-1}, x, x_{i+1}^{\rho(\sigma)-1}) \subseteq B$.

We say that B is closed (*resp.* invertible, conjugable) with respect to σ, if it is i-closed (*resp.* invertible, conjugable) with respect to σ, for each $1 \leq i \leq \rho(\sigma)$. Also, we say that B is closed (*resp.* invertible, conjugable) if it is closed (*resp.* invertible, conjugable) for each $\sigma \in \Sigma$. Sets of closed, invertible and conjugable subhyperalgebra of hyperalgebra $\mathfrak{A}$ are denoted by $Sub_{cl}(\mathfrak{A})$, $Sub_{in}(\mathfrak{A})$ and $Sub_{co}(\mathfrak{A})$, respectively.

5.2. Examples. (1) Consider 3-hypergroup $\mathfrak{A}$ in Examples 2.4. Let $I = \{a\}$. Clearly, I is a subhyperalgebra of $\mathfrak{A}$ which it is closed and invertible but it is not ultraclosed, since $\sigma(b, a, a) \cap \sigma(b, a, c) = b$, and it is not conjugable.

(2) Consider commutative 3-ary hypergroup $\mathfrak{B}$ in Examples 2.4. Let $J = \{b\}$. One can see that J is a subhyperalgebra of $\mathfrak{B}$ which it is closed, invertible, ultraclosed and conjugable.

The following theorems present some connections among the above types of subhyperalgebras.

5.3. Theorem. *Let* $\mathfrak{A}$ *be a* Σ*-hyperalgebra. Then* $(Sub_{cl}(\mathfrak{A}), \subseteq)$ *and* $(Sub_{in}(\mathfrak{A}), \subseteq)$ *are algebraic closed set systems. Furthermore,* $Sub_{in}(\mathfrak{A}) \subseteq Sub_{cl}(\mathfrak{A})$.

Proof. Obviously, $(Sub_{cl}(\mathfrak{A}), \subseteq)$ and $(Sub_{in}(\mathfrak{A}), \subseteq)$ are closed set systems. Also, similar to Theorem 4.3, we can conclude that $(Sub_{cl}(\mathfrak{A}), \subseteq)$ and $(Sub_{in}(\mathfrak{A}), \subseteq)$ are algebraic.

Now, suppose that $\sigma \in \Sigma$ and $1 \leq i \leq \rho(\sigma)$, $b_1^{\rho(\sigma)} \in B$ and $x \in A$ such that $x \in \sigma_i^{\mathfrak{A}}(b_1^{\rho(\sigma)})$. By hypothesis $x \in \sigma^{\mathfrak{A}}(b_1^{\rho(\sigma)}) \subseteq B$ and it holds the result. □

5.4. Lemma. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $B \in Sub(\mathfrak{A})$ and $\sigma \in \Sigma$ such that it is $(1, \rho(\sigma))$-associative. Then B is $1(\rho(\sigma))$-invertible if and only if $\{\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})\}_{x\in A}$ ($\{\sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, x)\}_{x\in A}$) is a disjoint family of subsets of A.*

Proof. Suppose that $\sigma \in \Sigma$, B is 1-invertible with respect to σ and $z \in \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})) \cap \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}))$, for some $x, y \in A$. Then we have

$$\begin{aligned} \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B}) &\subseteq \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}), \overset{(\rho(\sigma)-1)}{B}) \\ &= \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma))}{B})) \\ &\subseteq \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \end{aligned}$$

By hypothesis $x \in \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$ and similarly we can conclude that $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \subseteq \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$. So $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$. Also, we can get $\sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$ and it shows that $\{\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})\}_{x\in A}$ is a disjoint family of subsets of A.

Conversely, suppose that $x \in \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B})$. Then $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \subseteq \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$ whence $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B})$. Thus we have $x \in \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})$ hence we obtain that $y \in \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})$. Similarly, we can show that B is $\rho(\sigma)$-invertible if and only if $\{\sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, x)\}_{x\in A}$ is a disjoint family of subsets of A. □

5.5. Lemma. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and $B \in Sub(\mathfrak{A})$ such that A, B satisfies the reproduction axiom with respect to σ and B be $\rho(\sigma)(1)$-closed. Then $\sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c) = B^c$, ($\sigma^{\mathfrak{A}}(B^c, \overset{(\rho(\sigma)-1)}{B}) = B^c$).*

Proof. We have

$$\begin{aligned} A &= \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, A) \\ &= \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma))}{B}) \cup \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c) \\ &= B \cup \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c). \end{aligned}$$

On the other hand, assume that $b \in B \cap \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c)$. So there exist $c \in B^c$ and $b_1^{\rho(\sigma)-1} \in B$ such that $b \in \sigma^{\mathfrak{A}}(b_1^{\rho(\sigma)-1}, c)$. Now, since B is $\rho(\sigma)$-closed we obtain that $c \in B$. It is a contradiction. □

5.6. Theorem. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and $B \in Sub(\mathfrak{A})$ such that A, B satisfies the reproduction axiom with respect to σ and B be conjugable. Then B is ultraclosed on the right (on the left).*

Proof. Suppose that $x \in A$. Denote $C = \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cap \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c)$. Since B is conjugable it follows that B is closed and there exist $x_1^{\rho(\sigma)-1} \in A$ such that

$\sigma^{\mathfrak{A}}(x_1^{\rho(\sigma)-1}, x) \subseteq B$. By Lemma 5.5, we obtain that

$$\begin{aligned}
\sigma^{\mathfrak{A}}(x_1^{\rho(\sigma)-1}, C) &\subseteq \sigma^{\mathfrak{A}}(x_1^{\rho(\sigma)-1}, \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})) \cap \sigma^{\mathfrak{A}}(x_1^{\rho(\sigma)-1}, \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c)) \\
&\subseteq B \cap \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x_1^{\rho(\sigma)-1}, x), \overset{(\rho(\sigma)-2)}{B}, B^c) \\
&\subseteq B \cap \sigma^{\mathfrak{A}}(B, \overset{(\rho(\sigma)-2)}{B}, B^c) \\
&= B \cap B^c = \emptyset.
\end{aligned}$$

Hence $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cap \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c) = \emptyset$, which means that B is ultraclosed on the right. Similarly, we can show that B is ultraclosed on the left. □

5.7. Lemma. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ and B be an ultraclosed subhyperalgebra on the right (on the left) which satisfies the reproduction axiom with respect to σ. Then B is $\rho(\sigma)(1)$-closed with respect to σ.*

Proof. Let $x \in A \setminus B$ and $b \in B$. Suppose that $b \in \sigma^{\mathfrak{A}}(b_1^{\rho(\sigma)-1}, x) \subseteq \sigma^{\mathfrak{A}}(b_1, \overset{(\rho(\sigma)-2)}{B}, B^c)$. Since B is ultraclosed on the right we have $b \notin \sigma^{\mathfrak{A}}(b_1, \overset{(\rho(\sigma)-1)}{B})$. By the reproduction axiom we conclude that $b \notin B$. It is a contradiction. □

5.8. Theorem. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ such that it is $(1, \rho(\sigma))$-associative and B be an ultraclosed subhyperalgebra on the right (on the left) which satisfies the reproduction axiom with respect to σ. Then B is $1(\rho(\sigma))$-invertible with respect to σ.*

Proof. Suppose that B is ultraclosed on the right. Let $y \in \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})$, for some $x, y \in A$. By associativity we have

$$\sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}) \subseteq \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}), \overset{(\rho(\sigma)-1)}{B}) \subseteq \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}).$$

On the other hand, By Lemma 5.7 we get that B is $\rho(\sigma)$-closed and then we obtain Lemma 5.5 and by this lemma we have

$$\begin{aligned}
\sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-2)}{B}, B^c) &\subseteq \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}), \overset{(\rho(\sigma)-2)}{B}, B^c) \\
&= \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c)) \\
&= \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c)
\end{aligned}$$

Thus, by the reproduction axiom we have

$$A = \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cup \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c) = \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}) \cup \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-2)}{B}, B^c)$$

Since, B is ultraclosed on the right we get that $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B})$. It shows that $\{\sigma^{\mathfrak{A}}(x, B^{\rho(\sigma)-1})\}_{x \in A}$ is a disjoint family of subsets of A so by Lemma 5.4 we conclude that B is 1-invertible with respect to σ. Similarly, we can show that B is $\rho(\sigma)$-invertible, if it is ultraclosed on the left. □

Let $\mathfrak{A}$ be a Σ-hyperalgebra and $\sigma \in \Sigma$. We denote

$$I_\sigma = \{e_1^{\rho(\sigma)} \in A \mid \exists x \in A, \text{ such that } x \in \sigma^{\mathfrak{A}}(e_1^{i_1}, x, e_i^{\rho(\sigma)-1}), \text{ for some } 1 \leq i \leq \rho(\sigma)\}.$$

5.9. Lemma. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ such that it is $(1, \rho(\sigma))$-associative and B be $\rho(\sigma)(1)$-closed with respect to σ which satisfies the reproduction axiom with respect to σ such that $I_\sigma \subseteq B$. Then B is $1(\rho(\sigma))$-invertible.*

Proof. Let $y \in \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B})$, for some $x, y \in A$. Suppose that $x \notin \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B})$. Since, $x \in A = \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-1)}{B}) \cup \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-2)}{B}, B^c)$, we get $x \in \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-2)}{B}, B^c)$. Now, by Lemma 5.5, we have

$$\begin{aligned} x \in \sigma^{\mathfrak{A}}(y, \overset{(\rho(\sigma)-2)}{B}, B^c) &\subseteq \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}), \overset{(\rho(\sigma)-2)}{B}, B^c) \\ &= \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, \sigma^{\mathfrak{A}}(\overset{(\rho(\sigma)-1)}{B}, B^c)) \\ &= \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c). \end{aligned}$$

So there exist $b_1^{\rho(\sigma)-2} \in B$ and $c \in B^c$ such that $x \in \sigma^{\mathfrak{A}}(x, b_1^{\rho(\sigma)-2}, c)$. Hence $c \in I_\sigma \cap B^c$ and it is a contradiction. □

5.10. Theorem. *Let $\mathfrak{A}$ be a Σ-hyperalgebra, $\sigma \in \Sigma$ such that it is $(1, \rho(\sigma))$-associative and $B \in Sub(\mathfrak{A})$ which satisfies the reproduction axiom with respect to σ. If B is 1-closed and $\rho(\sigma)$-close with respect to σ and $I_\sigma \subseteq B$ then B is ultraclosed.*

Proof. Let B is $\rho(\sigma)$-close with respect to σ and $I_\sigma \subseteq B$. Suppose that $z \in \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cap \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c)$. By Lemma 5.9, associativity and Lemma 5.5, we obtain that

$$x \in \sigma^{\mathfrak{A}}(z, \overset{(\rho(\sigma)-1)}{B}) \subseteq \sigma^{\mathfrak{A}}(\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c), \overset{(\rho(\sigma)-1)}{B}) = \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c).$$

Thus, there exist $b_1^{\rho(\sigma)-2} \in B$ and $c \in B^c$ such that $x \in \sigma^{\mathfrak{A}}(x, b_1^{\rho(\sigma)-2}, c)$ and it implies that $c \in I_\sigma \cap B^c$. It shows that $\sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-1)}{B}) \cap \sigma^{\mathfrak{A}}(x, \overset{(\rho(\sigma)-2)}{B}, B^c) = \emptyset$. Hence B is ultraclosed on the right. □

6. Conclusion

The above discussion shows that we can extend some notions of hypergroup theory to a Σ-hyperalgebra for an arbitrary signature Σ. This paper provides suitable tools for doing more research in the area of hyperstructures, such as on homomorphisms and subhyperalgebras. Also, by consideration the operator P one can research on a variety of an arbitrary Σ-hyperalgebras. **Acknowledgements**

Authors are extremely grateful to the referees for giving them many valuable comments and helpful suggestions which helped to improve the presentation of this paper.

References

[1] G. Birkhoff, *Lattice Theory*, 3rd edition. Colloq. Publ., vol. 25, Amer. Math. Soc., Providence, 1967.
[2] P. Corsini, *Sur les homomorphismes d hypergroupes*, Rend. Sem. Univ., 52, Padova, Italy, (1974) 117-140.
[3] P. Corsini, *Prolegomena of Hypergroup Theory*, 2nd edn. (Aviani editor, 1993).
[4] P. Corsini, *Hypergraphs and Hypergroups*, Algebra Universalis, 35 (1996) 548-555.
[5] P. Corsini, V. Leoreanu, *Hypergroups and binary relations*, Algebra Universalis, 43 (2000) 321-330.
[6] P. Corsini, V. Leoreanu, *Applications of hyperstructure theory*, Kluwer, Dordrecht 2003.
[7] B. Davvaz, V. Leoreanu, *Hyperring Theory and Applications*, International Academic Press, 2007.
[8] B. Davvaz, T. Vougiouklis, *n-ary hypergroups*, Iran. J. Sci. Technol. 30(A2) (2006) 165-174.
[9] B. Davvaz. W. A. Dudek. S. Mirvakili, *Neutral elements, fundamental relations and n-ary hypergroups*, International Journal of Algebra and Computation, 19(4) (2009) 567-583.
[10] K. Denecke, S. L. Wismath, *Universal algebra and coalgebra*, World Scientific, 2009.
[11] M. Dresher, O. Ore, *Theory of multigroups*, Amer. J. Math., 60 (1938) 705-733.

[12] G. Grätzer, *A representation theorem for multi-algebras*, Arch. Math., 3 (1962) 452-456.
[13] G. Grätzer, *Universal algebra. Second edition*, Springer-Verlag 1979.
[14] H. Höft, P. E. Howard, *Representing multi-algebras by algebras, the axiom of choice, and the axiom of dependent choice*, Algebra Universalis, 13 (1981) 69-77.
[15] G. E. Hansoul, *A subdirect decomposition theorem for multialgebras*, Algebra universalis, 16 (1983) 275-281.
[16] J. Jantosciak, *Homomorphisms, equivalences and reductions in hypergroups*, Riv. Mat. Pura Appl., 9 (1991) 23-47.
[17] M. Koskas, *Equivalences et homomorphismes des demi-hypergroupes*, C. R. Acad. Sci. Paris, t. 256, (1963).
[18] M. Krasner, *Hypergroups extramoduliformeset moduli forms*, C. R. Acad. Sci.(paris) 219 (1944) 473-476.
[19] J. Kuntzmann, *Homomorphic rntre systèmrs multiformes*, C. R. Acad. Sci. Paris, 205, (1937).
[20] V. Leoreanu, *New results on the hypergroups homomorphisms*, J. Inform. Optim. Sci., 20(2) (1999) 287-298.
[21] V. Leoreanu-Fotea, B. Davvaz, *n-hypergroups and binary relations*, European J. Combinatorics, 29 (2008) 1207-1218.
[22] F. Marty, Sur une generalisation de la notion de groupe, in: some congres des Mathematiciens Scandinaves, Stockholm, (1934) 45-49.
[23] C. Pelea, *On the fundumental relation of a multialgebra*, Ital. J. Pure Appl. Mat., 10 (2001) 327-342.
[24] C. Pelea, *Identities and multialgebras*, Ital. J. Pure Appl. Mat., 15 (2004) 83-92.
[25] C. Pelea, I. Purdea, *Multialgebras, universal algebras and identities*, J. Aust. Math. Soc., 81 (2006) 121-139.
[26] H. E. Pickett, *Homomorphisms and subalgebras of multialgebras*. Pacific J. Math., 21 (1967) 327-342.
[27] S. Rasouli and B. Davvaz, *Lattices derived from hyperlattices*, Communications in Algebra, 38(8) (2010) 2720-2737.
[28] S. Rasouli, B. Davvaz, *Construction and spectral topology on Hyperlattices*, Mediterr. J. Math., 7 (2010) 249-262.
[29] S. Rasouli, D. Heidari, B. Davvaz, *η-Relations and Transitivity Conditions of η on Hyper-MV algebras*, J. of Mult-Valued Logics & Soft Computing, 15, (2009) 517-524.
[30] S. Rasouli, D. Heidari, B. Davvaz, *β-Relations and Transitivity Conditions of β on Hyper-BCK algebras*, Hacettepe Journal of Mathematics and Statistics, 39 (4) (2010), 461-469.
[31] S. Rasouli, B. Davvaz, *Homomorphism, ideals and binary relations on hyper-MV algebras*, J. of Mult-Valued Logics & Soft Computing, 17, (2011) 47-68.
[32] Y. Sureau, *Contribution a la thorie des hypergroupes et hypergroupes operant transivement sur un ensemble*, Doctoral Thesis, (1980).
[33] T. Vougiouklis, *Hyperstructures and their representations*, Hadronic Press, Inc., 115, Palm Harber, USA, 1994.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 985 – 991

A note on the endomorphism ring of finitely presented modules of the projective dimension ≤ 1

Serap Şahinkaya* , Arda Kör† and M. Tamer Koşan‡

Abstract

In this paper, we study the behavior of endomorphism rings of a cyclic, finitely presented module of projective dimension ≤ 1. This class of modules extends to arbitrary rings the class of couniformly presented modules over local rings.

2000 AMS Classification: 16D50.

Keywords: Couniformly presented module , Semilocal ring, Monogeny class, Epigeny class.

Received 25 : 11 : 2011 : Accepted 17 : 07 : 2013 Doi : 10.15672/HJMS.2014437518

1. Introduction

Throughout this paper, all rings will be associative with identity and modules will be unital right modules. For any ring R, the Jacobson radical of R will be denoted by $J(R)$

Recall that M_R is *couniform* if it has dual Goldie dimension one (if and only if it is non-zero and the sum of any two proper submodules of M_R is a proper submodule of M_R). It is well know that a projective right module P_R is couniform if and only if $\mathrm{E}nd(P_R)$ is a local ring, if and only if there exists an idempotent $e \in R$ with $P_R \cong eR$ and eRe a local ring, if and only if is a finitely generated module with a unique maximal submodule.

In [7], Facchini and Girardi introduced and studied the notion of couniformly presented modules. A module M_R is called *couniformly presented* if it is non-zero and there exists an exact sequence

$$0 \to C_R \xrightarrow{\iota} P_R \to M_R \to 0$$

*Department of Mathematics, Gebze Institute of Technology, Çayirova Campus, 41400, Gebze- Kocaeli, Turkey
Email: ssahinkaya@gyte.edu.tr
† Email:akor@gyte.edu.tr
‡ Email:mtkosan@gyte.edu.tr

with P_R projective and both C_R and P_R couniform modules. In this case, every endomorphism f of M_R lifts to an endomorphism f_0 of its projective cover P_R, and we will denote by f_1 the restriction to C_R of f_0. Hence we have a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & C_R & \stackrel{\iota}{\longrightarrow} & P_R & \to & M_R & \to & 0 \\ & & f_1 \downarrow & & \downarrow f_0 & & \downarrow f & & \\ 0 & \to & C_R & \stackrel{\iota}{\longrightarrow} & P_R & \to & M_R & \to & 0. \end{array}$$

In [7, Theorem 2.5], Facchini and Girardi proved that:

- Let $0 \to C_R \to P_R \to M_R \to 0$ be a couniform presentation of a couniformly presented module M_R. Set $K := \{ f \in \mathrm{E}nd(M_R) \mid f \text{ is not surjective} \}$ and $I := \{ f \in \mathrm{E}nd(M_R) \mid f_1 \colon C_R \to C_R \text{ is not surjective} \}$. Then K and I are completely prime two-sided ideals of $\mathrm{E}nd(M_R)$, and the union $K \cup I$ is the set of all non-invertible elements of $\mathrm{E}nd(M_R)$. Moreover, one of the following two conditions holds:

(a) Either $\mathrm{E}nd(M_R)$ is a local ring, or
(b) K and I are the two maximal right, maximal left ideals of $\mathrm{E}nd(M_R)$.

If M_R and M'_R are two couniformly presented modules with couniform presentations $0 \to C_R \to P_R \to M_R \to 0$ and $0 \to C'_R \to P'_R \to M'_R \to 0$, we say that M_R and M'_R *have the same lower part,* and we write $[M_R]_\ell = [M'_R]_\ell$, if there are two homomorphisms $f_0 \colon P_R \to P'_R$ and $f'_0 \colon P'_R \to P_R$ such that $f_0(C_R) = C'_R$ and $f'_0(C'_R) = C_R$.

Recall that a ring R is *semilocal* if $R/J(R)$ is semisimple artinian, that is, isomorphic to a finite direct product of rings $M_{n_i}(D_i)$ of $n_i \times n_i$ matrices over division rings D_i. A ring R is *homogeneous semilocal* if $R/J(R)$ is simple artinian, that is, isomorphic to the ring $M_n(D)$ of all $n \times n$ matrices for some positive integer n and some division ring D [2, 4]. Examples of such rings include all local rings and all simple Artinian rings. If R is a homogeneous semilocal ring, then so are the rings eRe and $M_n(R)$, where e is a nonzero idempotent element of R and $M_n(R)$ is the matrix ring over R. Also, homogeneous semilocal rings appear in a natural way when one localizes a right Noetherian ring with respect to a right localizable prime ideal.

In [4], Corisello and Facchini showed that:

- a homogeneous semilocal ring has a unique maximal proper two-sided ideal and a unique simple module up to isomorphism. Similarly, as in the case of local rings, a homogeneous semilocal ring has only one indecomposable projective module P_R up to isomorphism, and all projective modules are direct sums of copies of this P_R.
- for a module M over any ring R, the Krull-Schmidt theorem holds for M provided $\mathrm{E}nd_R(M)$ is homogeneous semilocal—that is, the direct sum decomposition of M into indecomposable summands is unique up to isomorphism.

In [2], Barioli-Facchini-Raggi proved that:

- The later result fails to extend to modules M_R with finite direct sum decompositions whose indecomposable summands have homogeneous semilocal endomorphism rings,
- If a module M over a ring R has two decompositions $M = M_1 \oplus \cdots \oplus M_t = N_1 \oplus \cdots \oplus N_s$ where all the summands are indecomposable with homogeneous semilocal endomorphism rings, then these two decompositions are isomorphic.

2. The endomorphism ring

The following results describe the endomorphism ring of a cyclic, finitely presented module of projective dimension ≤ 1 over a local ring. Throughout this paper, we will assume that $M_R \neq 0$.

2.1. Theorem. *Let R be a local ring and let $M_R := R_R/I$ be a cyclic, finitely presented module of projective dimension ≤ 1. Suppose $\mathrm{Ext}^1_R(M_R, R_R) = 0$.*

Assume $0 \neq I \neq R$ and let E be the idealizer of the right ideal I of R, that is, the set of all $r \in R$ with $rI \subseteq I$, so that $\mathrm{End}(M_R) \cong E/I$. Set $L := \{\, r \in R \mid rI \subseteq IJ(R) \,\}$ and $K := E \cap J(R)$. Let $\psi\colon E \to \mathrm{End}_R(I/IJ(R))$ be the ring morphism defined by

$$\psi(e)(x + IJ(R)) = ex + IJ(R),$$

for every $e \in E$ and $x \in I$. Let n be the dimension of the right vector space $I/IJ(R)$ over the division ring $R/J(R)$. Then:

(1) *L and K are prime two-sided ideals of E containing I and K is a completely prime ideal of E;*

(2) *For every $e \in E$, the element $e + I$ of E/I is invertible in E/I if and only if $e + J(R)$ is invertible in $R/J(R)$ and $\psi(e)$ is invertible in $\mathrm{End}_R(I/IJ(R))$.*

(3) *The quotient ring E/L is isomorphic to the ring $M_n(R/J(R))$ of all $n \times n$ matrices over the division ring $R/J(R)$.*

(4) *Exactly one of the following two conditions holds:*
(a) *Either $K \subseteq L$, in which case E/I is a homogeneous semilocal ring with Jacobson radical L/I, or*
(b) *L and K are not comparable.*

Proof. (1) and (3). Notice that L is contained in E and is the kernel of ψ, so that L is a two-sided ideal of E. Trivially, I is contained in L. Let us prove that ψ is onto. Let $f\colon I/IJ(R) \to I/IJ(R)$ be a morphism. Since $M_R := R_R/I$ is of projective dimension ≤ 1, the ideal I_R is projective, so that f lifts to a morphism $f'\colon I_R \to I_R$. Apply the functor $\mathrm{H}om(-, R_R)$ to the exact sequence $0 \to I_R \to R_R \to M_R \to 0$, getting a short exact sequence

$$0 \to \mathrm{H}om(M_R, R_R) \to \mathrm{H}om(R_R, R_R) \to \mathrm{H}om(I_R, R_R) \to 0$$

because $\mathrm{Ext}^1_R(M_R, R_R) = 0$. Hence f' can be extended to a morphism $f''\colon R_R \to R_R$, which is necessarily left multiplication by an element $r \in R$. Since f'' restricts to the endomorphism f' of I_R, we get that $r \in E$, and $\psi(e) = f$. This proves that ψ is an onto ring morphism, so that

$$E/L = E/\ker\psi \cong \mathrm{E}nd_R(I/IJ(R)) \cong M_n(R/J(R)).$$

This proves (3).

As $\mathrm{E}nd_R(I/IJ(R)) \cong M_n(R/J(R))$ is a simple ring, it follows that L is a prime ideal and a maximal two-sided ideal. Similarly, K is the kernel of the composite morphism $\varphi\colon E \to R/J(R)$ of the embedding $E \to R$ and the canonical projection $R \to R/J(R)$. Since $R/J(R)$ is a division ring, we get that K is a completely prime, two-sided ideal of E containing I. This concludes the proof of (1).

(2). (:$\Rightarrow$) Since $\varphi(I) = 0$ and $\psi(I) = 0$, the morphisms φ and ψ induce morphisms $\widetilde{\varphi}\colon E/I \to R/J(R)$ and $\widetilde{\psi}\colon E/I \to \mathrm{E}nd(I/IJ(R))$, respectively. Hence $e + I$ invertible implies $\varphi(e) = e + J(R)$ invertible in $R/J(R)$ and $\psi(e)$ is invertible in $\mathrm{E}nd_R(I/IJ(R))$.
($\Leftarrow$:) Assume that $e \in E$ and that $\varphi(e)$ and $\psi(e)$ are invertible in $R/J(R)$ and $\mathrm{E}nd_R(I/IJ(R))$, respectively. Then we have a commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & I & \longrightarrow & R_R & \xrightarrow{\pi} & R_R/I & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle e} & & \downarrow{\scriptstyle e} & & \downarrow{\scriptstyle e} & & \\ 0 & \longrightarrow & I & \longrightarrow & R_R & \xrightarrow{\pi} & R_R/I & \longrightarrow & 0. \end{array}$$

Now $\varphi(e) = e + J(R)$ invertible implies that $e \in R \setminus J(R)$, and so e is invertible in R. Hence the middle vertical arrow is an isomorphism. Since $\psi(e)$ is invertible, it is an automorphism of $I/IJ(R)$, and so $e(I/IJ(R)) = I/IJ(R)$, that is, $eI + IJ(R) = I$. By Nakayama's Lemma, $eI = I$. Hence the left vertical arrow is an epimorphism. By the Snake Lemma, the right vertical arrow is a monomorphism, hence an isomorphism. That is, $e + I$ is invertible in E/I.
(4) We have the three cases (a) $L \subset K$, (b) $K \subseteq L$, and (c) $L \not\subseteq K$ and $K \not\subseteq L$.

Assume $L \subset K$. In this case, $L \subset K \subset E$ implies that $0 \subset K/L \subset E/L$, so that $E/L \cong M_n(R/J(R))$ has a proper non-zero two-sided ideal. This is impossible, because $M_n(R/J)$ is a simple ring. Hence this case cannot occur.

Assume $K \subseteq L$. From (2), it follows that an element $e+I$ of E/I is invertible in E/I if and only if $e + J(R)$ is invertible in $R/J(R)$ and $e + L$ is invertible in E/L. Hence, in order to prove (4) in this case $K \subseteq L$, it suffices to prove that $J(E/I) = L/I$.
($\subseteq$) If $e + I \in J(E/I)$, then $1 - xey + I$ is invertible in E/I for every $x, y \in E$. Thus $1 - xey + L$ is invertible in E/L for all $x, y \in E$, so that $e + L \in J(E/L)$.But $E/L \cong M_n(R/J(R))$ has Jacobson radical 0 so that $e \in L$.
($\supseteq$) Take $l + I \in L/I$ with $l \in L$. Then $1 - xly + L = 1 + L$ in E/L for every $x, y \in E$. Hence $1 - xly + L$ is invertible in E/L. In particular, $1 - xly \notin L$. Thus $1 - xly \notin K$, so that $1 - xly \notin J(R)$. As $R/J(R)$ is a division ring, it follows that $1 - xly + J(R)$ is invertible in $R/J(R)$. Thus $1 - xly + I$ is invertible in E/I, and $l \in J(E/I)$. □

It is known that a finitely presented module over a semilocal ring always has a semilocal endomorphism ring. We have the following natural question.

2.2. Question. Characterize $J(E/I)$. This was done in [1] for cyclically presented modules.

As far as Question 2.2 is concerned, notice that, in the proof of Theorem 2.1(2), we have seen that the mapping

$$\widetilde{\varphi} \times \widetilde{\psi}\colon E/J \to R/J(R) \times \mathrm{E}nd(I/IJ(R))$$

is a local morphism, so that its kernel $K/I \cap L/I$ is contained in $J(E/I)$. In particular, when $K \subseteq L$, we have that $L/I = J(E/I)$ as we have seen in Theorem 2.1(4)(a). We are not able to describe $J(E/I)$ when K and K are not comparable.

2.3. Remark. Let R be a local right self-injective ring. Let M_R be a cyclic and finitely presented module of projective dimension ≤ 1. Since R_R is injective, we have that $\mathrm{Ext}^1_R(M_R, R_R) = 0$. Thus, Theorem 2.1 can be applied.

Let A and B be two modules. We say that:

- A and B *have the same monogeny class*, and write $[A]_m = [B]_m$, if there exist a monomorphism $A \to B$ and a monomorphism $B \to A$ [5];

• A and B *have the same epigeny class*, and write $[A]_e = [B]_e$, if there exist an epimorphism $A \to B$ and an epimorphism $B \to A$;

It is clear that a module A has the same monogeny (epigeny) class as the zero module if and only if $A = 0$.

• Two cyclically presented modules R/aR and R/bR over a local ring R are said to *have the same lower part*, denoted $[R/aR]_l = [R/bR]_l$, if there exist $r, s \in R$ such that $raR = bR$ and $sbR = aR$ [1].

• If M_R and M'_R are two couniformly presented modules with couniform presentations

$$0 \to C_R \to P_R \to M_R \to 0$$

and

$$0 \to C'_R \to P'_R \to M'_R \to 0,$$

we say that M_R and M'_R *have the same lower part,* and we write $[M_R]_\ell = [M'_R]_\ell$, if there are two homomorphisms $f_0 \colon P_R \to P'_R$ and $f'_0 \colon P'_R \to P_R$ such that $f_0(C_R) = C'_R$ and $f'_0(C'_R) = C_R$ [7].

2.4. Theorem. *Let R be a semiperfect ring and let R_R/L be a cyclic uniform right R-module with $L \neq 0$. Let E be the idealizer of the right ideal L of R, that is, the set of all $r \in R$ with $rL \subseteq L$, so that*

$$\mathrm{E}nd(R_R/L) \cong E/L.$$

Similarly, let E' be the idealizer of the right ideal $L + J(R)$ of R, so that

$$\mathrm{E}nd(R_R/(L + J(R))) \cong E'/(L + J(R)).$$

Set $I := \{\, e \in E \mid$ left multiplication by $e + I$ is a non-injective endomorphism of $R_R/L \,\}$ and $K := E \cap (L + J(R))$. Then:

(1) *I and K are two two-sided ideals of E containing L, and I is completely prime in E.*
(2) *For every $e \in E$, the element $e + L$ of E/L is invertible in E/L if and only if $e + L + J(R)$ is invertible in $E'/L + J(R)$ and $e \notin I$.*
(3) *Moreover:*
  (a) *If $I \subseteq K$, then every epimorphism $R_R/L \to R_R/L$ is an automorphism of R_R/L,*
  (b) *$K \nsubseteq I$ if and only if $[R_R/L]_m = [L + J(R)/L]_m$.*

Proof. (1) We know that $\mathrm{E}nd(R_R/L) \cong E/L$. Every endomorphism $e + L$ of R_R/L extends to an endomorphism e_1 of the injective envelope $E(R_R/L)$. Define a ring morphism

$$\varphi \colon E \to \mathrm{E}nd(E(R_R/L))/J(\mathrm{E}nd(E(R_R/L)))$$

by $\varphi(e) = e_1 + J(\mathrm{E}nd(E(R_R/L)))$ for every $e \in E$. Since R_R/L is uniform, the injective envelope $E(R_R/L)$ is indecomposable, the endomorphism ring $\mathrm{E}nd(E(R_R/L))$ is a local ring, and the Jacobson radical $J(\mathrm{E}nd(E(R_R/L)))$ consists of all non-injective endomorphisms of $E(R_R/L)$. It follows that I, which is equal to the kernel of the ring morphism φ, whose range is the division ring

$$\mathrm{E}nd(E(R_R/L))/J(\mathrm{E}nd(E(R_R/L))),$$

must be a completely prime two-sided ideal of E. The remaining part of statement (1) is easily checked.

(2) We have already seen that there is a ring morphism

$$\varphi \colon E \to \mathrm{E}nd(E(R_R/L))/J(\mathrm{E}nd(E(R_R/L)))$$

whose kernel is I. Hence if $e \in E$ and $e + L$ is invertible in E/L, then $\varphi(e)$ must be invertible in the division ring $\mathrm{E}nd(E(R_R/L))/J(\mathrm{E}nd(E(R_R/L)))$. Thus $\varphi(e) \neq 0$, that is, $e \notin \ker \varphi = I$. Similarly, we can consider the ring morphism

$$\psi\colon E \to \mathrm{E}nd(R_R/L + J(R))$$

defined by $\psi(e)(r + L + J(R)) = er + L + J(R)$ for every $e \in E$ and every $r \in R$. Its kernel is K, which contains L. Hence $e + L$ invertible in E/L implies $\psi(e)$ invertible in $\mathrm{E}nd(R_R/L + J(R))$. But

$$\mathrm{E}nd(R_R/(L + J(R))) \cong E'/(L + J(R)),$$

so that $e + L + J(R)$ must be invertible in $E'/L + J(R)$.

Conversely, assume $e \in E$, $e+L+J(R)$ invertible in $E'/L+J(R)$ and $e \notin I$. We want to show that $e + L$ is invertible in E/L. Since $E/L \cong \mathrm{E}nd(R_R/L)$, this is equivalent to showing that left multiplication $\mu_e\colon R_R/L \to R_R/L$ by e is an automorphism of R_R/L. Now $e \notin I$ is equivalent to μ_e is injective by definition of I. In order to show that μ_e is onto as well, it suffices to prove that μ_e induces an onto endomorphism

$$(R_R/L)/(R_R/L)J(R) \to (R_R/L)/(R_R/L)J(R)$$

by Nakayama's Lemma. But $(R_R/L)J(R) = L + J(R)/L$, so that

$$(R_R/L)/(R_R/L)J(R) \cong R_R/L + J(R).$$

Hence $e + L + J(R)$ invertible in $E'/L + J(R) \cong \mathrm{E}nd(R_R/(L + J(R)))$ means that the endomorphism $\psi(e)$ of $R_R/L + J(R)$ induced by μ_e is onto, as desired.
(3) (a) Assume $I \subseteq K$. Let $e + L\colon R_R/L \to R_R/L$ be an epimorphism with $e \in E$. Then the induced morphism $\psi(e)\colon R_R/L + J(R) \to R_R/L + J(R)$ is also an epimorphism, so that it is an automorphism because $R_R/L + J(R)$ is a semisimple module of finite Goldie dimension. In the isomorphism

$$\mathrm{E}nd(R_R/(L + J(R))) \cong E'/(L + J(R)),$$

we obtain that $e + L + J(R)$ is invertible in the ring $E'/(L + J(R))$. Thus $e \notin K$. Hence $e \notin I$. It follows from (2) that $e + L$ is invertible, that is, it is an automorphism of R_R/L.
(b) Assume $K \nsubseteq I$. Then there is an element $f \in K$, $f \notin I$. Thus $f \in E$ induces an endomorphism f of R_R/L. Now $f \notin I$ means that f is injective, and $f \in K$ means that the image of f is contained in $L+J(R)/L$. Hence $[R_R/L]_m = [L+J(R)/L]_m$. Conversely, if $[R_R/L]_m = [L + J(R)/L]_m$, then there is a monomorphism $f\colon R_R/L \to L + J(R)/L$. If we compose it with the inclusion $L + J(R)/L \to R_R/L$ we get an endomorphism of R_R/L which is in K but not in I. Hence $K \nsubseteq I$. □

We finish this study with the following result.

2.5. Theorem. *Let R be a semiperfect ring, let $R/L, R/L'$ be two cyclic uniform modules with $L \neq 0$ and $L' \neq 0$ proper right ideals of R. Assume that either*

(1) *every monomorphism $R_R/L \to R_R/L$ is an automorphism of R_R/L, or*
(2) *every epimorphism $R_R/L \to R_R/L$ is an automorphism of R_R/L, or*
(3) $[R_R/L]_m = [L + J(R)/L]_m$.

Then the followings are equivalent.

(*a*) $R_R/L \cong R_R/L'$
(*b*) $[R_R/L]_m = [R_R/L']_m$ *and* $[R_R/L]_e = [R_R/L']_e$.

Proof. Assume $[R_R/L]_m = [R_R/L']_m$ and $[R_R/L]_e = [R_R/L']_e$. Then there are monomorphisms $\alpha\colon R_R/L \to R_R/L'$ and $\beta\colon R_R/L' \to R_R/L$ and epimorphisms $\alpha\colon R_R/L \to R_R/L'$ and $\beta\colon R_R/L' \to R_R/L$. Then $\beta\alpha$ is a monomorphism $R_R/L \to R_R/L$ and $\beta'\alpha'$ is an epimorphism $R_R/L \to R_R/L$. If hypothesis (a) holds, then $\beta\alpha$ is an automorphism of R_R/L that factors through R_R/L', so that R_R/L is isomorphic to a direct summand of R_R/L'. But $R_R/L \neq 0$ and R_R/L' is uniform, so that $R_R/L \cong R_R/L'$. This proves our theorem under hypothesis (a). Dually one proves that the theorem holds when hypothesis (b) holds.

Assume now that hypothesis (c) holds, i.e., $[R_R/L]_m = [L+J(R)/L]_m$. Equivalently, there exists a monomorphism $\gamma\colon R_R/L \to R_R/L$ whose image is contained in $L+J(R)/L$. Now if either α or α' are isomorphisms, then the existence of α or α' shows that $R_R/L \cong R_R/L'$. This allows us to conclude. Thus we can assume that α is not an epimorphism and α' is not a monomorphism. Then $\alpha' + \alpha\gamma\colon R_R/L \to R_R/L'$ is an isomorphism, because:

(1) It is injective, because it is the sum of the injective morphism $\alpha\gamma\colon R_R/L \to R_R/L'$ and the non-injective morphism $\alpha'\colon R_R/L \to R_R/L'$, and R_R/L is uniform.

(2) The ideal $J(R)$ is superfluous in R_R by Nakayama's Lemma. Considering the canonical projection $R_R \to R_R/L$, it follows that $L + J(R)/L$ is superfluous in R_R/L. Applying the morphism $\alpha\colon R/L \to R/L'$, we get that the image of $\alpha\gamma$ is contained in $\alpha(L+J(R)/L)$, hence is a superfluous submodule of R/L'. Thus the sum of $\alpha\gamma$ and the surjective morphism $\alpha'\colon R/L \to R/L'$ is a surjective morphism $\alpha'+\alpha\gamma\colon R_R/L \to R_R/L'$.

Thus $\alpha + \alpha'\gamma$ is an isomorphism of R_R/L onto R_R/L'. □

2.6. Remark. By Theorem 2.4, the only case in which we cannot apply Theorem 2.5 is when K is properly contained in I. Namely, if $K \not\subseteq I$, then $[R_R/L]_m = [L + J(R)/L]_m$ and we can apply Theorem 2.5(a); if $K \subseteq I$, then either K is properly contained in I, which is the case still unknown, or $K = I$, but in the latter case every epimorphism $R_R/L \to R_R/L$ is an automorphism of R_R/L by Theorem 2.4(1).

Acknowledgments. The authors are grateful to the referee for drawing our attention to a number of typos. Also it is a pleasure to thank Prof. A. Facchini for his helpful comments.

References

[1] Amini, B., Amini, A. and Facchini, A. *Equivalence of diagonal matrices over local rings,* J. Algebra, **320**, 1288-1310, 2008.

[2] Barioli, F., Facchini, A. and Raggi, F. J. R. *Krull-Schmidt theorem and homogeneous semilocal rings*, Commun. Algebra, **29**, 1649-1658, 2001.

[3] Bumby, R. T. *Modules which are isomorphic to submodules of each other*, Arch. Math., **16**, 184-185, 1965.

[4] Corisello, R. and Facchini, A. *Homogeneous semilocal rings*, Commun. Algebra, **29**, 1807-1819, 2001.

[5] Facchini, A. *Krull-Schmidt fails for serial modules*, Trans. Amer. Math. Soc., **348**, 4561-4575, 1996.

[6] Facchini, A. *Module Theory: Endomorphism rings and direct sum decompositions in some classes of modules*, (Progress in Math., 167, Birkhäuser Verlag, Basel, 1998).

[7] Facchini, A. and Girardi, N. *Couniformly presented modules and dualities*, (in Advances in Ring Theory, Dinh Van Huynh and Sergio R. Lopez Permouth Eds., Trends in Math., Birkhauser Verlag, Basel, Heidelberg, London, New York, pp. 149-163 2010).

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 993 – 1000

On Lagrangian submersions

Hakan Mete Taştan*

Abstract

In this paper, we study Riemannian, anti-invariant Riemannian and Lagrangian submersions. We prove that the horizontal distribution of a Lagrangian submersion from a Kählerian manifold is integrable. We also give some applications of this result. Moreover, we investigate the effect of the submersion to the geometry of its total manifold and its fibers.

2000 AMS Classification: 53C15, 53B20.

Keywords: Riemannian submersion, Lagrangian submersion, horizontal distribution, Kählerian manifold.

Received 15 : 07 : 2013 : Accepted 04 : 11 : 2013 Doi : 10.15672/HJMS.2014437529

1. Introduction

The theory of Riemannian submersions was initiated by O'Neill [11]. In [18], the Riemannian submersions were considered between almost Hermitian manifolds by Watson under the name of almost Hermitian submersions. In this case, the Riemannian submersion is also an almost complex mapping and consequently the vertical and horizontal distribution are invariant with respect to the almost complex structure of the total manifold of the submersion. Afterwards, almost Hermitian submersions have been actively studied between different kind of subclasses of almost Hermitian manifolds, for example, see [5]. We note that almost Hermitian submersions have been extended to different kind of subclasses of almost contact manifolds, for example, see [14]. Most of the studies related to Riemannian or almost Hermitian submersions can be found in the book [4]. The study of anti-invariant Riemannian submersions from almost Hermitian manifolds were initiated by Şahin [15]. In this case, the fibres are anti-invariant with respect to the almost complex structure of the total manifold. A Lagrangian submersion is a special case of an anti-invariant Riemannian submersion such that the almost complex structure of the total manifold reverses the vertical and horizontal distributions. In this paper, we consider Riemannian, anti-invariant Riemannian and Lagrangian submersions. We will focus Lagrangian submersions from a Kählerian manifold onto a Riemannian manifold

*İstanbul Universty, Department of Mathematics, Vezneciler, 34134, İstanbul, Turkey
Email: hakmete@istanbul.edu.tr

and prove that the horizontal distribution of such a submersion is integrable and totally geodesic. Using this result we obtain that such a submersion is a totally geodesic map if and only if it has totally geodesic fibers. We also obtained other applications of the result. In the last section, we show that non-existence of a Lagrangian submersion with totally geodesic fibers from a non-flat Kählerian manifold. We also proved that if the fibers of a Lagrangian submersion are totally umbilical, then the fibers are minimal.

2. Riemannian submersions

In this section, we give necessary background for Riemannian submersions.

Let (M, g) and (N, g_N) be Riemannian manifolds, where $dim(M) > dim(N)$. A surjective mapping $\pi : (M, g) \to (N, g_N)$ is called a *Riemannian submersion* [11] if:

(S1) π has maximal rank, and

(S2) π_*, restricted to $(ker\pi_*)^\perp$, is a linear isometry.

In this case, for each $q \in N$, $\pi^{-1}(q)$ is a k-dimensional submanifold of M and called *fiber*, where $k = dim(M) - dim(N)$. A vector field on M is called *vertical* (resp. *horizontal*) if it is always tangent (resp. orthogonal) to fibers. A vector field X on M is called *basic* if X is horizontal and π-related to a vector field X_* on N, i.e., $\pi_* X_p = X_{*\pi(p)}$ for all $p \in M$. As usual, we denote by $\mathcal{V}$ and $\mathcal{H}$ the projections on the vertical distribution $ker\pi_*$ and the horizontal distribution $(ker\pi_*)^\perp$, respectively. The geometry of Riemannian submersions is characterized by O'Neill's tensors T and A, defined as follows:

$$\mathrm{T}_E F = \mathcal{V}\nabla_{\mathcal{V}E}\mathcal{H}F + \mathcal{H}\nabla_{\mathcal{V}E}\mathcal{V}F, \tag{2.1}$$

$$\mathrm{A}_E F = \mathcal{V}\nabla_{\mathcal{H}E}\mathcal{H}F + \mathcal{H}\nabla_{\mathcal{H}E}\mathcal{V}F \tag{2.2}$$

for any vector fields E and F on M, where ∇ is the Levi-Civita connection of g_M. It is easy to see that T_E and A_E are skew-symmetric operators on the tangent bundle of M reversing the vertical and the horizontal distributions. We summarize the properties of the tensor fields T and A. Let V, W be vertical and X, Y be horizontal vector fields on M, then we have

$$\mathrm{T}_V W = \mathrm{T}_W V, \tag{2.3}$$

$$\mathrm{A}_X Y = -\mathrm{A}_Y X = \frac{1}{2}\mathcal{V}[X, Y]. \tag{2.4}$$

On the other hand, from (2.1) and (2.2), we obtain

$$\nabla_V W = \mathrm{T}_V W + \mathcal{V}\nabla_V W, \tag{2.5}$$

$$\nabla_V X = \mathrm{T}_V X + \mathcal{H}\nabla_V X, \tag{2.6}$$

$$\nabla_X V = \mathrm{A}_X V + \mathcal{V}\nabla_X V, \tag{2.7}$$

$$\nabla_X Y = \mathcal{H}\nabla_X Y + \mathrm{A}_X Y, \tag{2.8}$$

and if X is basic, then $\mathcal{H}\nabla_V X = \mathrm{A}_X V$. It is not difficult to observe that T acts on the fibers as the second fundamental form while A acts on the horizontal distribution and measures of the obstruction to the integrability of this distribution. For details on the Riemannian submersions, we refer to O'Neill's paper [11] and to the book [4].

3. Anti-invariant Riemannian submersions

A smooth manifold M is called almost Hermitian [19] if its tangent bundle has an almost complex structure J and a Riemannian metric g such that

$$g(E,F) = g(JE,JF) \tag{3.1}$$

for any vector fields E and F on M. Let M be a $2m$-dimensional almost Hermitian manifold with Hermitian metric g and almost complex structure J, and N be a Riemannian manifold with Riemannian metric g_N. Suppose that there exists a Riemannian submersion $\pi : M \to N$ such that $ker\pi_*$ is anti-invariant with respect to J, i.e., $J(ker\pi_*) \subseteq (ker\pi_*)^\perp$. Then the Riemannian submersion π is called an *anti-invariant Riemannian submersion.* For the details, see [15].

There are some other recent paper which involve other structures such as almost product [6], almost contact [9], Sasakian [7] and cosymplectic [8]. In any cases, the definition of anti-invariant Riemannian submersion is the same as the above definition. Besides there are many other notions related with that of anti-invariant Riemannian submersion, such as slant submersion [16] and semi-invariant submersion [17]. The key of this definitions consists on considering the fibres as submanifolds of the almost Hermitian manifold M having the corresponding property. Because of that, we may consider that the following names are more convenient: totally real, instead of anti-invariant, but semi-invariant (cfr. [17]) of CR-submersion (cfr. e.g. [10]) because definition of a CR-submersion depends on certain CR-submanifold of the total manifold, instead of the fact the fibres are CR-submanifolds. As one can see, names are quite complex in this field.

An almost Hermitian manifold M is called a *Kählerian manifold* if

$$(\nabla_E J)F = 0 \tag{3.2}$$

for any vector fields E and F on M, where ∇ is the Levi-Civita connection on M. Let (M,g,J) be a Kählerian manifold. The Riemannian curvature tensor [19] of (M,g,J) is defined by $\mathrm{R}(E,F)G = \nabla_{[E,F]}G - [\nabla_E, \nabla_F]G$ for vector fields E, F and G on M. We put $\mathrm{R}(E,F,G,K) = g(\mathrm{R}(E,F)G,K)$ where K is a vector field on M. The *holomorphic sectional curvature* [19] of M is defined for any unit vector field E tangent to M via

$$\mathrm{H}(E) = \mathrm{R}(E,JE,E,JE). \tag{3.3}$$

We note that a Kählerian manifold with vanishing holomorphic sectional curvature is flat [19]. The manifold M is called a *complex space form* if it is of constant holomorphic sectional curvature. We denote by $M(c)$ a complex space form of constant holomorphic sectional curvature c. Then the Riemannian curvature tensor R of $M(c)$ is given by

$$\begin{aligned} \mathrm{R}(E,F)G = \tfrac{c}{4}\{&g(F,G)E - g(E,G)F + g(JF,G)JE \\ &-g(JE,G)JF + 2g(E,JF)JG\} \end{aligned} \tag{3.4}$$

for any vector fields E, F and G on $M(c)$ [19]. In this point, we give the following proposition.

3.1. Proposition. *Let $\pi : M(c) \to N$ be a Riemannian submersion from a complex space form $M(c)$ with $c \neq 0$ onto a Riemannian manifold N. Then the fibers of $M(c)$ are invariant or anti-invariant with respect to the almost complex structure J of $M(c)$ if and only if*

$$g((\nabla_U \mathrm{T})_V W, X) = g((\nabla_V \mathrm{T})_U W, X), \tag{3.5}$$

where U, V and W are vertical vector fields and X is a horizontal vector field on $M(c)$.

Proof. Let U, V and W be vertical vector fields and X be a horizontal vector field on $M(c)$. Then from (3.4), we have

$$\text{(3.6)} \qquad \begin{aligned} \mathrm{R}(U,V)W = \tfrac{c}{4}\{&g(V,W)U - g(U,W)V + g(JV,W)JU \\ &-g(JU,W)JV + 2g(U,JV)JW\} \end{aligned} .$$

From (14), we see that $\mathrm{R}(U,V)W$ is vertical, if the fibers are invariant or anti-invariant with respect to the almost complex structure J of $M(c)$. So, we get easily, $\mathrm{R}(U,V,W,X) = 0$. Therefore, (3.5) follows from the following O'Neill curvature formula {1} [11]:

$$\mathrm{R}(U,V,W,X) = g((\nabla_V \mathrm{T})_U W, X) - g((\nabla_U \mathrm{T})_V W, X).$$

Conversely, assume that (3.5) holds. Then for U, V and W, it is not difficult to see that $\mathrm{R}(U,V)W$ is vertical from the above O'Neill curvature formula. If we put $W = U$ in (3.6), then we have

$$\text{(3.7)} \qquad \mathrm{R}(U,V)U = \frac{c}{4}\{g(V,U)U - g(U,U)V + g(U,JV)JU\}.$$

Thus, we see that $g(U,JV)JU$ is vertical from (15), since $\mathrm{R}(U,V)U$ is vertical. So, we conclude that either JU is vertical or $g(U,JV) = 0$. It means that either $J(ker\pi_*) \subseteq ker\pi_*$ or $J(ker\pi_*) \subseteq (ker\pi_*)^{\perp}$, i.e., either the fibers are invariant or anti-invariant with respect to the almost complex structure J of $M(c)$. □

3.2. Corollary. *Let $\pi : M(c) \to N$ be an anti-invariant Riemannian submersion from a complex space form $M(c)$ with $c \neq 0$ onto a Riemannian manifold N. Then the equality (3.5) holds.*

4. Lagrangian submersions

Let M be a $2m$-dimensional almost Hermitian manifold with Hermitian metric g and almost complex structure J, and N be a Riemannian manifold with Riemannian metric g_N and let $\pi : M \to N$ be an anti-invariant Riemannian submersion. Then we call π a *Lagrangian Riemannian submersion* or briefly, a *Lagrangian submersion*, if $dim(ker\pi_*) = dim((ker\pi_*)^{\perp})$. In this case, the almost complex structure J of M reverses the vertical and the horizontal distributions, i.e., $J(ker\pi_*) = (ker\pi_*)^{\perp}$ and $J((ker\pi_*)^{\perp}) = ker\pi_*$.

In Symplectic Geometry, a Lagrangian submersion $\pi : (M,\omega) \to N$ from a symplectic manifold onto a manifold is a submersion having the fibres Lagrangian submanifolds (see, e.g. [1]), i.e., $\omega|_{\pi^{-1}(q)} = 0$.

An almost Hermitian structure (J, g) defines an almost symplectic structure $\omega(X,Y) = g(JX,Y)$, and then we can consider compare both definitions. It is easily shown that they coincide:

4.1. Lemma. *Let $\pi : (M,J,g) \to N$ be a submersion from an almost Hermitian manifold onto a manifold. Then the following conditions are equivalent:*
(1) The fibres of π are Lagrangian submanifolds.
(2) $J(ker\pi_*) = (ker\pi_*)^{\perp}$.
Moreover, the horizontal distribution $(ker\pi_*)^{\perp}$ is also Lagrangian.

Proof. (1) ⇒ (2). Let X and Y be vertical, that is; $X, Y \in ker\pi_*$. Then $g(JX,Y) = \omega(X,Y) = 0$, thus proving (2). Reversing the reasoning, one has the other implication. □

In order to have a Lagrangian submersion $\pi : (M, J, g) \to N$ dimensions must be related in the following way: $dim(M) = 2dim(N)$. The most natural examples of manifolds having this relation are given by the tangent (resp. cotangent) bundle of $M = TN \to N$ (resp. $M = T^*N \to N$). In the seminal paper [2], Dombrowski introduces the almost complex structure J on the tangent bundle TN of a manifold N having a linear connection, which is given by the conditions $J(X^H) = X^V$; $J(X^V) = -X^H$, H and V being the horizontal and vertical lifts. On the other hand, Sasaki [13] introduced the diagonal lift g^D, or Sasaki metric, over the tangent bundle of a Riemannian manifold (N, g), given by $g(X^H, Y^H) = g(X^V, Y^V) = g(X, Y)$; $g(X^H, Y^V) = 0$. Thus, the tangent bundle (TN, J, g^D) of a Riemannian manifold (N, g) is an almost Hermitian manifold. Then one easily obtains:

4.2. Lemma. *With the above notation,* $\pi : (TN, J, g^D) \to (N, g)$ *is a Lagrangian submersion.*

We want to emphasize that the same considerations can be done about the cotangent bundle.

Let M be a Kählerian manifold with Hermitian metric g and almost complex structure J, and N be a Riemannian manifold with Riemannian metric g_N. Now we examine how the Kählerian structure on M places restrictions on the tensor fields T and A of a Lagrangian submersion $\pi : M \to N$.

4.3. Lemma. *Let* $\pi : M \to N$ *be a Lagrangian submersion from a Kählerian manifold* M *onto a Riemannian manifold* N*. Then we have*

$$\textbf{\textit{a)}} \quad \mathrm{T}_V JE = J\mathrm{T}_V E \qquad\qquad \textbf{\textit{b)}} \quad \mathrm{A}_X JE = J\mathrm{A}_X E$$

where V *is a vertical vector field,* X *is a horizontal vector field, and* E *is a vector field on* M*.*

Proof. Using (2.5)-(2.8), we obtain easily both assertions from (3.2). □

We remark that Lemma 4.3 was proved partially in [15].

4.4. Corollary. *Let* $\pi : M \to N$ *be a Lagrangian submersion from a Kählerian manifold* M *onto a Riemannian manifold* N*. Then we have*

$$\mathrm{A}_X JY = -\mathrm{A}_Y JX$$

where X *and* Y *are any horizontal vector fields on* M*.*

Proof. Let X and Y be any horizontal vector fields on M, from Lemma 4.3-b), we have $\mathrm{A}_X JY = J\mathrm{A}_X Y$. Since the tensor A has the alternation property, we get $J\mathrm{A}_X Y = -J\mathrm{A}_Y X = -\mathrm{A}_Y JX$. □

4.1. The Horizontal Distribution. We now prove that the horizontal distribution $(ker\pi_*)^\perp$ is integrable and totally geodesic. It is well-known that the vertical distribution $ker\pi_*$ is always integrable.

4.5. Theorem. *Let* $\pi : M \to N$ *be a Lagrangian submersion from a Kählerian manifold* M *onto a Riemannian manifold* N*. Then the horizontal distribution* $(ker\pi_*)^\perp$ *is integrable and totally geodesic.*

Proof. Let X and Y be any horizontal vector fields on M, since $\mathrm{A}_X Y = \frac{1}{2}\mathcal{V}[X, Y]$, it is sufficient to show that $\mathrm{A}_X = 0$. If Z is a horizontal vector field on M, then using (2.4), (2.8), (3.1), (3.2) and Corollary 4.4, we have

$$g(\mathrm{A}_X JY, Z) = -g(\mathrm{A}_Y JX, Z) = -g(\nabla_Y JX, Z) = -g(J\nabla_Y X, Z)$$
$$= g(\nabla_Y X, JZ) = -g(\mathrm{A}_X Y, JZ) = g(\mathrm{A}_X JZ, Y) = -g(\mathrm{A}_Z JX, Y)$$
$$= g(\mathrm{A}_Z Y, JX) = -g(\mathrm{A}_Y Z, JX) = g(\mathrm{A}_Y JX, Z) = -g(\mathrm{A}_X JY, Z).$$

Therefore $\mathrm{A}_X JY = 0$. By Proposition 2.7-(e) ([18]), we get $\mathrm{A}_X = 0$. □

4.2. Applications. In this subsection, we give some applications of Theorem 4.5.

4.6. Corollary. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. Then we have*

$$(\nabla\pi_*)(X, JY) = (\nabla\pi_*)(JX, Y) = 0,$$

where X and Y are any horizontal vector fields on M, and $\nabla\pi_*$ is the second fundamental form [15] *of π.*

Proof. It follows immediately from our main result Theorem 4.5, Corollary 3.1([15]) and Corollary 3.2([15]). □

It is well-known that a differential map π between two Riemannian manifolds is called totally geodesic if $\nabla\pi_* = 0$. Now we give a necessary and sufficient condition for a Lagrangian submersion to be a totally geodesic map.

4.7. Theorem. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. Then π is a totally geodesic map if and only if it has totally geodesic fibers.*

Proof. Let V and W be any vertical vector fields on M, if $\mathrm{T}_V JW = 0$, then from Lemma 4.3, we get $\mathrm{T}_V W = 0$. On the other hand, from Proposition 2.7-(d)([18]), it follows that $\mathrm{T}_V = 0$, which means that the Lagrangian submersion π has totally geodesic fibers. Thus the assertion follows from Theorem 4.5 and Theorem 3.4([15]). □

Now, we simply decompose theorems given in [15]. First, we recall the following facts given in [12].

Let $B = M \times N$ be a Riemannian manifold with metric g. Assume that the canonical foliations $\mathcal{D}_M$ and $\mathcal{D}_N$ intersect perpendicularly everywhere. Then g is the metric tensor of

(i) a twisted product $M \times_f N$ if and only if $\mathcal{D}_M$ is a totally geodesic foliation and $\mathcal{D}_N$ is a totally umbilical foliation,

(ii) a usual product of Riemannian manifolds if and only if $\mathcal{D}_M$ and $\mathcal{D}_N$ are totally geodesic foliations.

Thus, from Theorem 4.5, Theorem 4.2([15]) and Theorem 4.3([15]), we have the following result.

4.8. Theorem. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. Then*
a) *M is a locally twisted product manifold of the form $M_{(ker\pi_*)^\perp} \times_f N_{ker\pi_*}$ if and only if π has totally umbilical fibers,*
b) *M is a locally product of manifold if and only if π has totally geodesic fibers.*

5. The Geometry of Total Manifold and Fibers

In this section, we prove some characterization results for a Lagrangian submersion from a Kählerian manifold onto a Riemannian manifold.

Let M be a Kählerian manifold with Hermitian metric g and almost complex structure J and let $\pi : M \to N$ be a Lagrangian submersion from the manifold M onto a Riemannian manifold N. Since $\mathrm{A} \equiv 0$, the O'Neill's curvature formula {2} [11] reduces to

(5.1) $\quad \mathrm{R}(X, V, Y, W) = g((\nabla_X \mathrm{T})_V W, Y) - g(\mathrm{T}_V X, \mathrm{T}_W Y),$

where V and W are vertical, and X and Y are horizontal vector fields on M.

5.1. Theorem. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. Then the holomorphic sectional curvature* H *of M satisfies*

a) $\mathrm{H}(X) = g_M((\nabla_X \mathrm{T})_{JX} JX, X) - \|\mathrm{T}_{JX} X\|^2,$

b) $\mathrm{H}(V) = g_M((\nabla_{JV} \mathrm{T})_V V, JV) - \|\mathrm{T}_V V\|^2,$

where X is a unit horizontal and V is a unit vertical vector field on M.

Proof. Both assertion **a)** and assertion **b)** follow easily from (3.3), (5.1), Lemma 4.3 and (3.1). □

We know from Proposition 1.2([3]) that if T is parallel, i.e., $\nabla_E \mathrm{T} = 0$, for any vector field E on M, then $\mathrm{T} = 0$. Therefore, by Theorem 5.1 we obtain the following result.

5.2. Theorem. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. If the tensor field* T *is parallel, then the holomorphic sectional curvature* H *of M vanishes. Namely, M is flat.*

We remark that Theorem 5.2 describes the geometry of the total manifold of the Lagrangian submersion studied above. On the other hand, if the tensor T vanishes, then the fibers are totally geodesic. Thus, from Theorem 5.1 and Theorem 5.2, we have the following result.

5.3. Corollary. *Let M be a non-flat Kählerian manifold. Then there is no Lagrangian submersion π with totally geodesic fibers from M onto a Riemannian manifold N.*

Now, we recall that any fiber of a Riemannian submersion $\pi : (M, g) \to (N, g_N)$ is called *totally umbilical* if

(5.2) $\quad \mathrm{T}_U V = g(U, V)\eta$

for any $U, V \in ker\pi_*$, where η is the mean curvature vector field of the fiber in M. The fiber is called *minimal*, if $\eta = 0$, identically [4].

5.4. Proposition. *Let $\pi : M \to N$ be a Lagrangian submersion from a Kählerian manifold M onto a Riemannian manifold N. If the fibers of M are totaly umbilical, then either $ker\pi_* = \{0\}$ or 1-dimensional or the mean curvature vector field η vanishes, i.e., the fibers are minimal.*

Proof. If $ker\pi_* = \{0\}$ or $Dim(ker\pi_*) = 1$, then the conclusion is obvious. If $Dim(ker\pi_*) \geq 2$, then we can choose $U, V \in ker\pi_*$, such that $g(U, V) = 0$ and $\|U\| = 1$. By Lemma 4.3-(a) and (5.2), we have

$g(\eta, JV) = g(\mathrm{T}_U U, JV) = -g(\mathrm{T}_U JV, U) = -g(J\mathrm{T}_U V, U) = 0$. Hence, it follows that $\eta = 0$. □

Acknowledgements. The author is deeply indebted to the referee(s) for useful suggestions and valuable comments.

References

[1] Castanõ-Bernard R., Matessi, D. *Lagrangian 3-torus fibrations,* J. Differential Geom. **81**(3), 483-573, 2009.

[2] Dombrowski, P. *On the geometry of tangent bundle,* J. Reine Angew. Math. **210**, 73-88, 1962.

[3] Escobales R.H. Jr. *Riemannian submersions with totally geodesic fibers,* J. Differential Geom. **10**(1), 253-276, 1975.

[4] Falcitelli M., Ianus S. and Pastore A.M. *Riemannian submersions and related topics* (World Scientific, River Edge, NJ, 2004).

[5] Falcitelli M., Pastore A.M. *A note on almost Kähler and nearly Kähler submersions,* J. Geom. **69**, 79-87, 2000.

[6] Gündüzalp, Y. *Anti-invariant Riemannian submersions from almost product Riemannnian manifolds,* Math. Sci. and Appl. E-notes, **1**, 58-66, 2013.

[7] Küpeli Erken, I., Murathan C. *Anti-invariant Riemannian submersions from Sasakian manifolds,* arXiv:1302.4906.

[8] Küpeli Erken, I., Murathan C. *Anti-invariant Riemannian submersions from cosymplectic manifolds,* arXiv:1302.5108.

[9] Lee, J. W. *Anti-invariant $\xi^{\perp}$ Riemannian submersions from almost contact manifolds,* Hacettepe Journal of Mathematics and Statistic, **42**(3), 231-241, 2013.

[10] Mangione, V. *Some recent contributions to CR-submersions,* Proceedings of The Conference of Apllied differential Geometry-General Relativity and The Worhshop on Global Analysis, Differential Geometry and Lie Algebras, 85-90, 2002.

[11] O'Neill B. *The fundamental equations of a submersion,* Mich. Math. J. **13**, 458-469, 1966.

[12] Ponge R. and Reckziegel H. *Twisted products in Pseudo-Riemannian geometry,* Geom. Dedicata. **48**(1), 15-25, 1993.

[13] Sasaki S. *On the differential geometry of tangent bundles of Riemannian manifolds* Tohô ku. Math. J. **10**, 338-354, 1958.

[14] Shahid M.H., Al-Solamy F.R. Yun J.B. and Ahmad M. *Submersion of semi-invariant submanifolds of Trans-Sasakian manifold,* Bull. Malays. Math. Sci. Soc. (2) **36**(1), 63-71, 2013.

[15] Ṣahin B. *Anti-invariant Riemannian submersions from almost Hermitian manifolds,* Cent. Eur. J. Math. **8**(3), 437-447, 2010.

[16] Ṣahin B. *Slant submersions from almost Hermitian manifolds,* Bull. Math. Soc. Sci. Math. Roumanie **54**(102), No. 1, 93-105, 2011.

[17] Ṣahin B. *Riemannian submersions from almost Hermitian manifolds,* Taiwanese J. Math. **17**(2), 629-659, 2013.

[18] Watson B. *Almost Hermitian submersions,* J. Differential Geom. **11**(1), 147-165, 1976.

[19] Yano K., Kon M., *Structures on manifolds* (World Scientific, Singapore, 1984).

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1001 – 1007

On the Fibonacci and Lucas numbers, their sums and permanents of one type of Hessenberg matrices

Fatih YILMAZ*† and Durmus BOZKURT‡

Abstract

At this paper, we derive some relationships between permanents of one type of lower-Hessenberg matrix family and the Fibonacci and Lucas numbers and their sums.

2000 AMS Classification: 15A36, 15A15, 11B37

Keywords: Hessenberg matrix, permanent, Fibonacci and Lucas number.

Received 05 : 10 : 2011 : Accepted 05 : 10 : 2013 Doi : 10.15672/HJMS.2014437527

1. Introduction

The well-known Fibonacci and Lucas sequences are recursively defined by

$$F_{n+1} = F_n + F_{n-1},\ n \geq 1$$
$$L_{n+1} = L_n + L_{n-1},\ n \geq 1$$

with initial conditions $F_0 = 0$, $F_1 = 1$ and $L_0 = 2$, $L_1 = 1$. The first few values of the sequences are given below:

n	0	1	2	3	4	5	6	7	8	9
F_n	0	1	1	2	3	5	8	13	21	34
L_n	2	1	3	4	7	11	18	29	47	76

The permanent of a matrix is similar to the determinant but all of the signs used in the Laplace expansion of minors are positive. The permanent of an n-square matrix is defined by

$$perA = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_i \sigma(i)$$

*Gazi University, Polatli Art and Science Faculty Department of Mathematics, 06900, Ankara, Turkey, Email: fatihyilmaz@gazi.edu.tr

†Corresponding Author.

‡Selcuk University, Science Faculty Department of Mathematics, 42250, Campus Konya, Turkey, Email: dbozkurt@selcuk.edu.tr

where the summation extends over all permutations σ of the symmetric group S_n [1].

Let $A = [a_{ij}]$ be an $m \times n$ matrix with row vectors $r_1, r_2, \ldots, r_m$. We call A is *contractible* on column k, if column k contains exactly two non zero elements. Suppose that A is contractible on column k with $a_{ik} \neq 0, a_{jk} \neq 0$ and $i \neq j$. Then the $(m-1) \times (n-1)$ matrix $A_{ij:k}$ obtained from A replacing row i with $a_{jk}r_i + a_{ik}r_j$ and deleting row j and column k is called the *contraction* of A on column k relative to rows i and j. If A is contractible on row k with $a_{ki} \neq 0, a_{kj} \neq 0$ and $i \neq j$, then the matrix $A_{k:ij} = [A^T_{ij:k}]^T$ is called the contraction of A on row k relative to columns i and j. We know that if B is a contraction of A[6], then

$$perA = perB. \tag{1.1}$$

It is known that there are a lot of relationships between determinants or permanents of matrices and well-known number sequences. For example, the authors [2] investigate relationships between permanents of one type of Hessenberg matrix and the Pell and Perrin numbers.

In [3], Lee defined a $(0-1)-$ matrix whose permanents are Lucas numbers.

In [4], the author investigate general tridiagonal matrix determinants and permanents. Also he showed that the permanent of the tridiagonal matrix based on $\{a_i\}$, $\{b_i\}$, $\{c_i\}$ is equal to the determinant of the matrix based on $\{-a_i\}$, $\{b_i\}$, $\{c_i\}$.

In [5], the authors give $(0, 1, -1)$ tridiagonal matrices whose determinants and permanents are negatively subscripted Fibonacci and Lucas numbers. Also, they give an $n \times n$ $(-1, 1)$ matrix S, such that per$A$=det$(A \circ S)$, where $A \circ S$ denotes Hadamard product of A and S.

In the present paper, we consider a particular case of lower Hessenberg matrices. We show that the permanents of this type of matrices are related with Fibonacci and Lucas numbers and their sums.

2. Determinantal representation of Fibonacci and Lucas numbers and their sums

Let $H_n = [h_{ij}]_{n\times n}$ be an n-square lower Hessenberg matrix as below:

$$H_n = [h_{ij}]_{n\times n} = \begin{cases} 2, \text{ if } i=j, \text{ for } i,j = 1,2,\ldots,n-1 \\ 1, \text{ if } j = i-2 \text{ and } i=j=n \\ (-1)^i, \text{ if } j = i+1 \\ 0, \text{ otherwise} \end{cases} \tag{2.1}$$

Then we have the following theorem.

2.1. Theorem. *Let H_n be as in (2.1), then*

$$perH_n = perH_n^{(n-2)} = F_{n+1}$$

where F_n is the nth Fibonacci number.

Proof. By definition of the matrix H_n, it can be contracted on column n. Let $H_n^{(r)}$ be the rth contraction of H_n. If $r = 1$, then

$$H_n^{(1)} = \begin{pmatrix} 2 & -1 & & & & & & \\ 0 & 2 & 1 & & & & & \\ 1 & 0 & 2 & -1 & & & & \\ & 1 & 0 & 2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \ddots & \\ & & & 1 & 0 & & 2 & (-1)^{n-2} \\ & & & & 1 & & (-1)^{n-1} & 2 \end{pmatrix}.$$

Since $H_n^{(1)}$ also can be contracted according to the last column,

$$H_n^{(2)} = \begin{pmatrix} 2 & -1 & & & & & & \\ 0 & 2 & 1 & & & & & \\ 1 & 0 & 2 & -1 & & & & \\ & 1 & 0 & 2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \ddots & \\ & & & 1 & 0 & & 2 & (-1)^{n-3} \\ & & & & 2 & & (-1)^{n-2} & 3 \end{pmatrix}.$$

Continuing this method, we obtain the rth contraction

$$H_n^{(r)} = \begin{pmatrix} 2 & -1 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r-1} \\ & & & & F_{r+1} & (-1)^r(F_{r+2}-F_{r+1}) & F_{r+2} \end{pmatrix}, \ n \text{ is even}$$

$$H_n^{(r)} = \begin{pmatrix} 2 & -1 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r} \\ & & & & F_{r+1} & (-1)^{r-1}(F_{r+2}-F_{r+1}) & F_{r+2} \end{pmatrix}, \ n \text{ is odd}$$

where $2 \leq r \leq n-4$. Hence

$$H_n^{(n-3)} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & 1 \\ F_{n-2} & (F_{n-2}-F_{n-1}) & F_{n-1} \end{pmatrix}$$

by contraction of $H_n^{(n-3)}$ on column 3,

$$H_n^{(n-2)} = \begin{pmatrix} 2 & -1 \\ F_{n-2} & F_n \end{pmatrix}.$$

By (1.1), we have $per H_n = per H_n^{(n-2)} = F_{n+1}$. □

Let $K_n = [k_{ij}]_{n\times n}$ be an n-square lower Hessenberg matrix in which the superdiagonal entries are alternating -1s and 1s starting with 1, except the first one which is -3, the main diagonal entries are 2s, except the last one which is 1, the subdiagonal entries are 0s, the lower-subdiagonal entries are 1s and otherwise 0. Clearly:

$$(2.2)\qquad K_n = \begin{pmatrix} 2 & -3 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{n-1} \\ & & & & 1 & 0 & 1 \end{pmatrix}$$

2.2. Theorem. *Let K_n be as in (2.2), then*

$$perK_n = perK_n^{(n-2)} = L_{n-2}$$

where L_n is the nth Lucas number.

Proof. By definition of the matrix K_n, it can be contracted on column n. By consecutive contraction steps, we can write down,

$$K_n^{(r)} = \begin{pmatrix} 2 & -3 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r-1} \\ & & & & F_{r+1} & (-1)^{r-2}(F_{r+2}-F_{r+1}) & F_{r+2} \end{pmatrix}, \ n \text{ is even}$$

$$K_n^{(r)} = \begin{pmatrix} 2 & -3 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r} \\ & & & & F_{r+1} & (-1)^{r-1}(F_{r+2}-F_{r+1}) & F_{r+2} \end{pmatrix}, \ n \text{ is odd}$$

for $1 \le r \le n-4$. Hence

$$K_n^{(n-3)} = \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 1 \\ F_{n-2} & F_{n-2}-F_{n-1} & F_{n-1} \end{pmatrix}$$

by contraction of $K_n^{(n-3)}$ on column 3, gives

$$K_n^{(n-2)} = \begin{pmatrix} 2 & -3 \\ F_{n-2} & F_n \end{pmatrix}.$$

By applying (1.1), we have $perK_n = perK_n^{(n-2)} = 2F_n - 3F_{n-2} = L_{n-2}$, which is desired. □

Let $M_n = [m_{ij}]_{n\times n}$ be an n-square lower Hessenberg matrix as below:

$$(2.3)\qquad M_n = [m_{ij}]_{n\times n} = \begin{cases} 2, \text{ if } i=j, \text{ for } i,j = 1,2,\ldots,n \\ 1, \ \text{if } j = i-2 \\ (-1)^i, \ \text{if } j = i+1 \\ 0, \text{ otherwise} \end{cases}$$

2.3. Theorem. *Let M_n be as in (2.3), then*

$$perM_n = perM_n^{(n-2)} = \sum_{i=0}^{n+1} F_i = F_{n+3} - 1$$

where F_n is the nth Fibonacci number.

Proof. By contraction method on column n, we have

$$M_n^{(r)} = \begin{pmatrix} 2 & -1 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^r \\ & & & & \sum_{i=0}^{r+1} F_i & (-1)^{r-1}\sum_{i=0}^{r} F_i & \sum_{i=0}^{r+2} F_i \end{pmatrix}, \ n \text{ is odd}$$

$$M_n^{(r)} = \begin{pmatrix} 2 & -1 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r-1} \\ & & & & \sum_{i=0}^{r+1} F_i & (-1)^{r-2}\sum_{i=0}^{r} F_i & \sum_{i=0}^{r+2} F_i \end{pmatrix}, \ n \text{ is even}$$

for $1 \leq r \leq n-4$. Hence

$$M_n^{(n-3)} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & 1 \\ \sum_{i=0}^{n-2} F_i & -\sum_{i=0}^{n-3} F_i & \sum_{i=0}^{n-1} F_i \end{pmatrix}$$

by contraction of $M_n^{(n-3)}$ on column 3, gives

$$M_n^{(n-2)} = \begin{pmatrix} 2 & -1 \\ \sum_{i=0}^{n-2} F_i & \sum_{i=0}^{n} F_i \end{pmatrix}.$$

By applying (1.1), we have

$$perM_n = perM_n^{(n-2)} = \sum_{i=0}^{n+1} F_i = F_{n+3} - 1$$

which is desired. □

Let $N_n = [n_{ij}]_{n\times n}$ be an n-square lower Hessenberg matrix in which the superdiagonal entries are alternating -1s and 1s starting with 1, except the first one which is -2, the main diagonal entries are 2s, except the first one is 3, the subdiagonal entries are 0s, the lower-subdiagonal entries are 1s and otherwise 0. In this content:

$$N_n = \begin{pmatrix} 3 & -2 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{n-1} \\ & & & & 1 & 0 & 2 \end{pmatrix} \tag{2.4}$$

2.4. Theorem. *Let N_n be an n-square matrix ($n \geq 2$) as in (2.4), then*

$$perN_n = perN_n^{(n-2)} = \sum_{i=0}^{n} L_i = L_{n+2} - 1$$

where L_n is the nth Lucas number.

Proof. By contraction method on column n, we have

$$N_n^{(r)} = \begin{pmatrix} 3 & -2 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^r \\ & & & & \sum_{i=0}^{r+1} F_i & (-1)^{r-1} \sum_{i=0}^{r} F_i & \sum_{i=0}^{r+2} F_i \end{pmatrix}, \ n \text{ is odd}$$

$$N_n^{(r)} = \begin{pmatrix} 3 & -2 & & & & & \\ 0 & 2 & 1 & & & & \\ 1 & 0 & 2 & -1 & & & \\ & 1 & 0 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 2 & (-1)^{r-1} \\ & & & & \sum_{i=0}^{r+1} F_i & (-1)^{r} \sum_{i=0}^{r} F_i & \sum_{i=0}^{r+2} F_i \end{pmatrix}, \ n \text{ is even}$$

for $1 \leq r \leq n-4$. Hence

$$N_n^{(n-3)} = \begin{pmatrix} 3 & -2 & 0 \\ 0 & 2 & 1 \\ \sum_{i=0}^{n-2} F_i & -\sum_{i=0}^{n-3} F_i & \sum_{i=0}^{n-1} F_i \end{pmatrix}$$

by contraction of $N_n^{(n-3)}$ on column 3, gives

$$N_n^{(n-2)} = \begin{pmatrix} 3 & -2 \\ \sum_{i=0}^{n-2} F_i & \sum_{i=0}^{n} F_i \end{pmatrix}.$$

By applying (1.1), we have

$$perN_n = perN_n^{(n-2)} = \sum_{i=0}^{n} L_i = L_{n+2} - 1$$

by the identity $F_{n-1} + F_{n+1} = L_n$. □

Acknowledgement We thank to referees for providing valuable suggestions and the careful reading.

References

[1] H. Minc, *Encyclopedia of Mathematics and Its Applications, Permanents*, Vol. 6, Addison-Wesley Publishing Company, London, 1978.

[2] F. Yılmaz, D. Bozkurt, *Hessenberg matrices and the Pell and Perrin numbers*, Journal of Number Theory, 131 (2011) 1390-1396.

[3] G. Y. Lee, *k-Lucas numbers and associated bipartite graphs*, Linear Algebra and Its Applications, 320 (2000) 51.

[4] D. H. Lehmer, *Fibonacci and related sequences in periodic tridiagonal matrices*, Fibonacci Quarterly 12 (1975) 150-158.

[5] E. Kılıç and D. Taşçı, *Negatively subscripted Fibonacci and Lucas numbers and their complex factorizations*, Ars Combinatoria 96 (2010) 275-288.

[6] R. A. Brualdi, P. M. Gibson, *Convex polyhedra of Doubly Stochastic matrices I: applications of the permanent function*, J. Combin. Theo. A 1977, 194-230.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1009 – 1015

The product of generalized superderivations on a prime superalgebra

He Yuan* and Yu Wang†

Abstract

In the paper, we extend the definition of generalized derivations to superalgebras and prove that a generalized superderivation g on a prime superalgebra A is represented as $g(x) = ax + d(x)$ for all $x \in A$, where a is an element of Q_{mr} (the maximal right ring of quotients of A) and d is a superderivation on A. Using the result we study two generalized superderivations when their product is also a generalized superderivation on a prime superalgebra A.

2000 AMS Classification: 16N60, 16W25, 16W55.

Keywords: Generalized derivation, Prime ring, Extended centroid, Generalized superderivation, Superalgebra.

Received 05 : 10 : 2011 : Accepted 09 : 10 : 2013 Doi : 10.15672/HJMS.2014437522

1. Introduction

Let R be a prime ring. According to Hvala [9] an additive mapping $g : R \to R$ is said to be a generalized derivation of R if there exists a derivation δ of R such that $g(xy) = g(x)y + x\delta(y)$ for all $x, y \in R$. In [14] Lee proved that every generalized derivation of A can be uniquely extended to Q_{mr} and there exists an element $a \in Q_{mr}$ such that $g(x) = ax + \delta(x)$ for all $x \in R$.

The study of the product of derivations in prime rings was initiated by Posner [18]. He proved that the product of two nonzero derivations can not be a derivation on a prime ring of characteristic not 2. Later a number of authors studied the problem in several ways (see [2], [4], [5], [9], [10], [12], [13], and [15]). Hvala [9] studied two generalized derivations f_1, f_2 when the product is also a generalized derivation on a prime ring R of characteristic not 2 in 1998. In 2001 Lee [13] gave a description of Hvala's Theorem without the assumption of char$R \neq 2$. In 2004 Fošner [5] extended Posner's Theorem to prime superalgebras.

*Department of Mathematics Jilin University Changchun 130012, P. R. of China and Department of Mathematics Jilin Normal University Siping 136000, P. R. of China
Email: yuanhe1983@126.com

†Department of Mathematics Shanghai Normal University Shanghai 200234, P. R. of China
Email:ywang2004@126.com

Superalgebras first appeared in physics, in the Theory of Supersymmetry, to create an algebraic structure representing the behavior of the subatomic particles known as bosons and fermions ([11]). Recently there has been a considerable authors who are interested in superalgebras. They extended many results of rings to superalgebras (see [3], [5], [6], [7], [8], [11], [16], [17] and [19]).

In Section 3, we will extend the definition of generalized derivations to superalgebras and prove that every generalized superderivation of a prime superalgebra A can be extended to Q_{mr} (the maximal right ring of quotients of A). Further, we will prove that a generalized superderivation of a prime superalgebra is a sum of a left multiplication mapping and a superderivation. Using the result we will study two generalized superderivations when their product is also a generalized superderivation on a prime superalgebra. As a result, Fošner's theorem [5, Theorem 4.1] is the special case of the main theorem of the paper.

2. preliminaries

Let Φ be a commutative ring with $\frac{1}{2} \in \Phi$. An associative algebra A over Φ is said to be an associative superalgebra if there exist two Φ-submodules A_0 and A_1 of A such that $A = A_0 \bigoplus A_1$ and $A_iA_j \subseteq A_{i+j}$, $i, j \in Z_2$. A superalgebra is called trivial if $A_1 = 0$. The elements of A_i are homogeneous of degree i and we write $|a_i| = i$ for all $a_i \in A_i$. We define $[a, b]_s = ab - (-1)^{|a||b|}ba$ for all $a, b \in A_0 \cup A_1$. Thus, $[a, b]_s = [a_0, b_0]_s + [a_1, b_0]_s + [a_0, b_1]_s + [a_1, b_1]_s$, where $a = a_0 + a_1$, $b = b_0 + b_1$ and $a_i, b_i \in A_i$ for $i = 0, 1$. It follows that $[a, b]_s = [a, b]$ if one of the elements a and b is homogeneous of degree 0. Let $k \in \{0, 1\}$. A superderivation of degree k is actually a Φ-linear mapping $d_k : A \to A$ which satisfies $d_k(A_i) \subseteq A_{k+i}$ for $i \in Z_2$ and $d_k(ab) = d_k(a)b+(-1)^{k|a|}ad_k(b)$ for all $a, b \in A_0 \cup A_1$. If $d = d_0 + d_1$, then d is a superderivation on A. For example, for $a = a_0 + a_1 \in A$ the mapping $\mathrm{ad}_s(a)(x) = [a, x]_s = [a_0, x]_s + [a_1, x]_s$ is a superderivation, which is called the inner superderivation induced by a. For a superalgebra A, we define $\sigma : A \to A$ by $(a_0 + a_1)^\sigma = a_0 - a_1$, then σ is an automorphism of A such that $\sigma^2 = 1$. On the other hand, for an algebra A, if there exists an automorphism σ of A such that $\sigma^2 = 1$, then A becomes a superalgebra $A = A_0 \bigoplus A_1$, where $A_i = \{x \in A | x^\sigma = (-1)^i x\}$, $i = 0, 1$. Clearly a superderivation d of degree 1 is a σ-derivation, i.e., it satisfies $d(ab) = d(a)b + a^\sigma d(b)$ for all $a, b \in A$. A superalgebra A is called a prime superalgebra if and only if $aAb = 0$ implies $a = 0$ or $b = 0$, where at least one of the elements a and b is homogeneous. The knowledge of superalgebras refers to [3], [5], [6], [7], [8], [16], [17] and [19].

In [17] Montaner obtained that a prime superalgebra A is not necessarily a prime algebra but a semiprime algebra. Hence one can define the maximal right ring of quotients Q_{mr} of A, and the useful properties of Q_{mr} can be found in [1]. By [1, proposition 2.5.3] σ can be uniquely extended to Q_{mr} such that $\sigma^2 = 1$. Therefore Q_{mr} is also a superalgebra. Further, we can get that Q_{mr} is a prime superalgebra.

3. the product of generalized superderivations

Firstly, we extend the definition of generalized derivations to superalgebras.

3.1. Definition. Let A be a superalgebra. For $i \in \{0, 1\}$, a Φ-linear mapping $g_i : A \to A$ is called a generalized superderivation of degree i if $g_i(A_j) \subseteq A_{i+j}$, $j \in Z_2$, and $g_i(xy) = g_i(x)y + (-1)^{i|x|}xd_i(y)$ for all $x, y \in A_0 \cup A_1$, where d_i is a superderivation of degree i on A. If $g = g_0 + g_1$, then g is called a generalized superderivation on A.

Let A be a prime superalgebra and $Q = Q_{mr}$ be the maximal right ring of quotients of A. Next, we prove that a generalized superderivation of a prime superalgebra is a sum of a left multiplication mapping and a superderivation. By [20, proposition 2] we have every σ-derivation d of a semiprime ring A can be uniquely extended to a σ-derivation of Q.

3.2. Theorem. *Let A be a prime superalgebra and $g : A \to A$ be a generalized superderivation. Then g can be extended to Q and there exist an element $a \in Q$ and a superderivation d of A such that $g(x) = ax + d(x)$ for all $x \in A$, where both a and d are determined by g uniquely.*

Proof. To prove that the generalized superderivation g on a prime superalgebra A can be extended to Q, it suffices to prove that g_0 and g_1 can be extended to Q, respectively. The generalized superderivation of degree 1 g_1 is represented as $g_1(xy) = g_1(x)y + x^\sigma d_1(y)$ for all $x, y \in A$, where d_1 is a superderivation of degree 1 on A. Note that $d_1(xy) = d_1(x)y + x^\sigma d_1(y)$. So combining the two equations we have $(g_1 - d_1)(xy) = (g_1 - d_1)(x)y$. Let $g_1 - d_1 = f$. Clearly f is a right A-module mapping. Then there exists $a_1 \in Q$ such that $f(x) = a_1 x$. So $g_1(x) = a_1 x + d_1(x)$ for all $x \in A$. Since d_1 can be extended to Q, then it follows that g_1 can be extended to Q. It is easy to prove that $g_0(x) = a_0 x + d_0(x)$ and g_0 can be extended to Q similarly, where a_0 is an element of Q and d_0 is a superderivation of degree 0 on A. So g can be extended to Q. Clearly $a_i \in Q_i$, $i \in \{0, 1\}$. Let $a = a_0 + a_1$ and $d = d_0 + d_1$. Then $g(x) = g_0(x) + g_1(x) = a_0 x + d_0(x) + a_1 x + d_1(x) = ax + d(x)$ for all $x \in A$, where a is an element of Q and d is a superderivation of A.

Now we claim both a and d are determined by g uniquely. It suffices to prove that $a = 0$ and $d = 0$ when $g = 0$. Since $g = 0$, we have $g_0 = g_1 = 0$. By $g_1 = 0$, we obtain $0 = g_1(yr) = a_1 yr + d_1(yr) = a_1 yr + d_1(y)r + y^\sigma d_1(r) = g_1(y)r + y^\sigma d_1(r) = y^\sigma d_1(r)$ for all $y, r \in A$. Then $A^\sigma d_1(A) = 0$. So $Ad_1(A) = 0$. Clearly $d_1(A) = 0$. Since $g_1(A) = 0$, it follows that $a_1 A = 0$. Hence $a_1 = 0$. Similarly we can prove the case when $g_0 = 0$. So $a = 0$ and $d = 0$. □

Next, we give two results which are used in the proof of the main result.

3.3. Lemma. *Let A be a prime superalgebra. If A satisfies*

(3.1) $$([a_0, x] + d_0(x))yk_0(z) + ([b_0, x] + k_0(x))yd_0(z) = 0 \qquad \text{for all } x, y, z \in A,$$

where a_0, $b_0 \in Q_0$ and both d_0 and k_0 are superderivations of degree 0 on A. Then one of the following cases is true:

(i) There exists $0 \neq \mu \in C_0$ such that $\mu k_0(x) + d_0(x) = 0$;

(ii) $[a_0, x] + d_0(x) = 0$;

(iii) $[b_0, x] + k_0(x) = 0$

for all $x \in A$.

Proof. Let $d_0 = k_0 = 0$. Clearly there exists $0 \neq \mu \in C_0$ such that $\mu k_0(x) + d_0(x) = 0$. Hence (i) is true.

Next we assume either $d_0 \neq 0$ or $k_0 \neq 0$. By [5, Theorem 3.3] there exist λ_1 and λ_2 not all zero such that $\lambda_1([a_0, x] + d_0(x)) + \lambda_2([b_0, x] + k_0(x)) = 0$. Let $\lambda_1 = \lambda_{10} + \lambda_{11}$ and $\lambda_2 = \lambda_{20} + \lambda_{21}$. Then $\lambda_{10}([a_0, x] + d_0(x)) + \lambda_{11}([a_0, x] + d_0(x)) + \lambda_{20}([b_0, x] + k_0(x)) + \lambda_{21}([b_0, x] + k_0(x)) = 0$ for all $x \in A$, where $\lambda_{10}, \lambda_{20} \in C_0$, $\lambda_{11}, \lambda_{21} \in C_1$. By $A_0 \cap A_1 = 0$, we have

(3.2) $$\lambda_{11}([a_0, x_0] + d_0(x_0)) + \lambda_{21}([b_0, x_0] + k_0(x_0)) = 0 \qquad \text{for all } x_0 \in A_0,$$

(3.3) $$\lambda_{11}([a_0, x_1] + d_0(x_1)) + \lambda_{21}([b_0, x_1] + k_0(x_1)) = 0 \qquad \text{for all } x_1 \in A_1.$$

Using (3.2) and (3.3) we obtain

(3.4) $$\lambda_{11}([a_0, x] + d_0(x)) + \lambda_{21}([b_0, x] + k_0(x)) = 0 \qquad \text{for all } x \in A.$$

We proceed by dividing three cases. Only one of λ_{11} and λ_{21} is nonzero. If $\lambda_{21} \neq 0$, then $[b_0, x] + k_0(x) = 0$. If $\lambda_{11} \neq 0$, then $[a_0, x] + d_0(x) = 0$. Hence either (ii) or (iii) is true.

Both $\lambda_{11} \neq 0$ and $\lambda_{21} \neq 0$. By (3.4) and [5, Lemma 3.1] we arrive at $[a_0, x] + d_0(x) = \lambda([b_0, x] + k_0(x))$, where $\lambda = -\lambda_{11}^{-1}\lambda_{21} \neq 0$. Using (3.1) we get $\lambda([b_0, x] + k_0(x))yk_0(z) + ([b_0, x] + k_0(x))yd_0(z) = 0$. That is, $([b_0, x] + k_0(x))y(\lambda k_0(z) + d_0(z)) = 0$. If there exists $z \in A$ such that $\lambda k_0(z) + d_0(z) \neq 0$, then $[b_0, x] + k_0(x) = 0$ for all $x \in A_0 \cup A_1$. It follows that $[b_0, x] + k_0(x) = 0$ for all $x \in A$. Hence either (i) or (iii) is true. Similarly, when $\rho([a_0, x] + d_0(x)) = [b_0, x] + k_0(x)$, where $\rho = -\lambda_{21}^{-1}\lambda_{11} \neq 0$, we have either (i) or (ii) is true by using (3.1) again.

When $\lambda_{11} = \lambda_{21} = 0$, i.e., $\lambda_1, \lambda_2 \in C_0$. If one of λ_1 and λ_2 is zero, then either (ii) or (iii) is true. If both λ_1 and λ_2 are nonzero, the proof is similar to the above paragraph. □

Similar to the proof of Lemma 3.3, we can get the following result.

3.4. Lemma. *Let A be a prime superalgebra. If A satisfies*

$$([a_1, x]_s + d_1(x))yk_1(z) - ([b_1, x]_s + k_1(x))yd_1(z) = 0 \qquad \text{for all } x, y, z \in A,$$

where a_1, $b_1 \in Q_1$ and both d_1 and k_1 are superderivations of degree 1 on A. Then one of the following cases is true:

(i) There exists $0 \neq \nu \in C_0$ such that $\nu k_1(x) + d_1(x) = 0$;

(ii) $[a_1, x]_s + d_1(x) = 0$;

(iii) $[b_1, x]_s + k_1(x) = 0$

for all $x \in A$.

Now, we are in a position to give the main result of this paper.

3.5. Theorem. *Let A be a prime superalgebra and let $f = a + d$ and $g = b + k$ be two nonzero generalized superderivations on A, where $a, b \in Q$ and both d and k are superderivations on A. If fg is also a generalized superderivation on A. Then one of the following cases is true:*

(i) There exists $0 \neq \omega \in C_0$ such that $\omega k_j(x) + d_j(x) = 0$;

(ii) $[a_i, x]_s + d_i(x) = 0$;

(iii) $[b_i, x]_s + k_i(x) = 0$

for all $x \in A$, where $i, j \in \{0, 1\}$, $a_i, b_i \in Q_i$ and both d_i and k_i are superderivations of degree i on A, as well as d_j and k_j.

Proof. According to Theorem 3.2 we assume $h(x) = fg(x) = cx + l(x)$ for all $x \in A$, where $c \in Q$ and l is a superderivation on A, then

$$\begin{aligned} fg(x) =& a(bx + k(x)) + d(bx + k(x)) \\ =& abx + ak(x) + d_0(b)x + bd_0(x) + d_1(b)x + b^\sigma d_1(x) + d_0k(x) + d_1k(x). \end{aligned}$$

Hence

$$\begin{aligned} c =& ab + d_0(b) + d_1(b) = ab + d(b), \\ l(x) =& ak(x) + bd_0(x) + b^\sigma d_1(x) + d_0k(x) + d_1k(x), \\ l_0(x) =& a_0k_0(x) + a_1k_1(x) + b_0d_0(x) - b_1d_1(x) + d_0k_0(x) + d_1k_1(x), \\ l_1(x) =& a_1k_0(x) + a_0k_1(x) + b_1d_0(x) + b_0d_1(x) + d_0k_1(x) + d_1k_0(x). \end{aligned}$$

On the one hand we get

$$\begin{aligned}l_0(xy) =&a_0k_0(xy) + a_1k_1(xy) + b_0d_0(xy) - b_1d_1(xy) + d_0k_0(xy) + d_1k_1(xy)\\ =&a_0k_0(x)y + a_0xk_0(y) + a_1k_1(x)y + a_1x^\sigma k_1(y)\\ &+ b_0d_0(x)y + b_0xd_0(y) - b_1d_1(x)y - b_1x^\sigma d_1(y)\\ &+ d_0k_0(x)y + k_0(x)d_0(y) + d_0(x)k_0(y) + xd_0k_0(y)\\ &+ d_1k_1(x)y + k_1(x)^\sigma d_1(y) + d_1(x^\sigma)k_1(y) + xd_1k_1(y)\end{aligned}$$

and on the other hand we get

$$\begin{aligned}l_0(xy) =&a_0k_0(x)y + a_1k_1(x)y + b_0d_0(x)y - b_1d_1(x)y + d_0k_0(x)y + d_1k_1(x)y\\ &+ x[a_0k_0(y) + a_1k_1(y) + b_0d_0(y) - b_1d_1(y) + d_0k_0(y) + d_1k_1(y)].\end{aligned}$$

Combining the two equations we have

$$\begin{aligned}0 =&[a_0, x]k_0(y) + a_1x^\sigma k_1(y) - xa_1k_1(y) + [b_0, x]d_0(y) - b_1x^\sigma d_1(y)\\ &+ xb_1d_1(y) + k_0(x)d_0(y) + d_0(x)k_0(y) + k_1(x)^\sigma d_1(y) - d_1(x)^\sigma k_1(y).\end{aligned} \tag{3.5}$$

In particular, replacing y by yz in (3.5) we get

$$\begin{aligned}0 =&[a_0, x]k_0(yz) + a_1x^\sigma k_1(yz) - xa_1k_1(yz) + [b_0, x]d_0(yz) - b_1x^\sigma d_1(yz)\\ &+ xb_1d_1(yz) + k_0(x)d_0(yz) + d_0(x)k_0(yz) + k_1(x)^\sigma d_1(yz) - d_1(x)^\sigma k_1(yz).\end{aligned}$$

Extending the identity above we arrive at

$$\begin{aligned}0 =&[a_0, x]k_0(y)z + [a_0, x]yk_0(z) + a_1x^\sigma k_1(y)z + a_1x^\sigma y^\sigma k_1(z)\\ &- xa_1k_1(y)z - xa_1y^\sigma k_1(z) + [b_0, x]d_0(y)z + [b_0, x]yd_0(z)\\ &- b_1x^\sigma d_1(y)z - b_1x^\sigma y^\sigma d_1(z) + xb_1d_1(y)z + xb_1y^\sigma d_1(z)\\ &+ k_0(x)d_0(y)z + k_0(x)yd_0(z) + d_0(x)k_0(y)z + d_0(x)yk_0(z)\\ &+ k_1(x)^\sigma d_1(y)z + k_1(x)^\sigma y^\sigma d_1(z) - d_1(x)^\sigma k_1(y)z - d_1(x)^\sigma y^\sigma k_1(z).\end{aligned}$$

Using (3.5) we have

$$\begin{aligned}0 =&[a_0, x]yk_0(z) + a_1x^\sigma y^\sigma k_1(z) - xa_1y^\sigma k_1(z) + [b_0, x]yd_0(z) - b_1x^\sigma y^\sigma d_1(z)\\ &+ xb_1y^\sigma d_1(z) + k_0(x)yd_0(z) + d_0(x)yk_0(z) + k_1(x)^\sigma y^\sigma d_1(z) - d_1(x)^\sigma y^\sigma k_1(z).\end{aligned}$$

[5, Corollary 3.6] gives

$$p_{ij} = [a_0, x_i]yk_0(z_j) + [b_0, x_i]yd_0(z_j) + k_0(x_i)yd_0(z_j) + d_0(x_i)yk_0(z_j) = 0, \tag{3.6}$$

$$\begin{aligned}q_{ij} =&a_1x_i^\sigma yk_1(z_j) - x_ia_1yk_1(z_j) - b_1x_i^\sigma yd_1(z_j) + x_ib_1yd_1(z_j)\\ &+ k_1(x_i)^\sigma yd_1(z_j) - d_1(x_i)^\sigma yk_1(z_j) = 0.\end{aligned} \tag{3.7}$$

for all $x_i \in A_i$, $y \in A$, $z_j \in A_j$, $i, j \in \{0, 1\}$. Therefore

$$\begin{aligned}p_{00} + p_{01} + p_{10} + p_{11} =&[a_0, x]yk_0(z) + [b_0, x]yd_0(z) + k_0(x)yd_0(z)\\ &+ d_0(x)yk_0(z) = 0,\end{aligned} \tag{3.8}$$

$$\begin{aligned}q_{00} + q_{01} + q_{10} + q_{11} =&a_1x^\sigma yk_1(z) - xa_1yk_1(z) - b_1x^\sigma yd_1(z) + xb_1yd_1(z)\\ &+ k_1(x)^\sigma yd_1(z) - d_1(x)^\sigma yk_1(z) = 0.\end{aligned} \tag{3.9}$$

According to (3.8) and Lemma 3.3 we see that either (i) or (ii) or (iii) is true.

By (3.9) we get

$$\begin{aligned}&[a_1, x_0]yk_1(z) - [b_1, x_0]yd_1(z) - k_1(x_0)yd_1(z)\\ &\qquad + d_1(x_0)yk_1(z) = 0 \text{ for all } x_0 \in A_0, y, z \in A,\end{aligned}$$

$$-[a_1,x_1]_s y k_1(z) + [b_1,x_1]_s y d_1(z) + k_1(x_1) y d_1(z) - d_1(x_1) y k_1(z) = 0 \text{ for all } x_1 \in A_1, y, z \in A.$$

Combining the identities above we give

$$[a_1,x]_s y k_1(z) - [b_1,x]_s y d_1(z) - k_1(x) y d_1(z) + d_1(x) y k_1(z) = 0 \text{ for all } x, y, z \in A.$$

By Lemma 3.4 we have that either (i) or (ii) or (iii) is true. Similarly, using the same way to $l_1(xy)$ we have

$$[a_0,x] y k_1(z) + [b_0,x] y d_1(z) + k_0(x) y d_1(z) + d_0(x) y k_1(z) = 0,$$

$$(3.10) \qquad \begin{aligned} & a_1 x y k_0(z) - x^\sigma a_1 y k_0(z) + b_1 x y d_0(z) \\ & - x^\sigma b_1 y d_0(z) + k_1(x) y d_0(z) + d_1(x) y k_0(z) = 0 \end{aligned}$$

and either (i) or (ii) or (iii) is true. □

In particular, taking $a = b = 0$ in Theorem 3.5 we obtain

3.6. Corollary. ([5, Theorem 4.1]) *Let A be a prime associative superalgebra and let $d = d_0 + d_1$ and $k = k_0 + k_1$ be nonzero superderivations on A. Then dk is a superderivation if and only if $d_0 = k_0 = 0$ and $k_1 = \lambda_0 d_1$ for some nonzero $\lambda_0 \in C_0$.*

Proof. We assume that both d_0 and k_0 are nonzero. Since d and k are nonzero superderivations and dk is also a superderivation of A, then there exists $0 \neq \mu \in C_0$ such that $k_0(x) = \mu d_0(x)$ by Theorem 3.5. We have $2\mu d_0(x) y d_0(x) = 0$ by taking $z = x$ in (3.8), that is, $d_0(x) A d_0(x) = 0$. Since A is a semiprime algebra, then $d_0(x) = 0$. But it contradicts $d_0 \neq 0$. We set $d_0 = 0$. Then $d_1 \neq 0$. When $k_1 \neq 0$. There exists $0 \neq \lambda_0 \in C_0$ such that $k_1(x) = \lambda_0 d_1(x)$ and $k_0(x) = d_0(x) = 0$ by Theorem 3.5. When $k_1 = 0$ and $k_0 \neq 0$, we have $d_1(x) = 0$ by (3.10). It contradicts that d is a nonzero superderivation. So $d_0 = k_0 = 0$ and $k_1 = \lambda_0 d_1$ for some nonzero $\lambda_0 \in C_0$ when dk is a superderivation. It is easy to prove that dk is a superderivation when $d_0 = k_0 = 0$ and $k_1 = \lambda_0 d_1$ for some nonzero $\lambda_0 \in C_0$ □

References

[1] Beidar, K. I., Martindale 3rd, W. S. and Mikhalev, A. V. *Rings with Generalized Identities*, Marccl Dekker, New York, 1996.

[2] Chebotar, M. A. and Lee, P. H. *A Note on Compositions of Derivations of Prime Rings*, Comm. Algebra **31**(6), 2965-2969, 2003.

[3] Chen, T. S. *Supercentralizing Superderivations on Prime Superalgebras*, Comm. Algebra **33**(12), 4457-4466, 2005.

[4] Chuang, C. L. *On Compositions of Derivations of Prime Rings*, Proc. Amer. Math. Soc. **108**(3), 647-652, 1990.

[5] Fošner, M. *On the Extended Centroid of Prime Associative Superalgebras with Applications to Superderivations*, Comm. Algebra **32**(2), 689-705, 2004.

[6] Fošner, M. *Jordan ϵ-homomorphisms and Jordan ϵ-derivations*, Taiwanese J. Math. **9**(4), 595-616, 2005.

[7] Fošner, M. *Jordan superderivations*, Comm. Algebra **31**(9), 4533-4545, 2003.

[8] Fošner, M. *Jordan superderivations II*, Int. J. Math. Math. Sci. **2004**(44), 2357-2369, 2004.

[9] Hvala, B. *Generalized Derivations in Rings*, Comm. Algebra **26**(4), 1147-1168, 1998.

[10] Jensen, D. *Nilpotency of Derivations in Prime Rings*, Proc. Amer. Math. Soc. **123**(9), 2633-2636, 1995.

[11] Laliena, J. and Sacristán, S. *On Certain Semiprime Associative Superalgebras*, Comm. Algebra **37**(10), 3548-3552, 2009.

[12] Lanski, C. *Differential Identities, Lie Ideals, and Posner's Theorems*, Pacific J. Math. **134**(2), 275-297, 1988.

[13] Lee, T. K. *Identities with Generalized Derivations*, Comm. Algebra **29**(10), 4451-4460, 2001.

[14] Lee, T. K. *Generalized Derivations of Left Faithful Rings*, Comm. Algebra **27**(8), 4057-4073, 1999.

[15] Lee, T. K. and Liu, C. K. *Nilpotency of q-skew Derivations*, Comm. Algebra **34**(2), 661-669, 2006.

[16] Lee, P. H. and Wang, Y. *Supercentralizing Maps in Prime Superalgebras*, Comm. Algebra **37**(3), 840-854, 2009.

[17] Montaner, F. *On the Lie Structure of Associative Superalgebras*, Comm. Algebra **26**(7), 2337-2349, 1998.

[18] Posner, E. C. *Derivations in Prime Rings*, Proc. Amer. Math. Soc. **8**, 79-92, 1957.

[19] Wang, Y. *Supercentralizing Superautomorphisms on Prime Superalgebras*, Taiwanese J. Math. **13**(5), 1441-1449, 2009.

[20] Wei, F. and Niu, F. W. *Extension of (1,σ) Derivations of Semiprime Rings*, Acta Mathematica Sinica (PRC) **44**(1), 169-174, 2001.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1017 – 1034

Fuzzy approach of group sequential test for binomial case

Sevil Bacanlı* , Duygu Icen† and Salih Emri‡

Abstract

The aim of this study is to present the fuzzy statistics into group sequential test when response variable has binomial case. Confidence intervals for fuzzy parameter estimation in group sequential test procedure is applied to construct the related fuzzy test statistic with the help of Buckley's approach with r-cuts. Afterwards, this present study is completed with a numerical application to real data. Finally it is concluded that the fuzzy approach is also applicable for group sequential tests when response variable has binomial case.

2000 AMS Classification: 62L05, 62L86.

Keywords: Group sequential test, fuzzy statistics, asthma prevalence, Turkey.

Received 30 : 07 : 2013 : Accepted 11 : 02 : 2014 Doi : 10.15672/HJMS.201457456

1. Introduction

One of the most important problems in medicine is the uncertainty between patients and medical relations. These relations are considered as inexact medical entities [2,3]. According to fuzzy set theory suggested by Zadeh [35], inexact medical entities can be defined as fuzzy sets. Theory of fuzzy sets is widely used for solving problems in which parameter or quantities cannot be expressed precisely. Buckley [6,7,8] introduced an approach that uses a set of confidence intervals. Furthermore, fuzzy sets present a number of powerful reasoning methods that can handle approximate inferences for medical data [9,19]. Several authors have proposed fuzzy approaches for medical researches. Reis [26] proposed a fuzzy expert model. This model could be used as a teaching or training tool that helps midwives, residents and medical students to identify and evaluate clinical risk factors. Duarte [11] tested a model to select patients for myocardial perfusion

*Hacettepe University, Department of Statistics, 06800, Ankara, Turkey.
Email: sevil@hacettepe.edu.tr
†Hacettepe University, Department of Statistics, 06800, Ankara, Turkey.
Email:duyguicn@hacettepe.edu.tr
‡Hacettepe University, Department of Chest Diseases, 06100, Ankara, Turkey.
Email:semri.editor@gmail.com

scintigraphy (MPS) based on fuzzy sets. Zolnoori [33,34] developed a fuzzy expert system for prediction of fatal asthma and evaluation of the level in asthma exacerbation.

Group sequential tests are not only used in clinical trials but also in medical studies due to their ethical, economical and administrative benefits. There is an extensive literature on group sequential tests and their application in clinical trials: an excellent summary is provided in Jennison and Turmbull [15]. As for medical studies, Pasternak and Shoe [23] demonstrated that the group sequential test had generally higher efficiency in a cohort study. Satagoban et al. [29] explain the use of a two stage group sequential test for gene-disease association studies. Aplenc et al. [4] give a description of group sequential test for molecular epidemiology study.

Group sequential tests have been applied to the normal, Binomial, inverse Gaussian and survival response variables [15,5]. Various group sequential test procedures have been suggested to analyze accumulated data in the literature [15]. Originally they were defined on the basis of a normalized Z statistic [24], or partial sum statistic [21,32]. Later, Kim and DeMets [16], Lan and DeMets [18], Pampallona and Anastasios [22], Chang, Hwang and Shih [10] proposed their designs based on error spending functions. Any of these procedures can have the overall type I (α) and type II (β) error while providing an opportunity for early stopping critical values [28].

Asthma is a chronic inflammatory disorder of the respiratory tract characterized by the infiltration of inflammatory cells, including mast cells, eosinophils, and lymphocytes [13,14,27]. It is a major cause of disability, utilization of health resource and poor quality of life around the world. In addition, asthma is the most common chronic disease among children and young adults. It causes considerable health care costs and loss of work productivity [31].

There is an epidemic of asthma affecting approximately 4% to 5% of people in developed countries. In United States, 20.1 million individuals are affected due to asthma and 6.3 million of them are children [1,30]. Emri [13] researched asthma prevalence in five urban regions in Turkey. It is found that the asthma prevalence 6.6%. After that, Kurt [17] evaluated the prevalence of risk factors for asthma and allergic diseases in Turkey.

In this study, it is indicated that group sequential test with $\alpha^*(t)$ functions for binomial response is applied to asthma data under the light of fuzzy approach. In many cases of real life, most of the data are approximately known. In addition to this, effects of measurement errors or unrecognized interactions are inevitable in every field of science. That is why, we use Buckley's fuzzy approach for estimating the asthma prevalence. Subsequently, fuzzy approach for group sequential test is applied to asthma prevalence in five Turkish urban regions. More information is used in the process of estimation and hypothesis testing with Buckley's approach than classical approach.

This paper is organized as follows; The definitions of fuzzy sets, triangular shaped fuzzy numbers, r-cut of triangular shaped fuzzy number and fuzzy probability are explained Section 2.1. Later, Buckley's approach for hypothesis testing is briefly reviewed in Section 2.2. Group sequential test based on $\alpha^*(t)$ spending functions and the adaptation of group sequential test according to Buckley's approach for a binomial case are given in Section 2.3. An illustrative example of the application of the fuzzy group sequential test to real asthma data from five Turkish urban regions is given in Section 3. Finally, concluding remarks are summarized in Section 4.

2. Theory and Methods

2.1. Fuzzy Sets and Triangular Shaped Fuzzy Numbers. A class of objects whose boundaries are not sharply defined is called as a fuzzy set. If $X = \{x\}$ denote a collection of objects, a fuzzy set $\widetilde{N}$ in X is a set of ordered pairs $\widetilde{N} = \{x, \mu_{\widetilde{N}}(x)\}, x \in X$ where $\mu_{\widetilde{N}}$ is the grade of membership of x in $\widetilde{N}$, $\mu_{\widetilde{N}}(x) : X \to M$ is a function from X to membership space M and produces values in $[0, 1]$ for all x. Hence the degree of membership of x in $\widetilde{N}$ is represented by $\mu_{\widetilde{N}}(x)$ which is a function having values between 0 and 1 [12].

The r-cuts of a fuzzy number, slices through a fuzzy number, is a non-fuzzy set defined as $\widetilde{N}(r) = \{x \in R, \mu_{\widetilde{N}}(x) \geq r\}$. Hence r-cut of a triangular shaped fuzzy number can be shown as $\widetilde{N}(r) = [N^L(r), N^U(r)]$, where $N^L(r)$ is the minimum value and $N^U(r)$ is the maximum value of the r-cut [12].

2.2. Hypothesis Testing using Buckley's Approach with r-cuts. One of the primary purposes of this statistical inference is to test the hypothesis. The problem of testing a hypothesis may be about the decision, since the decisions have to be made about the truth of two propositions, the null hypothesis H_0 and the alternative H_1. Furthermore, in traditional statistics, all parameters of the mathematical model should be very well defined. Sometimes these assumptions may appear too rigid for the real-life problems, especially dealing with imprecise requirements in medical studies. To lessen this rigidity, fuzzy methods are incorporated into statistics. In this section, Buckley's [6,7,8] approach for hypothesis testing that the parameter of crisp binomial distribution is defined as a triangular fuzzy number is summarized.

Let P be the probability of a success so that $Q = 1 - P$ is the probability of a failure. It is obtained x successes in a random sample size n so $p = x/n$ is the point estimate of P. The classical hypothesis for binomial distribution is defined as $H_0 : P = P_0$ versus $H_1 : P \neq P_0$. The test statistic

$$Z_0 = \frac{p - P_0}{\sqrt{P_0\, Q_0/n}} \tag{2.1}$$

is approximately standard normal distribution if n is sufficiently large. Then, decision rule is: (1) reject H_0 if $Z_0 \geq z_{\alpha/2}$ or $Z_0 \leq -z_{\alpha/2}$; and (2) do not reject H_0 when $-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$. In the above decision rule $\pm z_{\alpha/2}$ are called critical values (CV) for the test. In the decision rule $z_{\alpha/2}$ is the z value so that probability of random variable having the $N(0, 1)$ probability density, exceeding z is $\alpha/2$.

It is known that $(p - P)/\sqrt{PQ/n}$ is approximately $N(0, 1)$ if n is sufficiently large. At that case

$$P\left(p - z_{\alpha/2}\sqrt{p\,q/n} \leq\ P\ \leq p + z_{\alpha/2}\sqrt{p\,q/n}\right) = (1 - \alpha). \tag{2.2}$$

This interval can be arranged according to the method proposed by Buckley [7,8] with substituting $(1 - \alpha)100\%$ confidence interval for all $0.01 \leq \alpha \leq 1$. So equation (2.2) is defined by the following equation,

$$[p^L(\alpha), p^U(\alpha)] = [p - z_{\alpha/2}\sqrt{pq/n}, p + z_{\alpha/2}\sqrt{pq/n}]. \tag{2.3}$$

By placing these confidence intervals one after the other, a triangular shaped fuzzy number $\widetilde{p}$ whose r-cuts are the confidence intervals as

$$\widetilde{p}[r] = [p^L(r), p^U(r)], \tag{2.4}$$

is given

$$\widetilde{p}[r] = [p - z_{r/2}\sqrt{pq/n}, p + z_{r/2}\sqrt{pq/n}] \tag{2.5}$$

for $0.01 \leq r \leq 1$. Hence the fuzzy parameter estimation of P as triangular shaped fuzzy number is obtained.

By substituting equation (2.5) for p into equation (2.1), r-cuts of fuzzy test statistic are obtained as

$$\widetilde{Z}[r] = \frac{\widetilde{p}[r]-P_0}{\sqrt{P_0Q_0/n}} \tag{2.6}$$

$$= \left[Z_0 - z_{r/2}\sqrt{\frac{pq}{P_0Q_0}}, Z_0 + z_{r/2}\sqrt{\frac{pq}{P_0Q_0}}\right].$$

Each r-cut is put one over the other, in order to get a triangular fuzzy test statistic $\widetilde{Z}[r]$. Calculations are performed by r-cuts and interval arithmetic. Since test statistic is fuzzy, the critical values $\widetilde{CV}_i, i = 1, 2$, which are given with equation (2.7) and equation (2.8), will also be fuzzy. Let $\widetilde{CV}_1$ correspond to $-z_{\gamma/2}$; and let $\widetilde{CV}_2$ go with $z_{\gamma/2}$, in this way it is possible to write $\widetilde{CV}_1 = -\widetilde{CV}_2$.

$$\widetilde{CV}_2[r] = \left[z_{\alpha/2} - z_{r/2}\sqrt{\frac{p\,q}{P_0Q_0}}; z_{\alpha/2} + z_{r/2}\sqrt{\frac{p\,q}{P_0Q_0}}\right] \tag{2.7}$$

$$\widetilde{CV}_1[r] = \left[-z_{\alpha/2} - z_{r/2}\sqrt{\frac{p\,q}{P_0Q_0}}; -z_{\alpha/2} + z_{r/2}\sqrt{\frac{p\,q}{P_0Q_0}}\right] \tag{2.8}$$

Both $\widetilde{CV}_1$ and $\widetilde{CV}_2$ are triangular shaped fuzzy numbers. In addition to this, r ranges in the interval $[0.01, 1]$. Final decision rule depends on the positions of fuzzy critical values: (1) $\widetilde{CV}_2 < \widetilde{Z}$ reject H_0; (2) $\widetilde{CV}_1 < \widetilde{Z} \approx \widetilde{CV}_2$ no decision; (3) $\widetilde{CV}_1 < \widetilde{Z} < \widetilde{CV}_2$ do not reject H_0; (4) $\widetilde{CV}_1 \approx \widetilde{Z} < \widetilde{CV}_2$ no decision; (5) $\widetilde{Z} < \widetilde{CV}_1$ reject H_0 [6,7,8].

2.3. Group sequential test for a binomial case using Buckley's approach with r-cuts. Several authors have proposed group sequential tests according to the significance levels:(i) constant levels for Pocock [24] and (ii) slowly increasing levels for O'Brien and Flemmnig [21]. These tests can be used when the group sizes are equal. In 1980s, first generation methods were generalized by Kim and DeMets [16] with the $\alpha^*(t)$ which allows one to characterize the rate at which the α risk is spent. The time t is the so-called information fraction in which the information is observed at a given time and divided by the total information which is at the end of the study. For example: $t = n/N$ can be given as a quantitative endpoint which represents the division of the number of patients at a given time with the number of patients at the end of the study [30]. In this group sequential test, it is determined that a discrete sequential critical value $(c_1, c_2, ..., c_K)$ is constructed by choosing positive constants $\alpha_1, ..., \alpha_K$ so $\sum \alpha_i = \alpha$, $P(Z_1 \geq c_1) = \alpha_1$ and $i = 2, ..., K$, $P(Z_i \geq c_i, Z_j \leq c_j, j = 1, ..., i-1) = \alpha_i$ [18]. Several examples of functions are existed in the literature. In this study $\alpha_i^*(t)$'s $(i = 1, 2, ..., 5)$ are used as follows;

1. $\alpha_1^*(t) = 2[1 - \varphi(Z_{1-\alpha/2}/\sqrt{t})]$ $\quad 0 \leq t \leq 1$
2. $\alpha_2^*(t) = \alpha[ln[1 + (e-1)t]]$ $\quad 0 \leq t \leq 1$
3. $\alpha_3^*(t) = \alpha\, t$ $\quad 0 \leq t \leq 1$
4. $\alpha_4^*(t) = \alpha\, t^{3/2}$ $\quad 0 \leq t \leq 1$
5. $\alpha_5^*(t) = \alpha\, t^2$ $\quad 0 \leq t \leq 1.$

When the group sizes are equal, it generates $\alpha_1^*(t)$, discrete c_k approximate to those of O'Brien and Fleming [21] and it generates $\alpha_2^*(t)$, c_k approximate to those of Pocock [24]. Reboussin et al. [25] introduced a program to perform computations related to the design and analysis of group sequential clinical tests using [16] spending functions. The program and detailed information are publicly available at `http://www.biostat.wisc.edu/landemets`[25].

Firstly, it is considered that the primary outcome of group sequential test is binary. A sequence of independent Bernoulli random variables $X_1, X_2, \ldots$ is taken into consideration with $P(X_i = 1) = P$ and $P(X_i = 0) = 1 - P$. If data are divided with the total numbers of observations $n_1, n_2, \ldots, n_K$ as group sequentially in analysis 1 to K, the usual estimate of P in analysis k is given as such:

$$(2.9) \qquad p^{(k)} = \frac{1}{n_k} \sum_{i=1}^{n_k} X_i$$

which has variance $P(1-P)/n_k$ and expectation P. The standardized statistics which are given in equation (2.10), may be used for constructing a two sided test

$$(2.10) \qquad Z_k = (p^{(k)} - P_0)\sqrt{I_k} \quad , k = 1, \ldots, K$$

in which $I_k = n_k/\{P_0(1-P_0)\}$. The test statistic Z_k, is compared with c_k as follows; **1.** After group $k = 1, 2, \ldots, K-1$, if $Z_k \geq c_k$ stop reject H_0, otherwise continue to $k+1$. **2.** After group K, if $Z_K \geq c_K$ stop reject H_0, otherwise accept H_0.

After the classical approach of group sequential test is reviewed, let us proceed to fuzzy approach in which the estimate of P is a triangular shaped fuzzy number and its r-cuts are given with equation (2.5).

In order to perform $H_0 : P = P_0$ versus $H_1 : P > P_0$, equation (2.9) is calculated in every step of group sequential test. The uncertainty of this parameter is taken into account during the process and is taken as triangular shaped fuzzy number. Hence for each stage of the process in group sequential tests, fuzzy parameter estimation of $p^{(k)}$ is calculated with equation (2.11) for $0.01 \leq r \leq 1$.

$$(2.11) \qquad \widetilde{p}^{(k)}[r] = \left[p^{(k)} - z_{r/2}\sqrt{\frac{p^{(k)}\ q^{(k)}}{n_k}};\ p^{(k)} + z_{r/2}\sqrt{\frac{p^{(k)}\ q^{(k)}}{n_k}}\right]$$

Better results can be attained with fuzzy approach which considers all confidence intervals as ($\widetilde{p}^{(k)}[r]$, $0.01 \leq r \leq 1$) for unknown parameter $p^{(k)}$ in the process of group sequential test rather than classical approach for unknown parameter $p^{(k)}$. Calculations are performed with interval arithmetic. Substituting r-cuts of $\widetilde{p}^{(k)}$ into the equation (2.10) makes it possible to simplify by using interval arithmetic to produce $\widetilde{Z}_k[r]$ which is given below

$$(2.12) \qquad \begin{aligned} \widetilde{Z}_k[r] &= (\widetilde{p}^{(k)}[r] - P_0)\sqrt{I_k} \\ &= [Z_k - z_{r/2}\sqrt{\tfrac{p^{(k)}q^{(k)}}{P_0 Q_0}}; Z_k + z_{r/2}\sqrt{\tfrac{p^{(k)}q^{(k)}}{P_0 Q_0}}], \quad \text{for} \quad k = 1, 2, \ldots, K. \end{aligned}$$

Each r-cut is placed one after the other in order to get a fuzzy test statistic $\widetilde{Z}_k[r]$ at each step of group sequential test. Since the test statistic is fuzzy, the critical values will also, be fuzzy. Thus, substituting c_k for $\alpha_i^*(t)$ functions for z_α continues and then, each r-cut of fuzzy critical value $\widetilde{CV}^*_{(i)k}[r] = [cv_{1ik}(r); cv_{2ik}(r)]$ can be evaluated with given calculations below

$$(2.13) \qquad P\left(Z_k + z_{r/2}\sqrt{\frac{p^{(k)}q^{(k)}}{P_0\ Q_0}} \geq cv_{2ik}(r)\right) = \alpha$$

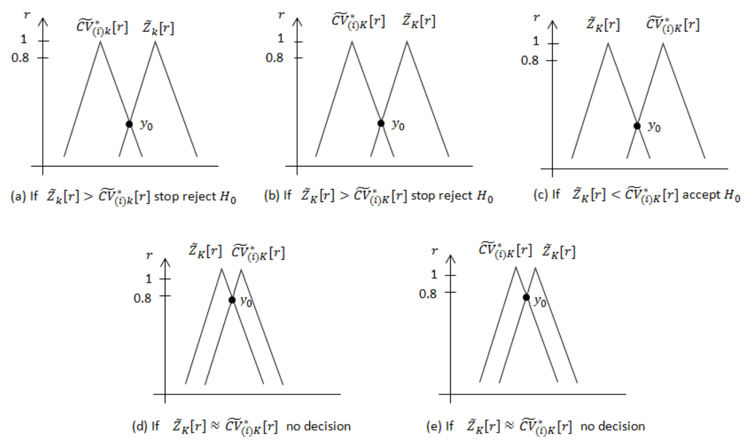
Therefore, fuzzy critical values of group sequential tests for binomial case can be defined as

$$(2.14) \qquad \widetilde{CV}^*_{(i)k}[r] = \left[c_k - z_{r/2}\sqrt{\frac{p^{(k)}\ q^{(k)}}{P_0\ Q_0}}; c_k + z_{r/2}\sqrt{\frac{p^{(k)}\ q^{(k)}}{P_0\ Q_0}}\right]$$

Therefore, the fuzzy test process is as follows; **1.** After group $k = 1, 2, \ldots, K-1$, if $\widetilde{Z}_k[r] > \widetilde{CV}^*_{(i)k}[r]$ stop reject H_0, otherwise continue to $k+1$. **2.** After group K, if

$\widetilde{Z}_K[r] > \widetilde{CV}^*_{(i)K}[r]$ stop reject H_0, if $\widetilde{Z}_K[r] < \widetilde{CV}^*_{(i)K}[r]$ accept H_0, if $\widetilde{Z}_K[r] \approx \widetilde{CV}^*_{(i)K}[r]$ no decision.

Figure 1. Decision criteria of r-cuts approach in group sequential test for binomial case



These situations are detailed in Figure 1. As a result final decision depends on the relationship between $\widetilde{Z}_k[r]$ and $\widetilde{CV}^*_{(i)k}[r]$ for $k = 1, 2, .., K$: (a) $\widetilde{Z}_k[r] > \widetilde{CV}^*_{(i)k}[r]$ reject H_0 (Fig.1-a), (b) $\widetilde{Z}_K[r] > \widetilde{CV}^*_{(i)K}[r]$ stop reject H_0 (Fig.1-b), (c) $\widetilde{Z}_K[r] < \widetilde{CV}^*_{(i)K}[r]$ accept H_0 (Fig.1-c), (d) $\widetilde{Z}_K[r] \approx \widetilde{CV}^*_{(i)K}[r]$ no decision (Fig.1-d,e).

In Figure 1, height of the intersection between two triangular shaped fuzzy number is given as y_0. Buckley and some of the works that uses Buckley's approach state that if $y_0 = 0.8$ than it is impossible to compare these two numbers [6,7,8]. Hence it is taken into account that $y_0 = 0.8$ value for the fuzzy test process decides how much $\widetilde{Z}_k[r]$ is bigger than or less than $\widetilde{CV}^*_{(i)k}[r]$ for $k = 1, 2, .., K$. In some cases it is possible to calculate $\widetilde{Z}_k[r] \approx \widetilde{CV}^*_{(i)k}[r]$ (Fig.1-d,e) for $k = 1, 2, .., K$, so the final decision is *"no decision"* on H_0. That is the result of the fuzzy numbers that incorporate all the uncertainty in confidence intervals [6,7,8]. It is also possible to describe the fuzzy hypothesis testing procedure in more detailed and realistic way when the value of the test statistic is very close to the quantile of the test statistic.

Within the framework of the information given in Section 2.2, group sequential test is modified based on α-spending function for binomial case according to Buckley's approach. In Buckley's approach, fuzzy test statistic is obtained by using more than one confidence interval as the r-cut of triangular shaped fuzzy number. Thus; in this hypothesis testing procedure, group sequential test is done by taking into consideration more than one r value instead of just one value ($r = 1$) and that is the advantage of Buckley's fuzzy approach. Therefore, in this study, it is intended to demonstrate how to use fuzzy approach proposed by Buckley, in group sequential test based on α-spending function for binomial case.

3. An illustrative example

In this section, the use of fuzzy approach to medical data in group sequential test based on $\alpha_i^*(t)$ functions will be described. The medical data of this study is taken from a representative sample of adult population of Turkey which takes parts in the first national fluid and food consumption survey. It is also indicated that, applied survey is intended to reveal the general health status of a representative Turkish population [13]. Emri at al. [13] researched asthma prevalence in five urban centers in Turkey. In Table 1, the prevalence of asthma is shown for five urban regions. Asthma prevelance is 5.6% in Kütahya, 9% in Eskisehir, 5.2% in Mersin, 8.7% in Aksaray and 4.3% in Sakarya. On the whole, one hundred and seven (6.6%) participants stated that they are diagnosed with asthma by a physician in Turkey.

Table 1. Prevalence of asthma cases by region (five urban regions, 2002)

	Kutahya	Eskisehir	Mersin	Aksaray	Sakarya	Total
n(%)	19(5.6)	32(9.0)	19(5.2)	26(8.7)	11(4.3)	107(6.6)
Total	337	357	365	300	255	1614

Traditional statistical analysis is based on crispness of data, random variable, point estimation and so on. However, in real life, it is known that there are many different situations in which the above mentioned concepts are imprecise. Moreover, effects of measurement errors or unrecognized interactions in the estimations of prevalence are inevitable [2,3]. In Buckley's approach, fuzzy asthma prevalence is obtained by using more than one confidence interval as the r-cut of triangular shaped fuzzy number. Thus, more information is used than in the classical approach.

The fuzzy estimations of asthma prevalence for each region are given in Table2. Calculations are performed within the scope of Maple 9 [20]. The fuzzy asthma prevalence for each region is estimated. For example, it is appropriate to say that the asthma prevalence for Kütahya is almost 5.6%, whose r-cuts are represented in Table2. Moreover, it is possible to see both lower (L) and upper (U) values of the estimated asthma prevalence for each region. Here, more information is used regarding not only one value but also all the confidence levels for the estimation of ashtma prevalence under the guidance of Buckley's approach. In more detail, the lower and upper values of estimated asthma prevalence are given such as $r = 0.01$, $r = 0.20$, $r = 0.40$, $r = 0.60$, $r = 0.80$ and lastly $r = 1$ for each region. In Table 2, it can be seen that, if r-cuts increase, lower and upper bounds get closer. If $r = 1$ is taken for each region, the classical results of asthma prevalence which are given in Table 1 is achieved. By estimating the fuzzy prevalence of asthma for each region, more information is used compared to the classical method. Besides, the measurement errors in calculation mistakes can be avoided by using these estimations.

Classical group sequential test is applied for $H_0 : P = 0.06$ versus $H_1 : P > 0.06$ with significance level $\alpha = 0.05$, $K = 3$. Classical group sequential test results are given for different $\alpha_i^*(t)$ functions in Table 3. In this study, $\alpha_2^*(t)$ values are taken into account to test hypothesis at each stage for each region.

Table 2. Fuzzy prevalence of asthma cases by region (five urban regions, 2002))

	p		$r = 0.01$	$r = 0.20$	$r = 0.40$	$r = 0.60$	$r = 0.80$	$r = 1.00$
Kutahya	5.6	L	5.580	5.584	5.590	5.593	5.597	**5.600**
		U	5.620	5.616	5.610	5.607	5.604	**5.600**
Eskisehir	9.0	L	8.975	8.980	8.987	8.992	8.996	**9.000**
		U	9.025	9.020	9.013	9.008	9.004	**9.000**
Mersin	5.2	L	5.181	5.186	5.191	5.195	5.198	**5.200**
		U	5.219	5.215	5.210	5.206	5.203	**5.200**
Aksaray	8.7	L	8.674	8.679	8.687	8.692	8.696	**8.700**
		U	8.728	8.728	8.714	8.708	8.704	**8.700**
Sakarya	4.3	L	4.278	4.284	4.289	4.294	4.297	**4.300**
		U	4.322	4.317	4.311	4.307	4.303	**4.300**
Total	6.6	L	6.589	6.592	6.595	6.597	6.598	**6.600**
		U	6.611	6.608	6.605	6.604	6.602	**6.600**

Table 3. Classical Group Sequential Test Results for different $\alpha_i^*(t)$ functions

		Stage		
Region		**k=1**	**k=2**	**k=3**
Kutahya	$t_i = \frac{n_i}{N_i}$	$200/337 = 0.594$	$270/337 = 0.801$	$337/337 = 1.000$
	$p_i = \frac{n_{\text{asthma}}}{n_i}$	$4/200 = 0.020$	$11/270 = 0.041$	$19/337 = 0.056$
	Z_i	−2.381	-1.335	-0.773
	$\alpha_1^*(t)$	2.292	1.195	1.739
	$\boldsymbol{\alpha_2^*(t)}$	**1.810**	**1.996**	**2.020**
	$\alpha_3^*(t)$	1.886	1.966	1.922
	$\alpha_4^*(t)$	1.988	1.950	1.836
	$\alpha_5^*(t)$	2.106	1.958	1.782
Eskisehir	$t_i = \frac{n_i}{N_i}$	$150/357 = 0.420$	$214/357 = 0.599$	$357/357 = 1.000$
	$p_i = \frac{n_{\text{asthma}}}{n_i}$	$6/150 = 0.040$	$15/214 = 0.070$	$32/357 = 0.089$
	Z_i	−1.031	0.616	1.796
	$\alpha_1^*(t)$	2.807	2.305	1.681
	$\boldsymbol{\alpha_2^*(t)}$	**1.924**	**2.074**	**1.950**
	$\alpha_3^*(t)$	2.033	2.093	1.857
	$\alpha_4^*(t)$	2.208	2.136	1.774
	$\alpha_5^*(t)$	2.373	2.201	1.727
Mersin	$t_i = \frac{n_i}{N_i}$	$76/365 = 0.208$	$220/365 = 0.603$	$365/365 = 1.000$
	$p_i = \frac{n_{\text{asthma}}}{n_i}$	$3/76 = 0.039$	$12/220 = 0.055$	$19/365 = 0.052$
	Z_i	−0.753	−0.344	−0.636
	$\alpha_1^*(t)$	4.139	2.271	1.680
	$\boldsymbol{\alpha_2^*(t)}$	**2.162**	**1.969**	**1.951**
	$\alpha_3^*(t)$	2.311	1.999	1.856
	$\alpha_4^*(t)$	2.594	2.048	1.770
	$\alpha_5^*(t)$	2.853	2.125	1.723
Aksaray	$t_i = \frac{n_i}{N_i}$	$100/300 = 0.333$	$200/300 = 0.667$	$300/300 = 1.000$
	$p_i = \frac{n_{\text{asthma}}}{n_i}$	$2/100 = 0.02$	$10/200 = 0.05$	$26/300 = 0.087$
	z_i	−1.853	−0.595	1.969
	$\alpha_1^*(t)$	3.200	2.141	1.695
	$\boldsymbol{\alpha_2^*(t)}$	**2.002**	**1.994**	**1.980**
	$\alpha_3^*(t)$	2.128	1.998	1.881
	$\alpha_4^*(t)$	2.341	2.019	1.792
	$\alpha_5^*(t)$	2.539	2.069	1.741
Sakarya	$t_i = \frac{n_i}{N_i}$	$119/255 = 0.467$	$194/255 = 0.761$	$255/255 = 1.000$
	$p_i = \frac{n_{\text{asthma}}}{n_i}$	$1/119 = 0.008$	$4/187 = 0.021$	$11/255 = 0.043$
	Z_i	−2.389	−2.228	−1.136
	$\alpha_1^*(t)$	2.642	1.989	1.722
	$\boldsymbol{\alpha_2^*(t)}$	**1.889**	**1.988**	**2.015**
	$\alpha_3^*(t)$	1.989	1.966	1.913
	$\alpha_4^*(t)$	2.150	1.957	1.822
	$\alpha_5^*(t)$	2.294	1.977	1.768

When the results for Kütahya are examined, it can be seen that asthma prevelance is 2% at stage 1 and test statistic is obtained as -2.381, this value is compared with the critical value $\alpha_2^*(t) = 1.81$. It takes us to the next step because $Z_1 = -2.381 < \alpha_2^*(t) = 1.81$. Then, in stage 2, it can bee seen that $Z_2 = -1.335 < \alpha_2^*(t) = 1.9964$, hence this leads us to next stage. In stage 3, $Z_3 = -0.773 < \alpha_2^*(t) = 2.020$ hence we stop and accept H_0.

Test statistic for Eskisehir is calculated as $Z_1 = -1.031$ in the first step and then comes the next step because $Z_1 = -1.031 < \alpha_2^*(t) = 1.9241$. In the second step, it is calculated that $Z_2 = 0.616 < \alpha_2^*(t) = 2.074$. Therefore it is proceeded with step 3. It is obtained that $Z_3 = 1.796 < \alpha_2^*(t) = 1.950$, thus we stop and accept H_0.

Test statistic for Mersin is calculated as $Z_1 = -0.753$ in the first stage, later, it leads us to the next step because $Z_1 = -0.753 < \alpha_2^*(t) = 2.162$. In the second step, it is obtained that $Z_2 = -0.344 < \alpha_2^*(t) = 1.969$ hence this takes us to last step. Calculation is performed as such $Z_3 = -0.636 < \alpha_2^*(t) = 1.951$ in the third step so we stop and accept H_0.

Test statistic and critical value is obtained as $Z_1 = -1.853 < \alpha_2^*(t) = 2.002$ for Aksaray so it proceeds to second step. It is calculated as $Z_2 = -0.595 < \alpha_2^*(t) = 1.994$, therefore this takes us to step 3. It is obtained that $Z_3 = 1.969 < \alpha_2^*(t) = 1.9802$, thus we stop and accept H_0.

Test statistic for Sakarya is calculated as $Z_1 = -2.389$ in the first stage, then this leads us to the next step because $Z_1 = -2.389 < \alpha_2^*(t) = 1.889$. In the second step it is obtained that $Z_2 = -2.228 < \alpha_2^*(t) = 1.988$ hence it carries us to last step. It is calculated that $Z_3 = -1.136 < \alpha_2^*(t) = 2.015$ in the third step so we stop and accept H_0.

In Buckley's approach, fuzzy test statistic is obtained with using more than one confidence interval as the r-cut of triangular shaped fuzzy number. Thus, more information is used in hypothesis testing procedure. However, sample size is fixed in this approach. Fixed sample size is not beneficial in the medical studies in which data comes sequentially. For this purpose, it is illustrated in this section how to use fuzzy approach proposed by Buckley in group sequential test based on α-spending function for binomial case for the prevalence of asthma. Table 4-8 show the results of fuzzy group sequential test based on $\alpha_i^*(t)$ functions for asthma prevalence for Kütahya, Eskisehir, Mersin, Aksaray and Sakarya respectively by using fuzzy test statistics. In all regions, no matter which $\alpha_i^*(t)$ function has been used, (H_0) hypothesis has been accepted at the end of step 3. However, in Eskisehir and Aksaray regions, only of $\alpha_2^*(t)$ function is used, (H_0) hypothesis has been accepted at the end of step 3. If other functions are used, H_0 hypothesis has been rejected at the end of step 3. These tables give fuzzy estimations of asthma prevalence $\widetilde{p}_i[r]$, fuzzy test statistics $\widetilde{Z}_i[r]$ and fuzzy critical values $\widehat{\alpha}_i^*(t)$ with the help of equation (2.11), (2.12) and (2.14) for each urban regions in every stage of group sequential test. As a result, Table 4-8 indicate fuzzy group sequential test for different r-cuts ($r = 0.01, 0.20, 0.40, 0.60, 0.80, 1.00$).

Table 4. Fuzzy Group Sequential Test Results for different $\widetilde{\alpha}_i^*(t)$ functions for Kutahya.

Stage			$r=0.01$	$r=0.20$	$r=0.40$	$r=0.60$	$r=0.80$	$r=1.00$
$k=1$ ($t_1=0.594$)	$\widetilde{p}_i[r]$	L	-0.005	0.001	0.004	0.005	0.019	**0.020**
		U	0.043	0.042	0.040	0.022	0.020	**0.020**
	$\widetilde{Z}_i[r]$	L	-3.911	-3.513	-3.106	-2.977	-2.412	**-2.381**
		U	-0.089	-1.912	-1.588	-2.103	-2.297	**-2.381**
	$\widetilde{\alpha}_1^*(t)$	L	0.765	1.531	1.784	1.973	2.156	**2.292**
		U	3.794	3.040	2.786	2.592	2.431	**2.292**
	$\widetilde{\alpha}_2^*(t)$	L	**0.291**	**0.932**	**1.302**	**1.645**	**1.713**	**1.810**
		U	**3.330**	**3.018**	**2.970**	**2.677**	**2.364**	**1.810**
	$\widetilde{\alpha}_3^*(t)$	L	0.364	1.119	1.383	1.566	1.728	**1.886**
		U	3.404	2.628	2.385	2.191	2.035	**1.886**
	$\widetilde{\alpha}_4^*(t)$	L	0.467	1.226	1.491	1.692	1.827	**1.988**
		U	3.491	2.738	2.478	2.299	2.136	**1.988**
	$\widetilde{\alpha}_5^*(t)$	L	0.579	1.345	1.595	1.786	1.959	**2.106**
		U	3.608	2.859	2.595	2.428	2.255	**2.106**
$k=2$ ($t_2=0.801$)	$\widetilde{p}_i[r]$	L	0.005	0.012	0.022	0.031	0.039	**0.041**
		U	0.077	0.062	0.053	0.052	0.042	**0.041**
	$\widetilde{Z}_i[r]$	L	-3.552	-3.098	-2.365	-2.131	-1.612	**-1.335**
		U	0.653	0.077	-0.879	-0.978	-1.091	**-1.335**
	$\widetilde{\alpha}_1^*(t)$	L	-0.954	0.144	0.485	0.755	0.973	**1.195**
		U	3.319	2.259	1.886	1.622	1.391	**1.195**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.084**	**0.715**	**1.210**	**1.764**	**1.874**	**1.996**
		U	**4.072**	**3.614**	**3.089**	**2.606**	**2.037**	**1.996**
	$\widetilde{\alpha}_3^*(t)$	L	-0.179	0.901	1.259	1.518	1.776	**1.966**
		U	4.069	2.996	2.658	2.359	2.154	**1.966**
	$\widetilde{\alpha}_4^*(t)$	L	-0.191	0.891	1.234	1.531	1.736	**1.950**
		U	4.073	2.997	2.647	2.383	2.159	**1.950**
	$\widetilde{\alpha}_5^*(t)$	L	-0.212	0.886	1.250	1.535	1.753	**1.958**
		U	4.061	3.016	2.652	2.401	2.156	**1.958**
$k=3$ ($t_1=1.000$)	$\widetilde{p}_i[r]$	L	-0.003	0.021	0.042	0.050	0.052	**0.056**
		U	0.081	0.080	0.074	0.061	0.058	**0.056**
	$\widetilde{Z}_i[r]$	L	-3.192	-2.658	-1.889	-1.367	-0.978	**-0.773**
		U	1.591	1.356	0.017	-0.029	-0.589	**-0.773**
	$\widetilde{\alpha}_1^*(t)$	L	-0.778	0.486	0.935	1.225	1.461	**1.739**
		U	4.233	3.007	2.551	2.233	1.994	**1.739**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.390**	**0.007**	**0.908**	**1.209**	**1.906**	**2.020**
		U	**4.540**	**3.968**	**3.307**	**3.0281**	**2.783**	**2.020**
	$\widetilde{\alpha}_3^*(t)$	L	-0.572	0.677	1.103	1.423	1.675	**1.922**
		U	4.402	3.145	2.711	2.444	2.147	**1.922**
	$\widetilde{\alpha}_4^*(t)$	L	-0.644	0.560	1.017	1.314	1.603	**1.836**
		U	4.315	3.020	2.708	2.334	2.075	**1.836**
	$\widetilde{\alpha}_5^*(t)$	L	-0.721	0.551	0.963	1.290	1.557	**1.782**
		U	4.276	3.050	2.593	2.296	2.104	**1.782**

Table 5. Fuzzy Group Sequential Test Results for different $\widetilde{\alpha}_i^*(t)$ functions for Eskisehir

Stage			$r = 0.01$	$r = 0.20$	$r = 0.40$	$r = 0.60$	$r = 0.80$	$r = 1.00$
$k = 1$, $t_1 = 0.420$	$\widetilde{p}_i[r]$	L	-0.001	0.019	0.027	0.032	0.036	**0.040**
		U	0.081	0.061	0.053	0.048	0.044	**0.040**
	$\widetilde{Z}_i[r]$	L	-3.064	-2.010	-1.685	-1.435	-1.226	**-1.031**
		U	0.986	-0.007	-0.349	-0.643	-0.832	**-1.031**
	$\widetilde{\alpha}_1^*(t)$	L	0.779	1.793	2.141	2.397	2.613	**2.807**
		U	4.827	3.804	3.451	3.204	2.997	**2.807**
	$\widetilde{\alpha}_2^*(t)$	L	**-1.116**	**0.901**	**1.262**	**1.515**	**1.731**	**1.924**
		U	**3.949**	**2.940**	**2.588**	**2.326**	**2.107**	**1.924**
	$\widetilde{\alpha}_3^*(t)$	L	0.004	1.023	1.372	1.614	1.861	**2.033**
		U	4.053	3.039	2.704	2.443	2.232	**2.033**
	$\widetilde{\alpha}_4^*(t)$	L	0.175	1.225	1.542	1.789	2.014	**2.208**
		U	4.228	3.262	2.905	2.631	2.407	**2.208**
	$\widetilde{\alpha}_5^*(t)$	L	0.344	1.372	1.711	1.958	2.170	**2.373**
		U	4.393	3.449	3.043	2.787	2.580	**2.373**
$k = 2$, $t_2 = 0.599$	$\widetilde{p}_i[r]$	L	0.026	0.047	0.055	0.060	0.066	**0.070**
		U	0.114	0.092	0.085	0.087	0.074	**0.070**
	$\widetilde{Z}_i[r]$	L	-2.026	-0.712	-0.259	0.081	0.358	**0.616**
		U	3.236	1.954	1.497	1.135	0.858	**0.616**
	$\widetilde{\alpha}_1^*(t)$	L	-0.321	0.996	1.450	1.792	2.058	**2.305**
		U	4.927	3.615	3.145	2.839	2.573	**2.305**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.563**	**0.769**	**1.208**	**1.530**	**1.805**	**2.074**
		U	**4.690**	**3.368**	**2.939**	**2.612**	**2.311**	**2.074**
	$\widetilde{\alpha}_3^*(t)$	L	-0.549	0.758	1.228	1.560	1.830	**2.093**
		U	4.725	3.408	2.954	2.642	2.351	**2.093**
	$\widetilde{\alpha}_4^*(t)$	L	-0.501	0.816	1.255	1.587	1.883	**2.136**
		U	4.758	3.461	3.001	2.675	2.404	**2.136**
	$\widetilde{\alpha}_5^*(t)$	L	-0.446	0.902	1.325	1.678	1.938	**2.201**
		U	4.817	3.505	3.071	2.734	2.448	**2.201**
$k = 3$, $t_3 = 1.000$	$\widetilde{p}_i[r]$	L	0.051	0.071	0.077	0.082	0.086	**0.089**
		U	0.27	0.108	0.102	0.097	0.094	**0.089**
	$\widetilde{Z}_i[r]$	L	-1.161	0.313	0.844	1.189	1.528	**1.796**
		U	4.736	3.245	2.754	2.398	2.104	**1.796**
	$\widetilde{\alpha}_1^*(t)$	L	-1.277	0.206	0.693	1.103	1.415	**1.681**
		U	4.649	3.144	2.630	2.280	1.973	**1.681**
	$\widetilde{\alpha}_2^*(t)$	L	**-1.008**	**0.467**	**1.015**	**1.355**	**1.695**	**1.950**
		U	**4.897**	**3.438**	**2.918**	**2.550**	**2.238**	**1.950**
	$\widetilde{\alpha}_3^*(t)$	L	-1.100	0.394	0.908	1.247	1.565	**1.857**
		U	4.804	3.332	2.812	2.495	2.139	**1.857**
	$\widetilde{\alpha}_4^*(t)$	L	-1.178	0.298	0.835	1.186	1.418	**1.774**
		U	4.720	3.194	2.745	2.378	2.034	**1.774**
	$\widetilde{\alpha}_5^*(t)$	L	-1.231	0.246	0.772	1.149	1.428	**1.727**
		U	4.673	3.185	2.703	2.309	2.019	**1.727**

Table 6. Fuzzy Group Sequential Test Results for different $\widetilde{\alpha}_i^*(t)$ functions for Mersin

Stage			$r = 0.01$	$r = 0.20$	$r = 0.40$	$r = 0.60$	$r = 0.80$	$r = 1.00$
$k = 1$, $t_1 = 0.208$	$\widetilde{p}_i[r]$	L	-0.018	0.010	0.021	0.028	0.034	**0.039**
		U	0.097	0.068	0.058	0.051	0.045	**0.039**
	$\widetilde{Z}_i[r]$	L	-2.774	-1.676	-1.416	-1.172	-0.954	**-0.753**
		U	1.264	0.247	-0.097	-0.350	-0.583	**-0.753**
	$\widetilde{\alpha}_1^*(t)$	L	2.126	3.121	3.475	3.731	3.918	**4.139**
		U	6.153	5.152	4.766	4.547	4.332	**4.139**
	$\widetilde{\alpha}_2^*(t)$	L	**0.138**	**1.177**	**1.498**	**1.754**	**1.957**	**2.162**
		U	**4.171**	**3.147**	**2.872**	**2.567**	**2.368**	**2.162**
	$\widetilde{\alpha}_3^*(t)$	L	0.294	1.306	1.651	1.919	2.110	**2.311**
		U	4.308	3.284	2.971	2.728	2.504	**2.311**
	$\widetilde{\alpha}_4^*(t)$	L	0.573	1.608	1.925	2.189	2.409	**2.594**
		U	4.602	3.595	3.237	3.002	2.815	**2.594**
	$\widetilde{\alpha}_5^*(t)$	L	0.836	1.872	2.225	2.436	2.660	**2.853**
		U	4.865	3.866	3.492	3.265	3.058	**2.853**
$k = 2$, $t_2 = 0.603$	$\widetilde{p}_i[r]$	L	0.015	0.035	0.042	0.047	0.051	**0.055**
		U	0.094	0.074	0.067	0.063	0.058	**0.055**
	$\widetilde{Z}_i[r]$	L	-2.686	-1.507	-1.108	-0.811	-0.572	**-0.344**
		U	1.993	0.850	0.420	0.132	-0.138	**-0.344**
	$\widetilde{\alpha}_1^*(t)$	L	-0.083	1.106	1.498	1.806	2.036	**2.271**
		U	4.611	3.412	3.030	2.731	2.516	**2.271**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.382**	**0.802**	**1.223**	**1.474**	**1.730**	**1.969**
		U	**4.310**	**3.136**	**2.733**	**2.435**	**2.198**	**1.969**
	$\widetilde{\alpha}_3^*(t)$	L	-0.361	0.813	1.225	1.537	1.769	**1.999**
		U	4.344	3.142	2.754	2.470	2.247	**1.999**
	$\widetilde{\alpha}_4^*(t)$	L	-0.302	0.891	1.265	1.577	1.814	**2.048**
		U	4.389	3.206	2.808	2.515	2.283	**2.048**
	$\widetilde{\alpha}_5^*(t)$	L	-0.237	0.962	1.342	1.664	1.901	**2.125**
		U	4.470	3.300	2.844	2.617	2.347	**2.125**
$k = 3$, $t_3 = 1.000$	$\widetilde{p}_i[r]$	L	0.022	0.037	0.043	0.046	0.049	**0.052**
		U	0.082	0.067	0.062	0.058	0.055	**0.052**
	$\widetilde{Z}_i[r]$	L	-2.374	-1.217	-0.817	-0.525	-0.282	**-0.636**
		U	2.224	1.081	0.681	0.398	0.173	**-0.636**
	$\widetilde{\alpha}_1^*(t)$	L	-0.621	0.524	0.954	1.243	1.467	**1.680**
		U	3.972	2.808	2.425	2.145	1.907	**1.680**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.345**	**0.799**	**1.196**	**1.453**	**1.729**	**1.951**
		U	**4.242**	**3.074**	**2.701**	**2.416**	**2.159**	**1.951**
	$\widetilde{\alpha}_3^*(t)$	L	-0.449	0.733	1.116	1.319	1.620	**1.856**
		U	4.148	3.003	2.606	2.321	2.101	**1.856**
	$\widetilde{\alpha}_4^*(t)$	L	-0.526	0.605	1.016	1.357	1.553	**1.770**
		U	4.026	2.936	2.534	2.230	2.011	**1.170**
	$\widetilde{\alpha}_5^*(t)$	L	-0.578	0.590	0.978	1.263	1.501	**1.723**
		U	4.015	2.884	2.468	2.197	1.964	**1.723**

Table 7. Fuzzy Group Sequential Test Results for different $\widetilde{\alpha}_i^*(t)$ functions for Aksaray

Stage			$r=0.01$	$r=0.20$	$r=0.40$	$r=0.60$	$r=0.80$	$r=1.00$
$k=1$ $t_1=0.333$	$\widetilde{p}_i[r]$	L	-0.016	0.002	0.009	0.012	0.017	**0.020**
		U	0.056	0.038	0.032	0.027	0.024	**0.020**
	$\widetilde{Z}_i[r]$	L	-3.302	-2.573	-2.314	-2.150	-1.979	**-1.853**
		U	-0.415	-1.129	-1.388	-1.564	-1.709	**-1.853**
	$\widetilde{\alpha}_1^*(t)$	L	1.748	2.479	2.718	2.894	3.056	**3.200**
		U	4.646	3.914	3.675	3.492	3.351	**3.200**
	$\widetilde{\alpha}_2^*(t)$	L	**0.557**	**1.309**	**1.153**	**1.703**	**1.844**	**2.002**
		U	**3.433**	**2.716**	**2.491**	**2.315**	**2.125**	**2.002**
	$\widetilde{\alpha}_3^*(t)$	L	0.697	1.421	1.688	1.829	1.984	**2.128**
		U	3.566	2.828	2.603	2.420	2.272	**2.128**
	$\widetilde{\alpha}_4^*(t)$	L	0.902	1.599	1.845	2.021	2.197	**2.341**
		U	3.758	3.055	2.752	2.633	2.492	**2.341**
	$\widetilde{\alpha}_5^*(t)$	L	1.101	1.818	2.057	2.226	2.395	**2.539**
		U	3.964	3.267	3.000	2.817	2.690	**2.539**
$k=2$ $t_2=0.667$	$\widetilde{p}_i[r]$	L	0.011	0.030	0.037	0.042	0.046	**0.050**
		U	0.089	0.069	0.062	0.058	0.054	**0.050**
	$\widetilde{Z}_i[r]$	L	-2.856	-1.723	-1.345	-1.085	-0.825	**-0.595**
		U	1.643	0.509	0.131	-0.152	-0.412	**-0.595**
	$\widetilde{\alpha}_1^*(t)$	L	-0.116	1.033	1.381	1.678	1.949	**2.141**
		U	4.363	3.266	2.480	2.598	2.375	**2.141**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.287**	**0.873**	**1.248**	**1.498**	**1.772**	**1.994**
		U	**4.193**	**3.095**	**2.270**	**2.483**	**2.209**	**1.994**
	$\widetilde{\alpha}_3^*(t)$	L	-0.275	0.874	1.324	1.536	1.774	**1.998**
		U	4.234	3.073	2.698	2.448	2.261	**1.998**
	$\widetilde{\alpha}_4^*(t)$	L	-0.238	0.891	1.293	1.569	1.795	**2.019**
		U	4.218	3.138	2.774	2.485	2.184	**2.019**
	$\widetilde{\alpha}_5^*(t)$	L	-0.203	0.978	1.333	1.600	1.816	**2.069**
		U	4.304	3.200	2.819	2.565	2.311	**2.069**
$k=3$ $t_3=1.000$	$\widetilde{p}_i[r]$	L	0.045	0.066	0.073	0.079	0.082	**0.087**
		U	0.129	0.108	0.100	0.095	0.091	**0.087**
	$\widetilde{Z}_i[r]$	L	-0.985	0.557	1.015	1.412	1.687	**1.969**
		U	4.847	3.442	2.908	2.542	2.236	**1.969**
	$\widetilde{\alpha}_1^*(t)$	L	-1.244	0.237	0.725	1.092	1.443	**1.695**
		U	4.558	3.153	2.619	2.298	1.932	**1.695**
	$\widetilde{\alpha}_2^*(t)$	L	**-0.959**	**0.492**	**0.980**	**1.347**	**1.698**	**1.980**
		U	**4.874**	**3.423**	**2.965**	**2.568**	**2.263**	**1.980**
	$\widetilde{\alpha}_3^*(t)$	L	-1.008	0.465	0.940	1.278	1.569	**1.881**
		U	4.713	3.302	2.873	2.458	2.121	**1.881**
	$\widetilde{\alpha}_4^*(t)$	L	-1.132	0.334	0.853	1.158	1.509	**1.792**
		U	4.655	3.250	2.731	2.380	2.059	**1.792**
	$\widetilde{\alpha}_5^*(t)$	L	-1.168	0.282	0.863	1.122	1.458	**1.741**
		U	4.619	3.245	2.695	2.328	1.947	**1.741**

Table 8. Fuzzy Group Sequential Test Results for different $\widetilde{\alpha}_i^*(t)$ functions for Sakarya

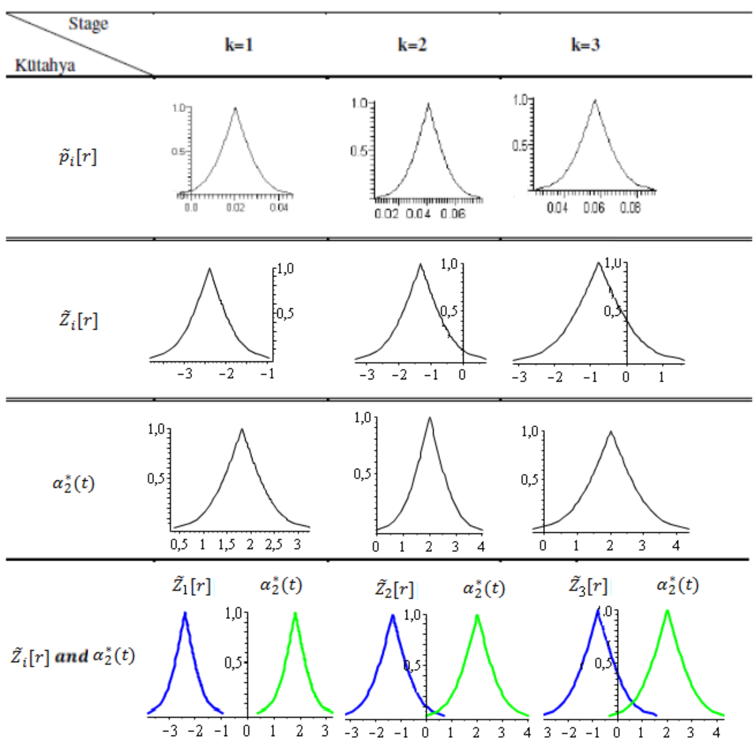
Stage			$r = 0.01$	$r = 0.20$	$r = 0.40$	$r = 0.60$	$r = 0.80$	$r = 1.00$
	$\widetilde{p}_i[r]$	L	-0.013	-0.003	0.001	0.004	0.006	**0.008**
		U	0.029	0.019	0.015	0.012	0.010	**0.008**
	$\widetilde{Z}_i[r]$	L	-3.30	-2.820	-2.670	-2.579	-2.472	**-2.389**
		U	-1.480	-1.936	-2.075	-2.198	-2.284	**-2.389**
	$\widetilde{\alpha}_1^*(t)$	L	1.734	2.179	2.331	2.444	2.553	**2.642**
		U	3.556	3.122	2.949	2.829	2.715	**2.642**
$k = 1$	$\widetilde{\alpha}_2^*(t)$	L	**0.970**	**1.433**	**1.591**	**1.728**	**1.788**	**1.889**
$t_1 = 0.467$		U	**2.807**	**2.332**	**2.185**	**2.076**	**1.995**	**1.889**
	$\widetilde{\alpha}_3^*(t)$	L	1.071	1.537	1.667	1.786	1.905	**1.989**
		U	2.892	2.447	2.252	2.160	2.062	**1.989**
	$\widetilde{\alpha}_4^*(t)$	L	1.222	1.682	1.854	1.959	2.056	**2.150**
		U	3.054	2.615	2.441	2.333	2.225	**2.150**
	$\widetilde{\alpha}_5^*(t)$	L	1.359	1.837	1.996	2.110	2.208	**2.294**
		U	3.209	2.758	2.586	2.593	2.374	**2.294**
	$\widetilde{p}_i[r]$	L	0.009	0.015	0.017	0.019	0.020	**0.021**
		U	0.033	0.027	0.025	0.024	0.023	**0.021**
	$\widetilde{Z}_i[r]$	L	-3.726	-2.982	-2.735	-2.545	-2.371	**-2.228**
		U	-0.739	-1.507	-1.759	-1.920	-2.086	**-2.228**
	$\widetilde{\alpha}_1^*(t)$	L	0.498	1.238	1.497	1.668	1.854	**1.989**
		U	3.472	2.744	2.477	2.284	2.131	**1.989**
$k = 2$	$\widetilde{\alpha}_2^*(t)$	L	**0.487**	**1.241**	**1.496**	**1.682**	**1.839**	**1.988**
$t_2 = 0.761$		U	**3.475**	**2.706**	**2.477**	**2.258**	**2.131**	**1.988**
	$\widetilde{\alpha}_3^*(t)$	L	0.468	1.222	1.467	1.643	1.810	**1.966**
		U	3.435	2.695	2.441	2.245	2.114	**1.966**
	$\widetilde{\alpha}_4^*(t)$	L	0.459	1.192	1.460	1.642	1.794	**1.957**
		U	3.433	2.679	2.440	2.258	2.106	**1.957**
	$\widetilde{\alpha}_5^*(t)$	L	0.486	1.226	1.480	1.662	1.821	**1.977**
		U	3.453	2.758	2.489	2.256	2.119	**1.977**
	$\widetilde{p}_i[r]$	L	0.027	0.035	0.038	0.040	0.042	**0.043**
		U	0.060	0.051	0.049	0.046	0.045	**0.043**
	$\widetilde{Z}_i[r]$	L	-3.210	-2.176	-1.183	-1.153	-1.329	**-1.136**
		U	0.949	-0.118	-0.481	-0.734	-0.954	**-1.136**
	$\widetilde{\alpha}_1^*(t)$	L	-0.438	0.685	1.033	1.314	1.482	**1.722**
		U	3.795	2.762	2.369	2.156	1.920	**1.722**
$k = 3$	$\widetilde{\alpha}_2^*(t)$	L	**-0.073**	**0.955**	**1.366**	**1.608**	**1.789**	**2.015**
$t_3 = 1.000$		U	**4.085**	**3.046**	**2.719**	**2.417**	**2.224**	**2.015**
	$\widetilde{\alpha}_3^*(t)$	L	-0.189	0.884	1.214	1.467	1.721	**1.913**
		U	3.973	2.923	2.570	2.310	2.122	**1.913**
	$\widetilde{\alpha}_4^*(t)$	L	-0.288	0.795	1.129	1.383	1.590	**1.822**
		U	3.872	2.846	2.489	2.235	2.028	**1.822**
	$\widetilde{\alpha}_5^*(t)$	L	-0.329	0.693	1.057	1.352	1.545	**1.768**
		U	3.863	2.807	2.398	2.170	1.977	**1.768**

The results from $r = 1$ are the same as the classical group sequential test. Thus, fuzzy group sequential test is carried out regarding more than one r value instead of just one ($r = 1$), which is the advantage of fuzzy approach. Furthermore, researcher can test hypothesis in different levels. If researcher thinks that uncertainty level is high, then hypothesis can be tested at r=0.01. However, if s/he thinks uncertainty is low, then hypothesis can be tested at r=0.80. Besides, it is indicated that crisp values are obtained for group sequential test if $r = 1$ is taken for each step. These values are given in Table 4-8.

For example, fuzzy asthma prevalence for Kütahya is calculated as $\widetilde{p}_1[r = 0.20] = [0.001, 0.042]$ which gets narrower at $r = 0.80$ as $\widetilde{p}_1[r = 0.80] = [0.019, 0.020]$ at stage one. Furthermore, at $r = 1.00$ it is obtained that $\widetilde{p}_1[r = 1.00] = [0.020, 0.020]$ which is equal to the classical approach results for Kütahya in stage one. Besides, fuzzy test statistic is obtained as $\widetilde{Z}_1[0.20] = [-3.513, -1.912]$ and fuzzy critical value $\widetilde{CV}^*_{(2)1}[0.20] = [0.932, 3.018]$ for $r = 0.20$. It is obtained that $\widetilde{Z}_1[0.20] < \widetilde{CV}^*_{(2)1}[0.20]$, then with the framework of the test procedure, we continue to the next step. In the second step, it is calculated that $\widetilde{Z}_2[0.20] = [-3.098, 0.077] < \widetilde{CV}^*_{(2)2}[0.20] = [0.715, 3.614]$ this takes us to the last step. It is obtained that $\widetilde{Z}_3[0.20] = [-2.658, 1.356] < \widetilde{CV}^*_{(2)3}[0.20] = [0.007, 3.968]$ so we stop and accept H_0. Here H_0 is tested according to $r = 0.20$ level.

Fuzzy prevalence of ashthma ($\widetilde{p}_i[r]$), fuzzy test statistics ($\widetilde{Z}_i[r]$), fuzzy critical values are given in detail with Figure 2 for Kütahya in all r-cuts ($0.01 \leq r \leq 1$).

Figure 2. Membership functions of the values in Table 5 for Kutahya



In general, taking into consideration of all the r-cuts for each step with Figure 2, it is clear that $\widetilde{Z}_1 < \widetilde{CV}^*_{(2)1}$ for stage 1 ($k = 1$). In this case, it will proceed to the second stage ($k = 2$). In second stage $\widetilde{Z}_2 < \widetilde{CV}^*_{(2)2}$ hence this leads us to the last step ($k = 3$). When last stage is examined, it is obtained that $\widetilde{Z}_3 < \widetilde{CV}^*_{(2)3}$. It is possible to accept the null hypothesis ($H_0 : P = 0.06$ versus $H_1 : P > 0.06$) at the third stage for Kütahya by taking into consideration of the all uncertainty within the process of using r-cuts. Moreover, as the number of steps increases, fuzzy group sequential test statistic and fuzzy critical value get closer to each other. Hence, closer results to the real values can be achieved in the fuzzy group sequential tests rather than classical group sequential tests.

Same calculations are done for other regions. Therefore, fuzzy asthma prevalence ($\widetilde{p}_i[r]$), fuzzy test statistic ($\widetilde{Z}_i[r]$) and fuzzy critical values ($\widetilde{\alpha}_1^*(t)$,$\widetilde{\alpha}_2^*(t)$,$\widetilde{\alpha}_3^*(t)$,$\widetilde{\alpha}_4^*(t)$,$\widetilde{\alpha}_5^*(t)$) are obtained in each step for each region. These results can be seen in Table 5-8.

4. Conclusion

In this study, hypothesis testing is adapted by using r-cuts for group sequential test based on α-spending function under the guidance of the information given in Section 1. The advantage of r-cuts (fuzzy) approach is that, instead of generating and processing a single confidence interval, all the confidence intervals are calculated in the process of corresponding fuzzy test statistics. Therefore, in this study it is intended to show that this advantage is also valid for the process of group sequential test based on α-spending function. Thus, the advantages of fuzzy set theory is combined with the advantages of group sequential test. If $r = 1$ is taken in each step, fuzzy group sequential test turns into the classical group sequential test procedure.

Consequently, in this paper fuzzy set theory and Buckley's approach are used to solve problems of impreciseness arising in group sequential test for binomial case. Since, in the traditional statistical tests, the parameters are assumed to be precise values, difficulties arise when the parameters become imprecise, especially in the field of medicine. Hence, the vagueness of p usually comes from personal judgment, experiment or estimation, whose accuracy is limited by the experimental or observational errors. It is clear that Buckley's approach, which uses several confidence intervals rather than only one value for estimating and testing fuzzy parameter, is a well known tool. Additionally, group sequential test provide ethical, economical and administrative advantages. As a result, in this study the benefits of two methodologies are combined and it leads us to propose group sequential test for binomial case under the light of Buckley's approach with r-cuts. It is intended to illustrate that how the fuzzy group sequential test could be applied to real life by using asthma data for five urban regions in Turkey.

References

[1] Adams, B.K. and Cydulka, R.K. *Asthma evalution and management*, Emerg Med Clin N Am **21(2)**, 315-330, 2003.

[2] Adlassnig, K.P. *Fuzzy set theory in medical diagnosis*, Transactions on systems, Man,and Cybernetics, **16(2)**, 260-265, 1986.

[3] Adlassnig, K.P., Scheithauer, W., Grabner, G. *CADIAG-2/PANCREAS: An artificial intelligence system based on fuzzy set theory to diagnose pancreatic diseases*, Third International Conference on System Science in Health Care Health Systems Research, 396-399, 1984.

[4] Alpenc, R., Zhao, H., Rebbeck, T.R., Propert, K.J. *Group Sequential Methods and Sample Size Saving in Biomerker-Disease Association Studies*, Genetics **163(3)**, 1215-1219, 2003.

[5] Bacanlı, S., Demirhan, Y.P. *A group sequential test for inverse Gaussian mean*, Statistical Papers **49**, 377-386, 2008.

[6] Buckley, J.J. *Fuzzy Statistics* (Springer-Verlage, Hiedelberg:Germany), 2004.

[7] Buckley, J.J. *Fuzzy statistics:hypothesis testing*, Soft Comput **9**, 512-518, 2005.

[8] Buckley, J.J. *Fuzzy Probability and Statistics.* (Springer-Verlage, Hiedelberg: Germany), 2006.

[9] Castanho, M.J.P.,Barros, L.C.,Yamakami, A. and Venditei,L.L. *Fuzzy Receiver Operating Characteristic Curve: An Option to Evaluate Diagnostic Tests*, IEEE Transactions on Information Technology ın Biomedicine **11(13)**, 244-250, 2007.

[10] Chang, M.N., Hwang, I.K, Shih, W.J. *Group sequential designs using both type I and type II error probability spending functions*, Commun. Statist.Part A Theory Methods **27(6)**, 1323-39, 1998.

[11] Duarte, P.S., Mastrocolla, L.E., Farsky, P.S., Sampaio, C.R.E.P.S., Tonelli, P.A., Barros, L.C., et al. *Selection of patients for myocardial perfusion scintigraphy based on fuzzy sets theory applied to clinical-epidemiological data and treadmill test results*, Brazilian Journal of Medical and Biological research **39(1)**,9-18, 2006.

[12] Dubois, D. and Prade, H. *Operations with fuzzy numbers*, Int J Syst Sci **9(6)**,613-626, 1978.

[13] Emri, S.,Turnagöl, H.,Basoglu, S.,Bacanlı, S.,Guven, G.S. and Aslan, D. *Asthma-like symptoms prevalence in five Turkish urban centers*, Allergol at Immunopathol **33(5)**,270-276, 2005.

[14] Green, R.H., Brighling, C.E., Pavord, I.D" Wardlaw, A.J. *Management of asthma in adults:Current therapy and future directions*, Postgrad Med J **79**, 259-67, 2003.

[15] Jennison, C., Turnbull, B.W. *Group sequential methods with applications to clinical trials*,(Chapman and Hall,New York), 2000.

[16] Kim, K., DeMets, D.L. *Design and analysis of group sequential tests based on the type I error spending rate function*, Biometrika **74(1)**, 149-154, 1987.

[17] Kurt, E., Metintas, S., Basyigit, I., Bulut, I., Coskun, E., Dabak, S. et al. *Prevalence and riskfactors of allergies in Turkey (PARFAIT): results of multicenter cross-sectional study in adults*, Eur Respir J **33**, 724-733, 2009.

[18] Lan, K.K.G., DeMets, D.L. *Discrete sequential boundaries for clinical trials*, Biometrika **70(3)**, 659-663, 1983.

[19] Liang, R.L., Shiyong, L., Xuena, W., Yi, L., Vinay, M., Dorrelyn, P. and Deepak, K. *FM-test: a fuzzy-set-theory-based approach to differential gene expression data analysis, BMC.* Bioinformatics **4**,7, 2006.

[20] Maple 9, Waterloo Maple Inc.,Waterloo,Canada.

[21] O'Brien, P.C., Fleming, T.R. *A multiple testing procedure for clinical trials*, Biometrics **35(3)**,549–556, 1979.

[22] Pampallona, S., Anastasios TA. *Group sequential desgns for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis*, J Stat Plan Inference **42**, 19-35, 1994.

[23] Pasternak, B.S., Shore, R.E. *Group Sequential Methods for Cohort and Case-Control Studies*, J.Choronic Dis., **33(6)**, 365-373, 1980.

[24] Pocock, S.J. *Group sequential methods in the design and analysis of clinical trials.* Biometrika **64(2)**, 191–199, 1977.

[25] Reboussin, D.M., DeMets, D.L., Kim, K., Lan, G.K.K. *Computations for Group Sequential Boundaries Using the Lan-DeMets spending function method*, Controlled Clinical Trials **21(3)**, 190-207, 2000.

[26] Reis, M.A.M., Ortega, N.R.S, Silveira, P.S.P. *Fuzzy expert system in the prediction of neonatal resuscitation*, Brazilian Journal of Medical and Biological Research **37(5)**, 755-764, 2004.

[27] Roy, R.S. *Asthma*, Southern Medical Journal **96(11)**, 1061-1067, 2003.

[28] Rudser, K.D., Emerson, S.S. *Implementing type I and type II error spending for two-sided group sequential designs*, Contemporary Clinical Trials **29(3)**, 351-358, 2008.

[29] Satagopan, J.M., Verbel, D.A., Venkatraman, E.S., Offit, K.E., Begg, C.B. *Two-Stage Designs for Gene-Disease Association Studies*, Biometrics **58**, 163-170, 2002.

[30] Sebille, V., Bellissant, E. *Sequential methods and group sequential designs for comparative clinical trials*, Fundamental and Clinical Pharmacology **17(5)**, 505-516, 2003.

[31] To, T., Stanojevic, S., Moores, G., Gershon, A.S., Bateman, E.D., Cruz, A.A., Boulet, L.P. *Global asthma prevelance in adults: findings from the cross-sectional world health survey*, BMC Public Health **12**, 204-211, 2012.

[32] Whitehead, J., Stratton, I. *Group sequential clinical trials with triangular continuation regions (corr:V39 p1137)*, Biometrics **39**, 227-36, 1983.

[33] Zolnoori, M., Zarandi, M.H.F., Moin, M. F*uzzy rule-base expert system for evaluation possibility of fatal asthma*, Journal of Health Informatics in Developing Countries **5(1)**, 171-184, 2010.

[34] Zolnoori, M., Zarandi, M.H.F., Moin, M. *Application of intelligent systems in asthma disease:designing a fuzzy rule-based system for evaluating level of asthma exacerbation*, J Med Syst **36(4)**, 2071-2083, 2012.

[35] Zadeh, L.A. *Fuzzy sets*. Information and control **8(3)**, 338-353, 1965.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1035 – 1046

Approximation of some discrete-time stochastic processes by differential equations

Raúl Fierro *

Abstract

This work deals with solutions of ordinary differential equations as approximations of some discrete-time stochastic processes. Similarly, these stochastic processes may be seen as schemes of approximation for this solution. Indeed, these stochastic schemes are defined and their convergence to the solution of a differential equation is proven. Moreover, the asymptotic distribution of the fluctuations about the limit solution is studied. This fact gives the asymptotic distribution of a random global error of approximation. Main results are illustrated by means of the so called SIS epidemic model and numerical simulations are carried out.

2000 AMS Classification: 60F17, 34K28.

Keywords: Law of Large Numbers; Central Limit Theorem; Numerical schemes; SIS epidemic model.

Received 26 : 06 : 2013 : Accepted 04 : 11 : 2013

1. Introduction

Often processes associated to population dynamics are mathematically modeled by differential equations and/or stochastic processes, which are of continuous or discrete time. Because the analysis of a model based on differential equations is less cumbersome and more efficient, both from a mathematical point of view as computational, by introducing a stochastic model for a given process is desirable that it can be approximated by the solution of an Ordinary Differential Equation (ODE), as is also the case studied in this work. Some authors such as Kurtz [9, 10, 11] and Darling and Norris [3] have studied the approximation of continuous-time Markov processes with pure jump by solving an ODE. The convergence shown by these authors is almost surely and based on the Markov property of these processes. Our interest is to analyze such an approach for a class of discrete

*Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Avda. Brasil 2950, Valparaíso, Chile. Instituto de Matemática, Universidad de Valparaíso.
Email: rfierro@ucv.cl

time stochastic processes which are not necessarily Markovian, but including discrete-time Markov chains. Conversely, given an ordinary differential equation, it is possible to approximate its solution through this kind of processes. Indeed, our class of stochastic processes can be seen as a stochastic variant of the Euler scheme to approximate the solution to an ODE. These schemes of approximation are presented as discrete-time stochastic processes, which includes but are not limited to Markov chains. Recently some authors, such as Abbasbandy and Bervillier in [1], Eslahchi *et al.* in [4] Parand *et al.* in [5], among others, have studied the problem of approximation for ordinary differential equations by different deterministic methods. Also, stochastic schemes of approximation have been developed by Fierro and Torres in [6] and Kloeden and Platen in [8]. This latter reference deals with schemes of approximation for stochastic differential equations (SDE). In [12], Kushner and Dupuis present a stochastic scheme of approximation for SDE based on a Markov chain, which, in particular, can be applied for approximating solutions to ODEs. Even though, in general, our schemes need not be Markovian, this model can be included in our setting, whenever the noise part of the equation is zero.

The main results presented in this work are the convergence of the mentioned schemes to the solution to the ODE and a central limit theorem, which allows to know the asymptotic distribution of the global error of the approximation. These results are applied to an example coming from the biomathematical literature. Indeed, the differential equation modeling the well-known SIS epidemic model is analyzed under our framework by means of two natural schemes of approximation.

In order to quantify the probability of error in the approximation, the central limit theorem presented allows to know the asymptotic distribution of the global error, i.e. of the fluctuations of the process around the solution to the ODE. By this result it is possible to establish confidence bands around this solution, which are determined by a preassigned probability. Therefore, when a particular heuristic model is defined by a solution to an ODE, it is possible to perform an asymptotical statistical test to validate the model. Indeed, by considering the stochastic model as the observed process, these observations should be close to the solution to the ODE insofar this solution to be a good model for the heuristic situation. Hence, the asymptotic distribution of the global error allows to carry out a goodness of fit test for both the random and deterministic model.

The plan of this paper is as follows. In Section 2, by means of a recursive condition, we define a family of discrete-time stochastic processes, which approximate the solution of the ODE. Main results of this work, along with their proofs, are stated in Section 3. Since our schemes are stochastic, the global error is so. The asymptotic distribution of it is analyzed in Section 4. Moreover, some dispersion measures and their estimators are defined in this section. An example is included in Section 5. Indeed, the solution to the differential equation defining the so called SIS epidemic model is approximated through two schemes included in our framework. Both schemes are compared and numerical simulations are carried out.

2. Preliminaries

Let $x_0 \in \mathbb{R}^d$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function satisfying the following Lipchitz condition:

(L) $\quad \|b(t,x) - b(t,y)\| \leq K\|x-y\|$, for all $t \in \mathbb{R}_+$,

where K is a positive constant and $\|\cdot\|$ stands for the usual norm in $\mathbb{R}^d$. Hence, the initial value problem

$$\dot{x}(t) = b(t, x(t)) \qquad x(0) = x_0, \tag{2.1}$$

has one and only one solution.

In this work, a stochastic scheme of approximation for the solution to (2.1) is stated.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{t_k^n\}_{k\in\mathbb{N}}$ the sequence of non-negative real numbers defined as $t_k^n = Ck/n$, $(C > 0, n \in \mathbb{N}\setminus\{0\})$. In what follows and without loss of generality, we assume $C = 1$. An approximation of the solution to (2.1) is obtained by means of a sequence $\{Z^n\}_{n\in\mathbb{N}}$ of stochastic processes defined on $\mathbb{R}_+ \times \Omega$. Such an approximation is obtained by defining $\mathcal{F}_k^n = \sigma(Z^n(t_1^n), \dots, Z^n(t_k^n))$ as the sigma algebra generated by $Z^n(t_1^n), \dots, Z^n(t_k^n)$, $x^n = x^n(0) + \frac{1}{n}Z^n$ and assuming the following condition:

(C) $\quad \mathbb{E}(\Delta Z^n(t_k^n)|\mathcal{F}_{k-1}^n) = b(t_{k-1}^n, x^n(t_{k-1}^n)), \qquad (k \geq 1),$

where for any stochastic process Z, $\Delta Z(t_k^n) = Z(t_k^n) - Z(t_{k-1}^n)$.

For a real number x, $[x]$ stands for the integer part of x and

$$L^n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} \xi_k^n, \quad (t \geq 0),$$

where $\xi_k^n = \Delta Z^n(t_k^n) - b(t_{k-1}^n, x^n(t_{k-1}^n))$. By defining $\mathcal{F}_t^n = \mathcal{F}_{[nt]}^n$, $(t \geq 0)$, we have L^n is a d-dimensional $\mathcal{F}_t^n-$martingale and

$$Z^n(t) = Z^n(0) + n\sum_{k=1}^{[nt]} b(t_{k-1}^n, x^n(t_{k-1}^n))\Delta t^n + nL^n(t), \quad (t > 0). \tag{2.2}$$

Given any d-dimensional martingale L, its predictable quadratic variation, at time t, is denoted by $\langle L\rangle(t)$. Thus, $\langle L\rangle(t)$ is a $d \times d$-matrix and it directly follows that

$$\langle L^n\rangle(t) = \frac{1}{n^2}\sum_{k=1}^{[nt]} \mathbb{E}(\xi_k^n \xi_k^{n\top}|\mathcal{F}_{k-1}^n), \quad (t \geq 0) \tag{2.3}$$

From (2.1) and (2.2), we have

$$x^n(t) - x(t) = x^n(0) - x(0) + \int_0^{[nt]/n} \{b([nu]/n, x^n(u)) - b(u, x(u))\}\,\mathrm{d}u + L^n(t) + \epsilon^n(t), \tag{2.4}$$

where, $\epsilon^n(t) = x([nt]/n) - x(t)$. Note that $\sup_{0\leq u\leq t}\|\epsilon^n(s)\| \leq S_t/n$, where $S_t = \sup_{0\leq u\leq t}\|b(u, x(u))\|$.

3. Main results

In the sequel, x stands for the solution to (2.1). In this section, the convergence of x^n to x is stated, which means $\{x^n\}_{n\in\mathbb{N}}$ converges uniformly in probability, on compact subsets of $\mathbb{R}_+$, to x as n goes to ∞.

3.1. Theorem. *Assume conditions (C) and (L) are satisfied. Moreover, suppose the following two conditions hold:*

- **(3.1.1):** $x^n(0) \xrightarrow{\mathbb{P}} x_0$.
- **(3.1.2):** *For each* $t \geq 0$,

$$\frac{1}{n^2}\sum_{k=1}^{[nt]} \mathbb{E}(\|\xi_k^n\|^2|\mathcal{F}_{k-1}^n) \xrightarrow{\mathbb{P}} 0, \quad \textit{as} \quad n \to \infty.$$

Then, for each $T > 0$, $\sup\limits_{0\leq t\leq T}\|x^n(t) - x(t)\| \xrightarrow{\mathbb{P}} 0, \quad$ *as* $\quad n \to \infty$.

Proof. Fix $T > 0$ and let $g^n(t) = \sup_{0\leq s\leq t}\|x^n(s) - x(s)\|$, $(t \in [0, T])$. From (2.4) and (L), we obtain

$$g^n(t) \leq \alpha_n + K\int_0^t g^n(u)\,\mathrm{d}u,$$

where $\alpha_n = g^n(0) + \sup_{0\le t\le T} \|L^n(t)\| + S_T/n$. Since, by (3.1.1), $\{g^n(0)\}_{n\in\mathbb{N}}$ converges in probability to zero, by Gronwall's inequality, it suffices to verify that $\{\sup_{0\le t\le T} \|L^n(t)\|\}_{n\in\mathbb{N}}$ converges in probability to zero.

From Theorem 1 by Lenglart [13], for any $\epsilon, \eta > 0$, we have

$$\begin{aligned}\mathbb{P}(\sup_{0\le t\le T} \|L^n(t)\|^2 > \epsilon) &\le \frac{1}{\epsilon}\,\mathbb{E}(\operatorname{tr}\langle L^n\rangle(T) \wedge \eta) + \mathbb{P}(\operatorname{tr}\langle L^n\rangle(T) > \eta)\\ &< \frac{\eta}{\epsilon} + \mathbb{P}(\frac{1}{n^2}\sum_{k=1}^{[nT]} \operatorname{tr}\mathbb{E}(\xi_k^n {\xi_k^n}^\top | \mathcal{F}_{k-1}^n) > \eta)\\ &= \frac{\eta}{\epsilon} + \mathbb{P}(\frac{1}{n^2}\sum_{k=1}^{[nT]} \mathbb{E}(\|\xi_k^n\|^2 | \mathcal{F}_{k-1}^n) > \eta)\end{aligned}$$

and hence, by (3.1.2),

$$\lim_{n\to\infty} \mathbb{P}(\sup_{0\le t\le T} \|L^n(t)\|^2 > \epsilon) = 0,$$

which concludes the proof. ∎

For each $t \in \mathbb{R}_+$, let $\widetilde{b}_t : \mathbb{R}^d \to \mathbb{R}^d$ be such that $\widetilde{b}_t(x) = b(t,x)$ and suppose for each $t \in \mathbb{R}_+$, $\widetilde{b}_t$ has continuous partial derivatives. The following result aims to the problem of finding confident bands for the approximate solution to (2.1). Before stating it, for each $(t,a) \in \mathbb{R}_+ \times \mathbb{R}^d$, let $\mathrm{D}(b)(t,a)$ denote the Jacobian matrix of $\widetilde{b}_t$ at a.

A function $v : \mathbb{R}_+ \to \mathbb{R}^{d\times d}$ is said to be a positive-definite, if for each $\theta \in \mathbb{R}^m$, $f^\theta : \mathbb{R}_+ \to \mathbb{R}$ defined as $f^\theta(t) = \theta^\top v(t)\theta$ is continuous, increasing and $f^\theta(0) = 0$.

3.2. Theorem. *Let v be a positive-definite function and $y^n = \sqrt{n}(x^n - x)$. Suppose the following conditions hold:*

- **(3.2.1):** *The partial derivatives of $\widetilde{b}_t$ exist and are continuous in $\mathbb{R}^d$.*
- **(3.2.2):** *For each $\epsilon > 0$ and $t \ge 0$, $\frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}(\|\xi_k^n\|^2 \mathrm{I}_{\{\|\xi_k^n\| > \epsilon\sqrt{n}\}} | \mathcal{F}_{k-1}^n) \xrightarrow{\mathbb{P}} 0$, as n goes to ∞.*
- **(3.2.3):** *$\{y^n(0)\}_{n\in\mathbb{N}}$ converges in distribution to a random variable η.*
- **(3.2.4):** *For each $t \ge 0$, $\sup_{0\le s\le t} \|\frac{1}{n}\sum_{k=1}^{[ns]} \mathbb{E}(\xi_k^n {\xi_k^n}^\top | \mathcal{F}_{k-1}^n) - v(s)\| \xrightarrow{\mathbb{P}} 0$, as n goes to ∞.*

Then, the sequence $\{y^n\}_{n\in\mathbb{N}}$ converges in law to the solution y satisfying the following stochastic differential equation:

$$\mathrm{d}y(t) = \mathrm{D}(b)(t, x(t))y(t)\mathrm{d}t + \mathrm{d}m(t), \quad y(0) = \eta, \tag{3.1}$$

where m is a d-dimensional continuous martingale starting at zero with predictable quadratic variation, at $t \ge 0$, given by $\langle m\rangle(t) = v(t)$.

Proof. Condition (3.2.2) implies the jump asymptotic rarefaction condition in [14, Theorem 8, Chapter II.5] by Rebolledo, for the sequence of martingales $\{m^n\}_{n\in\mathbb{N}}$, where $m^n = \sqrt{n}L^n$. This fact along with condition (3.2.4) imply $\{m^n\}_{n\in\mathbb{N}}$ converges in law to a continuous martingale m starting at zero and having predictable quadratic variation $\langle m\rangle$ given by $\langle m\rangle(t) = v(t)$.

Let b_i be the i-th coordinate of b, $(i = 1, \ldots, d)$. By the Value Mean Theorem, there exists $\theta_i^n(t) \in \mathbb{R}^d$ between $x(t)$ and $x^n(t)$ such that $b_i(t, x^n(t)) - b_i(t, x(t)) = \mathrm{D}(b_i)(t, \theta_i^n(t))(x^n(t) - x(t))$, where $\mathrm{D}(b_i)(t,a)$ is the Jacobian matrix of $b_i(t,\cdot)$ at $a \in \mathbb{R}^d$.

From (2.4), it is derived

$$(3.2)\qquad y^n(t) = y^n(0) + \int_0^t \mathrm{D}_n(u) y^n(s)\,\mathrm{d}s + m^n(t) + \sqrt{n}\epsilon^n(t),$$

where $\mathrm{D}_n(u) = (\mathrm{D}(b_1)(u, \theta_1^n(u)), \ldots, \mathrm{D}(b_d)(u, \theta_d^n(u)))^\top$.

Consequently,

$$\sup_{0\le u\le t} \|y^n(u)\| \le \|y^n(0)\| + C(t)\int_0^t \sup_{0\le u\le s} \|y^n(u)\|\,\mathrm{d}s + \sup_{0\le u\le t}\|m^n(u)\| + \frac{S_t}{\sqrt{n}},$$

where $C_n(t) = \sup_{0\le u\le t}\sup_{\|y\|=1}\|\mathrm{D}_n(u)y\|$ and $C(t) = \sup_{n\in\mathbb{N}} C_n(t)$. From (3.2.1) and Theorem 3.1, $\{\sup_{0\le u\le t}\|\mathrm{D}_n(u)\|\}_{n\in\mathbb{N}}$ converges in probability to $\sup_{0\le u\le t}\|\mathrm{D}(b)(u, x(u))\|$ and thus, $C(t) < \infty$.

Hence, from a standard application of the Gronwall inequality, we obtain

$$(3.3)\qquad \sup_{0\le u\le t}\|y^n(u)\| \le (\|y^n(0)\| + \sup_{0\le u\le t}\|m^n(u)\| + S_t/\sqrt{n})\,\mathrm{e}^{tC(t)}\,.$$

In order to prove the convergence in law of $\{y^n\}_{n\in\mathbb{N}}$ and that its limit has continuous trajectories, Theorem 15.5 by Billingsley (1968) is used. Since $\{y^n(0)\}_{n\in\mathbb{N}}$ converges in distribution, Theorem 6.2 in Billingsley (1968) implies this sequence is tight, which means for each $\epsilon > 0$, there exists $a > 0$ such that $\sup_{n\in\mathbb{N}}\mathbb{P}(\|y^n(0)\| > a) < \epsilon$. Hence

$$(3.4)\qquad \lim_{a\to\infty}\sup_{n\in\mathbb{N}}\mathbb{P}(\|y^n(0)\| > a) = 0.$$

Fix $T > 0$ and let us define the modulus of continuity ω_T as

$$\omega_T(z, \delta) = \sup_{|s-t|<\delta}\|z(s) - z(t)\|,$$

where $\delta > 0$ and $z : [0, T] \to \mathbb{R}^d$ is right continuous and left-hand limited.

From (3.2) we have

$$(3.5)\qquad \omega_T(y^n, \delta) \le \delta C(T)\sup_{0\le t\le T}\|y^n(t)\| + \omega_T(m^n, \delta) + 2S_T/\sqrt{n}.$$

Since $\{m^n\}_{n\in\mathbb{N}}$ converges in distribution to m, it follows from Theorem 15.2 in Billingsley [2] that for each $\epsilon > 0$, $\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\mathbb{P}(\omega_T(m^n, \delta) > \epsilon) = 0$. Hence, from (3.3), for each $\epsilon > 0$,

$$(3.6)\qquad \lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\mathbb{P}(\omega_T(y^n, \delta) > \epsilon) = 0.$$

Conditions (3.4) and (3.6) imply the sequence $\{P_n\}_{n\in\mathbb{N}}$ of probabilities measures, where P_n is the law of y^n, satisfies the hypotheses of Theorem 15.5 in Billingsley [2] and hence, $\{P_n\}_{n\in\mathbb{N}}$ is tight and every limit point P of this sequence satisfies $P(C) = 1$, where C is the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$. This fact, along Theorem 6.1 in Billingsley [2], imply that $\{P_n\}_{n\in\mathbb{N}}$ is relatively compact. Let $\{y^{n_k}\}_{k\in\mathbb{N}}$ a subsequence converging in distribution to a process y. Since, by Theorem 3.1, $\{\mathrm{D}_n\}_{n\in\mathbb{N}}$ converges uniformly in probability to $\mathrm{D}(b)(\cdot, x(\cdot))$, $\{m^n\}_{n\in\mathbb{N}}$ converges in law to m and $\{\sqrt{n}\epsilon^n\}_{n\in\mathbb{N}}$ converges uniformly to 0, it follows from (3.2) that y is a solution to (3.1). Finally, uniqueness of solutions to (3.1) implies $\{y^n\}_{n\in\mathbb{N}}$ converges in distribution to this solution y, which concludes the proof. ∎

Remarks

R1: By Itô's rule, the unique solution to (3.1) is given by

$$y(t) = \Psi(t)\left(\eta + \int_0^t \Psi(s)^{-1}\,\mathrm{d}m(s)\right), \quad 0 \le t \le 1,$$

where Ψ is the unique solution to the matrix differential equation

$$\Psi'(t) = \mathrm{D}(b)(t, x(t))\Psi(t), \quad \Psi(0) = \text{identity matrix}.$$

R2: Condition (3.2.4) holds whenever for each $t \geq 0$ and $\epsilon > 0$,

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}(\|\xi_k^n\|^2 \mathrm{I}_{\{\|\xi_k^n\| > \epsilon\sqrt{n}\}}) \xrightarrow{\mathbb{P}} 0, \text{ as } n \text{ goes to } \infty. \qquad \text{(Lindeberg condition).}$$

4. Random global discretization error

In this section, we assume the partial derivatives of $\widetilde{b}_t$ exist and are continuous in $\mathbb{R}^d$ and for each $n \in \mathbb{N}$, $x^n(0) = x(0)$.

4.1. Some definitions. In order to analyze the error produced by the discretization scheme introduced here, for a fixed $T > 0$, we define the random global error to be $\widehat{e}_T^n = \|x^n(T) - x(T)\|$ and, for $p \geq 1$, the p-mean global error to be $e_T^n(p) = \mathbb{E}(\|x^n(T) - x(T)\|^p)^{1/p}$, whenever $\mathbb{E}(\|x^n(T)\|^p) < \infty$, i.e. $e_T^n(p)$ is the usual norm of $\widehat{e}_T^n$ defined on $\mathrm{L}^p(\Omega, \mathcal{F}, \mathbb{P})$, the space of random variables x such that $\mathbb{E}(|x^p|) < \infty$. We refer to $e_T^n(2)$ as the square mean global error. Since, even in simple cases, it is not possible to know or calculate $e_T^n(p)$, an estimator of this one is obtained by defining

$$\widehat{e}_T^n(p, m) = \left(\frac{1}{m}\sum_{i=1}^m \widehat{e}_T^n(i)^p\right)^{1/p},$$

where $\widehat{e}_T^n(1), \ldots, \widehat{e}_T^n(m)$ are independent random variables with the same distribution than $\widehat{e}_T^n$. Anyway the distribution of $\widehat{e}_T^n$ needs to be known. Theorem 3.2 allows to obtain an approximation of this distribution.

It follows from the Strong Law of Large Numbers by Kolmogorov that the estimator $\widehat{e}_T^n(p, m)$ is strongly consistent, i.e.

$$\lim_{m\to\infty} \widehat{e}_T^n(p, m) = e_T^n(p).$$

Consequently, by carrying out simulations of $\widehat{e}_T^n$, an approximation of $e_T^n(p)$ can be obtained. In particular, the sample variance of $\widehat{e}_T^n$ can be consistently estimated by means of

$$S_m^{n,2} = \frac{1}{m}\sum_{i=1}^m \left(\widehat{e}_T^n(i) - \frac{1}{m}\sum_{i=1}^m \widehat{e}_T^n(i)\right)^2 = \widehat{e}_T^n(2, m)^2 - \widehat{e}_T^n(1, m)^2.$$

Since $(S_m^{n,2}; m \in \mathbb{N})$ converges $\mathbb{P}$-a.s. to $\mathrm{Var}(\widehat{e}_T^n) = e_T^n(2)^2 - e_T^n(1)^2$ and $\mathrm{Var}(\widehat{e}_T^n)$ is a measure of dispersion, small values of $S_m^{n,2}$ suggest no much simulations of $\widehat{e}_T^n$ are necessary to carry out a suitable estimation of the square mean global error.

4.2. Asymptotic distribution of the global error. In this subsection we examine the asymptotical distribution of $\widehat{e}_T^n$. Indeed, let y^n be as in Theorem 3.2 and suppose the hypotheses of this theorem hold. Thus, $\widehat{e}_T^n = (\Delta t^n)^{1/2}\|y^n(T)\|$ and from Theorem 3.2, $\widehat{e}_T^n$ is asymptotically distributed as $\|y(T)\|/\sqrt{n}$, where y is the solution to (3.1) with $y(0) = 0$. From Remark R1,

$$y(T) = \int_0^T B(T, s)\,\mathrm{d}m(s)$$

where $B(t, s) = \Psi(t)\Psi(s)^{-1}$. Hence, by taking into account that, for almost sure $s \geq 0$, there exists $v'(s)$, the derivative of v at s, we have

$$\mathbb{E}(\|y(T)\|^2) = \int_0^T B(T, s)v'(s)B(T, s)^\top\,\mathrm{d}s \tag{4.1}$$

and for large values of n, $e_T^n(2)$, the square mean global error can be approximated by $\sqrt{\mathbb{E}(\|y(T)\|^2)/n}$, whenever $\{\|y^n(T)\|^2\}_{n\in\mathbb{N}}$ is uniformly integrable (see Theorem 5.4 in Billingsley [2]).

4.3. Hypothesis testing. Fix $T > 0$. The global error could be used to develop an asymptotic hypothesis testing to reject or not the validity of the model. This procedure is performed in the following natural manner: given a significance level $\alpha \in (0, 1)$, we choose $t_\alpha > 0$ such that $\mathbb{P}(\|y(T)\| > t_\alpha) = \alpha$. Then, we compare the statistic $\sqrt{n}\widehat{e}_T^n$ with t_α. If $\sqrt{n}\widehat{e}_T^n > t_\alpha$ we reject the hypothesis as false, while if $\sqrt{n}\widehat{e}_T^n \le t_\alpha$, we conclude that there is no sufficient evidence that the model is incorrect. In this case, although the null hypothesis need not to be true, no change in the model is recommended.

5. An example

In this section, we apply the results of this work to a known differential equation coming from the biomathematical literature. A brief description of this model is given in the first subsection and two probabilistic schemes, which are approximated by the solution to this equation, are presented. Results of this work are illustrated by means numerical simulations in Subsection 2.

5.1. The SIS epidemic model. One of the most commonly used differential equations in the biomathematical literature is that correspondig to the SIS epidemic model. In this model it is assumed that at time $t \ge 0$, $x(t)$ and $y(t)$ represent the densities of infective and susceptible individuals, respectively, and they satisfy the following system of ordinary differential equations:

$$\begin{aligned}\tfrac{\mathrm{d}x}{\mathrm{d}t}(t) &= \beta x(t)y(t) - \gamma y(t)\\ \tfrac{\mathrm{d}y}{\mathrm{d}t}(t) &= -\beta x(t)y(t) + \gamma y(t).\end{aligned}$$

Since for each $t \ge 0$, $x(t) + y(t) = 1$, this model is completely determined by the ordinary differential equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = \beta(1 - x(t))x(t) - \gamma x(t). \tag{5.1}$$

This test equation, given $x(0) = x_0 \in]0, 1[$, has the unique solution

$$x(t) = \begin{cases} \frac{x_0}{x_0\beta t+1} & \text{if} \quad \beta = \gamma \\ \frac{x_0(\beta-\gamma)\,\mathrm{e}^{(\beta-\gamma)t}}{\beta-\gamma+x_0\beta(\mathrm{e}^{(\beta-\gamma)t}-1)} & \text{if} \quad \beta \neq \gamma. \end{cases}$$

Let $x^n(0) = [nx_0]/n$. In accordance with our setting, let $Z^n = A^n - B^n$ and $x^n = x^n(0) + \frac{1}{n}Z^n$, where A^n and B^n are independent. Two probability distributions are defined below, in order to Z^n takes values in the nonnegative integers. The first one, which is labeled by distribution D1 is recursively defined as follows. Let $I^n(t_{k-1}^n) = [nx_0] + Z^n(t_{k-1}^n)$ and $S^n(t_{k-1}^n) = n - I^n(t_{k-1}^n)$, conditional on $\mathcal{F}_{k-1}^n$, $\Delta A^n(t_k^n)$ and $\Delta B^n(t_k^n)$ have Binomial distribution with parameters $(S^n(t_{k-1}^n), \beta x^n(t_{k-1}^n)\Delta t^n)$ and $(I^n(t_{k-1}^n), \gamma\Delta t^n)$. I.e., for each $a \in \{0, \dots, S^n(t_{k-1}^n)\}$ and $b \in \{0, \dots, I^n(t_{k-1}^n)\}$,

$$\mathbb{P}(\Delta A^n(t_k^n) = a|\mathcal{F}_{k-1}^n) = \binom{S^n(t_{k-1}^n)}{a} p_{n,k-1}^a (1 - p_{n,k-1})^{S^n(t_{k-1}^n)-a}$$

and

$$\mathbb{P}(\Delta B^n(t_k^n) = b|\mathcal{F}_{k-1}^n) = \binom{I^n(t_{k-1}^n)}{b} q_n^b (1 - q_n)^{I^n(t_k^n)-b},$$

where $p_{n,k-1} = \beta x^n(t_{k-1}^n)\Delta t^n$ and $q_n = \gamma \Delta t^n$. Here, it is suppose n is large enough to $\beta/n \leq 1$ and $\gamma/n \leq 1$.

Since for each $k \in \mathbb{N}$, $0 \leq \Delta A^n(t_k^n), \Delta B^n(t_k^n) \leq n$, $\mathbb{P}$-a.s., on a time interval $[0, T]$, $(T > 0)$, the state space of x^n is a finite subset of $\frac{1}{n}\mathbb{Z}_+ = \{k/n : k \in \mathbb{Z}_+\}$.

By using notations before, we have

$$\xi_k^n = \Delta A^n(t_k^n) - \Delta B^n(t_k^n) - \{\beta S^n(t_{k-1}^n)x^n(t_{k-1}^n)\Delta t^n - \gamma I^n(t_{k-1}^n)\Delta t^n\}$$

and consequently,

$$\mathbb{E}(|\xi_k^n|^2|\mathcal{F}_{k-1}^n) = \beta(1 - x^n(t_{k-1}^n))x^n(t_{k-1}^n)(1 - \beta x^n(t_{k-1}^n)\Delta t^n) + \gamma x^n(t_{k-1}^n)(1 - \gamma \Delta t^n).$$

Since

$$\mathbb{E}(|\xi_k^n|^3|\mathcal{F}_{k-1}^n) = S^n(t_{k-1}^n)p_{n,k-1}(1 - p_{n,k-1})^2 + I^n(t_{k-1}^n)q_n(1 - q_n)^2 \leq \beta + \gamma,$$

$\{|\xi_k^n|^2; n, k \geq 1\}$ is uniformly integrable and Lindeberg condition stated in R2 holds. In addition, $\mathbb{E}(|\xi_k^n|^2|\mathcal{F}_{k-1}^n) \leq 1$ and hence conditions (3.1.2) and (3.2.2) of Theorem 3.1 and 3.2, respectively, are satisfied. Consequently, Theorem 3.1 implies for each $t \geq 0$,

$$\sup_{0 \leq s \leq t} |\frac{1}{n}\sum_{k=1}^{[ns]} \mathbb{E}(|\xi_k^n|^2|\mathcal{F}_{k-1}^n) - v_1(s)| \xrightarrow{\mathbb{P}} 0, \text{ as } n \text{ goes to } \infty,$$

where

$$v_1(s) = \int_0^s \{\beta(1 - x(u)) + \gamma\}x(u)\,\mathrm{d}u. \tag{5.2}$$

As shown in [7], the distribution of x^n has a biomathematical sense, where n denotes the population size. However, other distributions allow x^n is approximated by the solution to (5.1). Indeed, a second distribution for x^n, which has less variability, and we label by D2, is defined as follows. Assumed that, conditional to $\mathcal{F}_{k-1}^n$, $\Delta A^n(t_k^n)$ and $\Delta B^n(t_k^n)$ have Bernoulli distribution with parameters $\beta S^n(t_{k-1}^n)x^n(t_{k-1}^n)\Delta t^n$ and $\gamma I^n(t_{k-1}^n)\Delta t^n$, respectively. For large enough values of n, these conditional parameters are equal or less than one and it is easy to see the hypotheses of Theorems 3.1 and 3.2 hold with a positive-definite function v_2 defined by

$$v_2(s) = \int_0^s \{\beta(1 - x(u))(1 - \beta(1 - x(u))x(u)) + \gamma(1 - \gamma x(u))\}x(u)\,\mathrm{d}u \tag{5.3}$$

and satisfying, for each $t \geq 0$,

$$\sup_{0 \leq s \leq t} |\frac{1}{n}\sum_{k=1}^{[ns]} \mathbb{E}(|\xi_k^n|^2|\mathcal{F}_{k-1}^n) - v_2(s)| \xrightarrow{\mathbb{P}} 0, \text{ as } n \text{ goes to } \infty,$$

This latter probabilistic scheme has some advantages regarding the conditional Binomial jumps case, labeled by D1. One of them is that, according to D2 distribution, ΔZ^n takes values in the set $\{-1, 0, 1\}$ instead of $\{-n, \ldots, 0, \ldots, n\}$ as in the D1 distribution. In addition, from (5.2) and (5.3), for each $s \geq 0$,

$$v_1'(s) - v_2'(s) = \{\beta(1 - x(s))x(s)\}^2 + \{\gamma x(s)\}^2 \geq 0. \tag{5.4}$$

This inequality and (4.1) imply the square mean global error is lesser, for the D2 distribution, than the corresponding error for the D1 distribution.

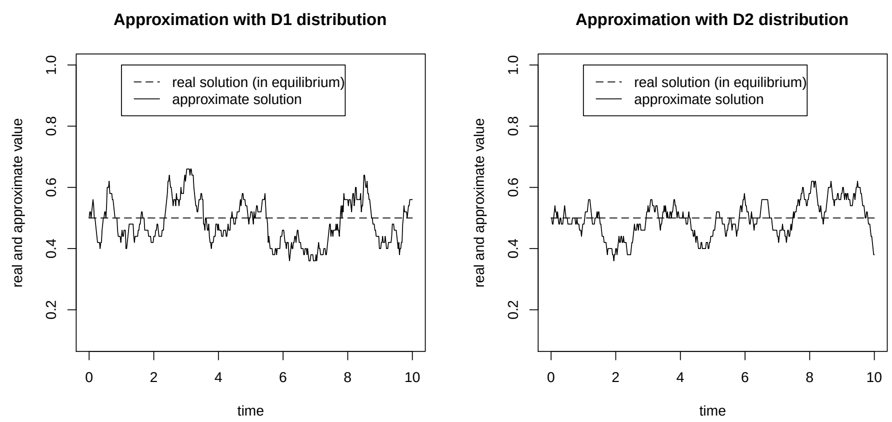


Figure 1. Approximations of the equilibrium solution with $n = 50$, $\beta = 2$ and $\gamma = 1$.

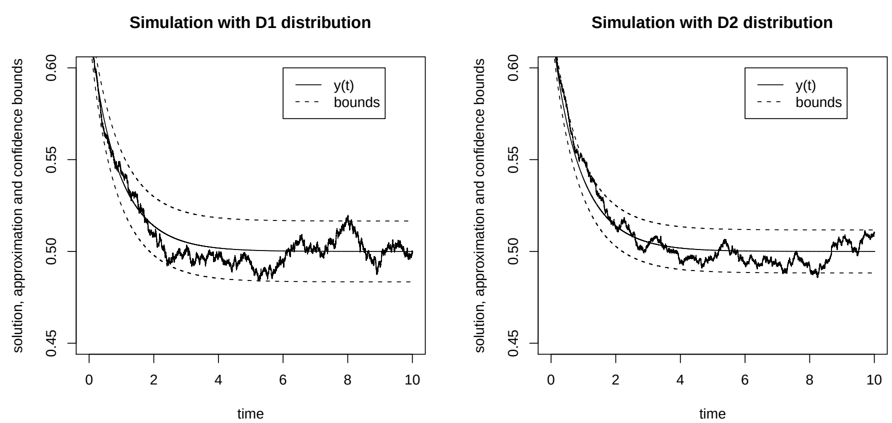


Figure 2. Approximations for the solution and confidence bounds with $n = 7,000$, $\beta = 2$ and $\gamma = 1$.

5.2. Numerical simulations. In the sequel, some of the concepts presented before are applied to the solution to (5.1) with the scheme of approximations labeled by D1 and D2. First, in order to appreciate the difference in variability of the schemes D1 and D2, the equilibrium solution to (5.1) is considered, i.e. $x_0 = 1 - \gamma/\beta$. We simulated both approximations for $T = 10$, $\beta = 2$, $\gamma = 1$ and $n = 50$; see Figure 1.

Let $a(t) = \beta - \gamma - 2\beta x(t)$ and, v_1 and v_2 defined by (5.2) and (5.3), respectively. From (4.1), $Var_1(y(t))$ and $Var_2(y(t))$, the variances of $y(t)$ according to the D1 and D2 distribution, respectively, are given by

$$Var_i(y(t)) = \int_0^t v_i'(u)\, \mathrm{e}^{2\int_u^t a(s)\,\mathrm{d}s}\, \mathrm{d}u,$$

where $v_1'(u) = \{\beta(1-x(u)) + \gamma\}x(u)$ and $v_2'(u) = \{\beta(1-x(u))(1-\beta(1-x(u))x(u)) + \gamma(1-\gamma x(u))\}x(u)$.

Let $0 < \alpha < 1$ and Φ be the cumulative function of a standard normal distribution. Since for the D1 approximation scheme and for large n, $\sqrt{n}(x^n(t) - x(t))$ has approximately normal distribution with mean zero and variance $Var_1(y(t))$, by defining $u_\alpha^\pm(t) = x(t) \pm w_{\alpha/2}\sqrt{Var_1(y(t))/n}$, we have $x^n(t) \in [u_\alpha^-(t), u_\alpha^+(t)]$ with an approximate probability $1-\alpha$ for large values of n. Analogously, $x(t) \pm w_{\alpha/2}\sqrt{Var_2(y(t))/n}$ allow to obtain confidence bands for the scheme of approximation based upon the D2 distribution. In Figure 2, simulations of x^n, starting at $x_0 = 5/8$, are carried out according to the D1 and D2 distributions with $n = 7,000$, $T = 10$, $\beta = 2$ and $\gamma = 1$. In both cases, the bounds u_α^- and u_α^+ are pictured with dash lines for $\alpha = .05$, which gives $w_{\alpha/2} = 1.96$.

5.3. About the appropriate value of n. In order to choose an appropriate value of n that provides a good approximation for the global error to a normal distribution, a goodness-of-fit test is developed for each of the both distributions we are considering.

Let $CHI_i^2(n) = n(\widehat{e}_T^n)^2/Var_i(y(T))$, $(i = 1,2)$. The values of the variances are given by $Var_1(y(T)) = 0.5016656$ and $Var_2(y(T)) = 0.2508328$. For large enough values of n, it is expected $CHI_i^2(n)$, $(i = 1,2)$, has an approximate χ^2-distribution with one degree of freedom, whether is D1 or D2, respectively, the assumed distribution for the model. Let F be the accumulative distribution function corresponding to a χ^2-distribution with one degree of freedom. Consequently, we expect $F(CHI_i^2(n))$ has an approximately uniform distribution for large values of n. We use the goodness-of-fit χ^2-test to evaluate this concordance. For this purpose, we partition the positive part of the real straight line by m subintervals determined by $0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = \infty$, where $t_0, \ldots, t_m$ have been chosen in such a way that $F(t_v) - F(t_{v-1}) = 1/m$. Then, $CHI_i^2(n)$ is simulated repeatedly, recording the number of times that $CHI_i^2(n)$ fall into each subinterval $[t_{v-1}, t_v[$, for each $v = 1, \ldots, m$. By choosing $m = 10$, we have $t_1 = 0.016$, $t_2 = 0.064$, $t_3 = 0.148$, $t_4 = 0.275$, $t_5 = 0.455$, $t_6 = 0.708$, $t_7 = 1.074$, $t_8 = 1.642$ and $t_9 = 2.706$. In addition, for $m = 10$, the expected percentage falling into each subinterval is 10%. A χ^2 test is performed for different values of n.

First, we analyzed the approximate normality of $CHI_1^2(n)$. To this end, $CHI_1^2(n)$ is simulated 10^3 times and the percentages of $CHI_1^2(n)$ falling into these subintervals are determined by the values in Table 1.

Table 1. Percentages of observations of $CHI_2^2(n)$ for the indicated value of n and p-values of the corresponding χ^2 test ($\beta = 2$, $\gamma = 1$ and $T = 10$).

n	$[t_0,t_1[$	$[t_1,t_2[$	$[t_2,t_3[$	$[t_3,t_4[$	$[t_4,t_5[$	$[t_5,t_6[$	$[t_6,t_7[$	$[t_7,t_8[$	$[t_8,t_9[$	$[t_9,t_{10}[$	p-value
40	7.3	0.0	15.9	14.7	13.6	0.0	14.4	7.5	12.4	14.2	0.00015575
50	8.3	15.7	0.0	14.5	13.9	10.1	7.5	8.5	10.8	10.7	0.03813725
60	6.1	13.8	13.1	0.0	11.3	10.8	9.4	15.0	9.9	10.6	0.05308202
70	6.8	11.9	11.8	12.9	0.0	9.0	15.8	9.0	11.1	11.7	0.05671326
80	5.8	15.0	11.0	10.4	9.4	9.2	9.3	11.6	8.8	9.5	0.83829960
90	5.6	12.3	10.1	11.2	8.7	9.1	9.7	13.0	11.2	9.1	0.91180700
100	6.7	12.1	9.0	8.8	9.5	10.5	12.8	13.1	7.6	9.9	0.90154290

Observed percentages of the values of $CHI_2^2(n)$ falling in the corresponding time intervals, for the seven values of n given in Table 1, are organized in the matrix

$$\mathbf{A} = (A_{uv}) = \begin{pmatrix} 7.3 & 0.0 & 15.9 & 14.7 & 13.6 & 0.0 & 14.4 & 7.5 & 12.4 & 14.2 \\ 8.3 & 15.7 & 0.0 & 14.5 & 13.9 & 10.1 & 7.5 & 8.5 & 10.8 & 10.7 \\ 6.1 & 13.8 & 13.1 & 0.0 & 11.3 & 10.8 & 9.4 & 15.0 & 9.9 & 10.6 \\ 6.8 & 11.9 & 11.8 & 12.9 & 0.0 & 9.0 & 15.8 & 9.0 & 11.1 & 11.7 \\ 5.8 & 15.0 & 11.0 & 10.4 & 9.4 & 9.2 & 9.3 & 11.6 & 8.8 & 9.5 \\ 5.6 & 12.3 & 10.1 & 11.2 & 8.7 & 9.1 & 9.7 & 13.0 & 11.2 & 9.1 \\ 6.7 & 12.1 & 9.0 & 8.8 & 9.5 & 10.5 & 12.8 & 13.1 & 7.6 & 9.9 \end{pmatrix}.$$

For the purpose of carrying out the test, the statistics

$$\chi_u^2 = \sum_{v=1}^{10} \frac{(O_{uv} - E_{uv})^2}{E_{uv}} \sim \chi^2(9), \quad u = 1, \ldots, 7,$$

have been defined, where $O_{uv} = 10 \times A_{uv}$ and $E_{uv} = 100$, for $u = 1, \ldots, 7$ and $v = 1, \ldots, 10$. For each $u = 1, \ldots,$, the rejection region is defined as $\{\chi_u^2 > c\}$, where c is chosen in such a way that $\mathbb{P}(\chi_u^2 > c) = .05$.

We compute the p-values associated with the χ^2 test statistic to evaluate the goodness-of-fit of $CHI_1^2(n)$; see Table 1. Since for $n = 60$ the p-value is approximately the significance level .05, we think, the distribution of $CHI_1^2(60)$ is well approximated by the χ^2 distribution with one degree of freedom.

Next, the former test is performed for D2 distribution and 9 values of n are considered. The simulated values of $CHI_2^2(n)$, along with the corresponding p-values for each n, are shown in Table 2.

Table 2. Percentages of observations of $CHI_2^2(n)$ for the indicated value of n and p-values of the corresponding χ^2 test ($\beta = 2$, $\gamma = 1$ and $T = 10$).

n	$[t_0, t_1[$	$[t_1, t_2[$	$[t_2, t_3[$	$[t_3, t_4[$	$[t_4, t_5[$	$[t_5, t_6[$	$[t_6, t_7[$	$[t_7, t_8[$	$[t_8, t_9[$	$[t_9, t_{10}[$	p-value
50	9.7	0.0	20.5	0.0	21.3	0.0	14.9	10.3	8.9	14.4	0.000
80	9.0	17.5	0.0	13.8	15.3	0.0	13.9	9.1	12.7	8.7	0.000
100	7.0	14.5	0.0	15.7	15.3	10.8	9.8	6.7	11.0	9.2	0.01612624
110	6.6	14.0	14.5	0.0	13.0	12.1	9.1	7.6	12.4	10.7	0.04275268
115	15.9	0.0	13.8	11.5	10.7	11.3	9.4	7.1	9.8	10.5	0.06137558
120	8.4	14.8	11.7	0.0	12.5	11.5	8.0	13.9	10.1	9.1	0.07337148
150	6.0	13.4	10.5	11.4	11.6	10.9	8.2	5.6	13.1	9.3	0.68034140
200	5.0	9.5	10.2	11.9	10.8	10.8	13.6	9.7	7.4	11.1	0.82372460
500	10.6	7.9	13.7	6.6	11.3	11.2	8.1	11.3	9.1	10.2	0.91595900

It is obtained for n between 110 and 115 the p-value of the corresponding χ^2 test is approximately the significance level .05. Hence, for these values of n, it is reasonable to assume $CHI_2^2(60)$ has an approximated χ^2 distribution with one degree of freedom.

Although, under D1, x^n has more variability than under D2, the conducted simulation shows that the global error, under D1, attains approximate normality for lower values on n than under D2.

6. Conclusions

A family of discrete-time stochastic processes is presented and it is proven that these processes can be approximate by means of the solution to an ODE. Conversely, these processes may be seen as schemes of approximation for this solution. For this reason, a stochastic version of the global error associated to these schemes are defined and its asymptotic distribution is studied. The uniform convergence in probability, on compact subsets of the positive real numbers, is proven and a central limit theorem for the fluctuations of the stochastic processes is derived. This fact allows us to find confidence bands, where with a preassigned probability the trajectories of the stochastic processes are bounded by these bands. Our results are illustrated by an emblematic model coming from the mathematical literature. Indeed, two discrete time stochastic processes are approximated by the solution of the differential equation corresponding to the SIS epidemic model. Simulations of their trajectories are carried out and compared with the solution of the SIS deterministic model. Moreover, χ^2 tests are carried out to evaluate the goodness of the discretization, in order to obtain approximate normality for the global error.

Acknowledgments

The author wish to thank two anonymous referees for their helpful comments that aided for improving this article. This work was partially supported by project number 124.733/2012 of Dirección de Investigación e Innovación de la Pontificia Universidad Católica de Valparaíso and by FONDECYT project number 1120879.

References

[1] S. Abbasbandy, C. Bervillier, Analytic continuation of taylor series and the boundary value problems of some nonlinear ordinary differential equations, Applied Mathematics and Computation vol. 218 (no. 5) (2011) 2178–2199.

[2] P. Billingsley, Convergence of Probability Measures, Wiley & Sons, New York, 1968.

[3] R.W.R Darling, J.R. Norris, Differential equation approximations for Markov Chain, Probability Surveys Vol. 5 (2008) 37-79.

[4] M. D. M.R. Eslahchi, S. Ahmadi-Asl, The general jacobi matrix method for solving some nonlinear ordinary differential equations, Applied Mathematical Modelling (In Press) (2012) 1–9.

[5] S. K. K. Parand, S. Abbasbandy, J. Rad, A novel application of radial basis functions for solving a model of first-order integro-ordinary differential equation, Communications in Nonlinear Science and Numerical Simulation Vol. 16 (No. 11) (2012) 4250–4258.

[6] R. Fierro, S. Torres, A stochastic scheme of approximation for ordinary differential equations, Electronic Communications in Probability Vol. 13 (2008) 1–9.

[7] R. Fierro, A class of stochastic epidemic models and its deterministic counterpart, Journal of the Korean Statistical Society Vol. 39 (No. 4) (2010) 397–407.

[8] P. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, New York, 1995.

[9] T.G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, Journal of Applied Probability, Vol. 7 (No.1) (1970) 49-58.

[10] T.G. Kurtz, Limit theorems for sequences of jump Markov processes approximating ordinary differential processes, Journal of Applied Probability, Vol. 8 (No.2) (1971) 344-356.

[11] T.G. Kurtz, Approximation of Population Processes, Series in Applied Mathematics 36. Philadelphia: SIAM, 1981.

[12] H. Kushner, P. Dupuis, Numerical Methods for Stochastic Control Problems in Continuous Time, Springer, New York, 1991.

[13] E. Lenglart, Relation de domination entre deux processus, Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques Vol. 13 (No. 2) (1977) 171–179.

[14] R. Rebolledo, La méthode des martingales appliquée á l'etude de la convergence en loi de processus, Mémoires de la Société Mathématique de France Vol. 62 (No. 4) (1979) 1–125.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1047 – 1061

Computing credibility Bonus-Malus premiums using the total claim amount distribution

Emilio Gómez-Déniz* , Agustín Hernández-Bastida† and M. Pilar Fernández-Sánchez‡

Abstract

Assuming a bivariate prior distribution for the two risk parameters appearing in the distribution of the total claim amount when the primary distribution is geometric and the secondary one is exponential, we derive Bayesian premiums which can be written as credibility formulas. These expressions can be used to compute bonus-malus premiums based on the distribution of the total claim amount but not for the claims which produce the amounts. The methodology proposed is easy to perform, and the maximum likelihood method is used to compute the bonus-malus premiums for a real set of automobile insurance data, one that is well known in actuarial literature.

Keywords: Automobile Insurance, Bonus-Malus; Bivariate Distribution, Credibility; Premium

Received 01 : 08 : 2013 : Accepted 27 : 11 : 2013

1. Introduction

Bayesian methods have been successfully applied in actuarial statistics, and have proved to be a good tool for resolving problems related to credibility theory and the setting of insurance premiums. Given a risk group, it is usual to assume that the level of risk of each policy is represented by a risk parameter, or risk profile. It is also assumed that across the group there exists a random variable whose realizations are the values of the risk parameter for policies belonging to that group; its distribution or density function is called the prior distribution or structure function. Most automobile insurance schemes employ Bonus-Malus Systems (BMS). In this context, there is a finite number of classes and the premium applicable depends on the class to which the policyholder belongs. In each period (usually a year), a policyholder's class is determined on the basis

*Department of Quantitative Methods in Economics. University of Las Palmas de Gran Canaria, Spain.
Email: egomez@dmc.ulpgc.es
†Department of Quantitative Methods. University of Granada, Spain.
‡Department of Quantitative Methods. University of Granada, Spain.

of that assigned for the previous period and on the number of claims made during the period. The main purpose of a BMS is to decrease the premiums for good risks and to increase them for bad ones. With Bayesian methodology, this is achieved by dividing a posterior expectation by a prior expectation according to an estimate derived by means of an appropriate loss function (see Lemaire, 1979, 1985, 1995; Gómez-Déniz et al., 2002; Sarabia et al., 2004; Denuit et al., 2007, among others).

Nevertheless, it is obvious that not all accidents produce the same individual claim size and thus it does not seem fair to penalize all policyholders in the same way when they present a claim. In other words, when the bonus-malus premium is based only on the number of claims a policyholder who has an accident with an individual claim size of 100$ is penalized by the same amount as if the accident had produced an individual claim size of 500$. As different claims produce different claim amounts, it would seem that the best way to build a BMS would be based on both the number of claims and on the individual claim size. As Lemaire (2004) points out, when the claim amount is not incorporated into the bonus-malus premium this implies an assumption of independence between the variables "number of claims" and "claim amount", an assumption that is open to question.

In recent years, attempts have been made to include factors other than the number of claims in calculating bonus-malus premiums. Thus, Frangos and Vrontos (2001) and Mert and Saykan (2005) introduced a model where the number of claims and the individual claim size were used jointly to compute the bonus-malus premiums. Based on the independence assumption assumed in the collective risk model between these two random variables, they computed the premium by multiplying the bonus-malus premiums based only on the number of claims with the bonus-malus premiums based only on the individual claim size. Their empirical results show there is a positive correlation between these two random variables, and thus the assumption of some kind of dependence between them should be taken into account in calculating bonus-malus premiums.

If we wish to replace the distribution of the number of claims by the distribution of the total claim amount, this will depend on two parameters, one related to the random variable "number of claims" and the other related to the random variable "claim cost". It is possible to transfer a relation of dependence between these risk profiles, by assuming that both profiles fit a joint bivariate prior distribution.

Apart from looking for some kind of dependence, the bivariate prior distribution can be justified in the following manner. The aim of the actuary is to design a tariff system that will distribute the exact weight of each risk fairly within the portfolio when policyholders present different risks. For instance, in the automobile insurance market, the first approach to solving this problem, called tariff segmentation, consists in dividing policyholders into homogeneous classes according to certain variables believed to be influencing factors (a priori factors), such as the model and use of the car, the age and sex of the driver, the duration of the driving licence, etc. Once the actuary has classified policyholders, the premium can be established for each type of risk. However, some factors cannot be measured or introduced into the rates to calculate premiums according to tariff-segmentation methods. Consequently, heterogeneity continues to exist in every class defined with a priori factors. Some of these unmeasured or unknown characteristics probably have a significant effect on the number of claims and also on the individual claim size; for instance, in automobile insurance, swiftness of reflexes, knowledge of the Highway Code or the behaviour patterns of the driver. Given that many claims could be explained by these hidden features, they should be included in the tariff system. This is the goal of experience rating or credibility theory, the underlying idea of which is that past experience reveals information about hidden features.

In this paper, we assume a bivariate prior distribution for the two risk parameters appearing in the distribution of the total claim amount when the distribution of the random variable number of claims (primary distribution) is geometric and the distribution of the random variable individual claim size (secondary distribution) is exponential. This allows us to use the unconditional distribution of the total claim amount to compute the premiums, which can then be written as a credibility formula. These premiums are then used to obtain the bonus-malus premiums, based on the distribution of the total claim amount and not only on the claims which produced the amounts. The maximum likelihood method is used to estimate the parameters of the distribution in a real data set concerning automobile insurance and well known in actuarial literature.

The rest of this paper is structured as follows. Section **??** presents the basic collective risk model based on the geometric and the exponential distribution as the primary and the secondary distribution, respectively. The bivariate prior distribution is presented in Section **??**, where we also show the marginal and the unconditional distribution of the claim size. Credibility premiums are obtained in Section **??** and the parameters of the unconditional distribution of the claim amount are estimated in Section **??**. A numerical application with a real data set is presented in Section **??** and the main conclusions are drawn in the last Section.

2. The basic model

One of the main objectives of risk theory is to model the distribution of the aggregate claim amount for portfolios of policies, so that the insurance firm can take decisions taking into account just two aspects of the insurance business: the number of claims and the individual claim size. Therefore, the total claim amount over a fixed time period is modelled by considering the number of claims and the individual claim size separately. In this paper, we assume that the premiums in a bonus-malus system should be computed by taking into account both the number of claims and the individual claim size.

In the collective risk theory, the random variable of interest is the aggregate claim defined by $X = \sum_{i=1}^{N} X_i$, where N is the random variable denoting the number of claims and X_i, for $i = 1, 2, ...$ is the random variable denoting the individual claim size of the i-th claim. Assuming that $X_1, X_2, \ldots,$ are independent and identically distributed random variables which are also independent of the random variable number of claims N, it is well-known (see Klugman et al. (2008) and Rolski et al. (1999), among others) that the probability density function of the aggregate claim (total claim amount) is given by $f_X(x) = \sum_{n=0}^{\infty} p_n f^{n^*}(x)$, where p_n denotes the probability of n claims (primary distribution) and $f^{n^*}(x)$ is the n-th fold convolution of $f(x)$, the probability density function of the claim amount (secondary distribution).

In automobile insurance, when the portfolio is considered to be heterogeneous, all policyholders have a constant but unequal underlying risk of having an accident. That is, the expected number of claims varies from policyholder to policyholder. As the mixed Poisson distributions have thicker tails than the Poisson distribution, the former provide a good fit to claim frequency data when the portfolio is heterogeneous. Frangos and Vrontos (2001), Gómez-Déniz (2002), Mert and Saykan (2005), among many others, consider the Poisson parameter, i.e. the expected number of claims, to follow a Gamma distribution. In this case, the unconditional distribution of the number of claims follows a negative binomial distribution. The advantage of this model is that the distribution of the total claim amount can be obtained in closed form expression when the secondary distribution is assumed to be exponential. Other models considered in the actuarial literature are the Poisson-inverse Gaussian distribution (Willmot (1987)) and the negative binomial-inverse Gaussian distribution (Gómez-Déniz et al. (2008)). Both models provide a good fit to the claim frequency data, and a recursive computation of the total claim amount can be obtained using Panjer's algorithm or a simple modification.

Assuming that the number of claims is represented by random variable N and that it follows a Poisson distribution with parameter $\lambda > 0$ denoting the differing underlying risk of each policyholder reporting a claim. Assume, moreover, that λ is distributed according to the exponential distribution with parameter $\theta_1/(1-\theta_1)$, with $0 < \theta_1 < 1$, i.e. $\pi(\lambda) \propto \exp\left(-\frac{\theta_1 \lambda}{1-\theta_1}\right)$, where $\pi(\lambda)$ represents the prior distribution of λ. It is a simple exercise to show that the unconditional distribution of the number of claims is given by

$$\Pr(N = n) = \theta_1(1-\theta_1)^n, \quad n = 0, 1, \ldots,$$

and therefore a geometric distribution with parameter θ_1.

Assuming that the individual claim size follows an exponential distribution (secondary distribution) with parameter $\theta_2 > 0$, the n-th fold convolution of exponential distribution has a closed form that is given as follows (see Klugman et al. (2008) and Rolski et al. (1999))

$$f^{*n}(x) = \frac{\theta_2^n}{(n-1)!} x^{n-1} e^{-\theta_2 x}, \ n = 1, 2, \ldots$$

i.e. it is a gamma distribution with shape parameter n and scale parameter θ_2. Now, it is easy to see that the probability density function of the random variable $X = \sum_{i=1}^{N} X_i$ is given by

$$f_X(x|\theta_1, \theta_2) = \begin{cases} \theta_1, & x = 0, \\ \theta_1(1-\theta_1)\theta_2 \exp(-\theta_1\theta_2 x), & x > 0. \end{cases} \tag{2.1}$$

Observe that the probability density function of the claim amount has a jump of size θ_1 at the origin.

3. A suitable bivariate distribution

In this section we introduce a new continuous probability density function that will be used to derive, by mixing, the unconditional probability density function of the total claim amount in (**??**) and also to compute the bonus-malus premiums proposed in this paper.

We begin by introducing the new continuous bivariate probability density function, as follows. It can be shown straightforwardly that

$$f(x, y) = \frac{\sigma^\gamma}{B(\alpha-\gamma, \beta)\Gamma(\gamma)} x^{\alpha-1}(1-x)^{\beta-1} y^{\gamma-1} \exp(-\sigma x y), \tag{3.1}$$

for $0 < x < 1$, $y > 0$, $\alpha > 0$, $\beta > 0, \gamma > 0, \sigma > 0$ and $\alpha > \gamma$ is a proper bivariate probability density function. In (**??**) we have that

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

is the gamma function and $B(z_1, z_2)$ is the beta function given by

$$B(z_1, z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1}\, dt.$$

To the best of our knowledge, the bivariate distribution presented here has not been previously addressed in statistical literature.

Some computations provide that the distribution is unimodal with modal value at the point

$$\begin{aligned} x &= \frac{\alpha-\gamma}{\alpha+\beta-\gamma-1}, \\ y &= \frac{(\gamma-1)(\alpha+\beta-\gamma-1)}{\sigma(\alpha-\gamma)}. \end{aligned}$$

Now by a straightforward calculation we see that

$$E(XY) = \frac{\gamma}{\sigma}. \tag{3.2}$$

The marginal distribution of X and Y, which can be obtained by integrating (**??**) with respect to y and x, respectively, can be shown to be a known univariate distribution. Thus, the marginal distribution of X is a beta distribution with parameters $\alpha-\gamma$ and β, i.e.

$$f_X(x) = \frac{1}{B(\alpha-\gamma,\beta)} x^{\alpha-\gamma-1}(1-x)^{\beta-1}. \tag{3.3}$$

The marginal distribution of Y is given by

$$f_Y(y) = \frac{\sigma^\gamma \Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha+\beta)B(\alpha-\gamma,\gamma)} y^{\gamma-1}\, {}_1F_1(\alpha,\alpha+\beta,-\sigma y), \tag{3.4}$$

where ${}_1F_1(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function, also called Kummer's function, given by

$${}_1F_1(m,n,z) = \sum_{k=0}^{\infty} \frac{(m)_k z^k}{(n)_k k!},$$

and $(m)_j = \Gamma(m+j)/\Gamma(m)$, $j \geq 1$, $(m)_0 = 1$ is the Pochhammer symbol.

Using Kummer's first theorem we have that (**??**) can be rewritten as

$$f_Y(y) = \frac{\sigma^\gamma \Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha+\beta)B(\alpha-\gamma,\gamma)} y^{\gamma-1} e^{-\sigma y}\, {}_1F_1(\beta,\alpha+\beta,\sigma y). \tag{3.5}$$

Expression (**??**) is reminiscent of the generalized exponential distribution given in Bhattacharya (1966), see expression (3.1) in this paper.

Since

$$E_Y(Y) = \frac{\gamma(\alpha+\beta-\gamma-1)}{\sigma(\alpha-\gamma-1)}, \tag{3.6}$$

by using (**??**) together with (**??**) and the mean of the distribution given in (**??**) we obtain the covariance of (**??**), which is given by

$$cov(X,Y) = \frac{\beta\gamma}{\sigma(\alpha+\beta-\gamma)(\gamma-\alpha+1)}, \tag{3.7}$$

which admits correlation of any sign. Thus, we have

$$cov(X,Y) \begin{cases} >0 & \text{if } 0<\alpha-\gamma<1, \\ >0 & \text{if } \alpha-\gamma>1. \end{cases}$$

Let us now assume that the two parameters of the distribution of the total claim amount in (**??**) are random and that they follow a bivariate prior distribution as in (**??**), i.e. we have that

$$\pi(\theta_1,\theta_2)=\frac{\sigma^\gamma}{B(\alpha-\gamma,\beta)\Gamma(\gamma)}\theta_1^{\alpha-1}(1-\theta_1)^{\beta-1}\theta_2^{\gamma-1}\exp(-\sigma\theta_1\theta_2), \tag{3.8}$$

for $0<\theta_1<1$, $\theta_2>0$, $\alpha>0$, $\beta>0, \gamma>0, \sigma>0$ and $\alpha>\gamma$. The unconditional distribution of the total claim amount in (**??**) can be obtained by mixing, by computing the following integral

$$f_X(x|\alpha,\beta,\gamma,\sigma)=\int_0^\infty\int_0^1 f_X(x|\theta_1,\theta_2)\pi(\theta_1,\theta_2)\,d\theta_1\,d\theta_2.$$

Some algebra provides the following probability density function for the unconditional distribution of the total claim amount.

$$f(x|\alpha,\beta,\gamma,\sigma)=\begin{cases}\dfrac{\alpha-\gamma}{\alpha+\beta-\gamma}, & x=0,\\[2ex] \dfrac{\beta\gamma\sigma^\gamma}{\alpha+\beta-\gamma}\dfrac{1}{(x+\sigma)^{\gamma+1}}, & x>0,\end{cases} \tag{3.9}$$

which is obviously a two piece distribution with a jump of size $\frac{\alpha-\gamma}{\alpha+\beta-\gamma}$ at the origin.

Moments of order r of (**??**) are as follows:

$$E(X^r)=\frac{\beta\sigma^r r!\Gamma(\gamma-r)}{(\alpha+\beta-\gamma)\Gamma(\gamma)},\quad \gamma>r. \tag{3.10}$$

In particular, we have that

$$\begin{aligned} E(X) &= \frac{\beta\sigma}{(\alpha+\beta-\gamma)(\gamma-1)},\quad \gamma>1,\\ E(X^2) &= \frac{2\beta\sigma^2}{(\alpha+\beta-\gamma)(\gamma-1)(\gamma-2)},\quad \gamma>2,\end{aligned}$$

from which we can obtain the variance of the distribution, given by

$$var(X)=\frac{\beta\sigma^2(2\alpha(\gamma-1)+\beta-2(\gamma-1)\gamma)}{(\alpha+\beta-\gamma)^2(\gamma-2)(\gamma-1)^2},\quad \gamma>2.$$

4. Credibility premiums

When the premiums are based only on the number of claims, the distribution to be considered is, in this case, the geometric distribution with parameter $0<\theta_1<1$. Suppose now that the prior distribution on θ_1 is the beta distribution given in (**??**). Given a sample information $n_1,\dots,n_t$ the posterior distribution is a beta distribution with parameters $\alpha-\gamma+t$ and $\beta+t\bar{n}$ and it is simple to see that the unconditional distribution of the number of claims is given by

$$\Pr(N=n)=\frac{B(\alpha-\gamma+1,\beta+n)}{B(\alpha-\gamma,\beta)}, \tag{4.1}$$

which is the geometric–beta distribution.

In the actuarial context, the premium charged to a policyholder is computed on the basis of the past claims made and on that of the accumulated past claims of the corresponding portfolio of policyholders. To obtain an appropriate formula for this, various methods have been proposed, mostly in the field of Bayesian decision methodology. The procedure for premium calculation is modelled as follows. The number of claims made with respect to a given contract in a given period is specified by a random variable X following a probability density function $f(x|\theta)$ depending on an unknown risk parameter θ. A premium calculation principle (Gómez-Déniz et al. (2006) and Heilmann (1989)) assigns to each risk parameter θ a premium within the set $P \in \mathbb{R}$, the action space. Let $L : \Theta \times P \to \mathbb{R}$ be a loss function that assigns to any $(\theta, P) \in \Theta \times P$ the loss sustained by a decision-maker who takes the action P and is faced with the outcome θ of a random experience. The premium must be determined such that the expected loss is minimized.

From this parameter, the unknown premium $\mu(\theta)$, called the risk premium, can be obtained by minimizing the expected loss $E_f\left[L(\theta, P)\right]$. L is usually taken as the weighted squared-error loss function, i.e. $L(a, x) = h(x)(x - a)^2$. Using different functional forms for $h(x)$ different premium principles are obtained. For example, for $h(x) = 1$ we obtain the net premium principle (Heilmann (1989), Gerber (1979) and Klugman et al. (2008); among others). For a review of the net premium and the different premiums defined in the actuarial setting, see Bühlmann and Gisler (2005), Gerber (1979), Gómez et al. (2002, 2006), Heilmann (1989) and Rolski et al. (1999).

If experience is not available, the actuary computes the collective premium, μ, which is given by minimizing the risk function, i.e. minimizing $E_\pi\left[L(\mu(\theta), \theta)\right]$, where $\pi(\theta)$ is the prior distribution on the unknown parameter θ. On the other hand, if experience is available, the actuary takes a sample $\mathbf{x}$ from the random variables $X_i,\ i = 1, 2, \ldots, t$, assuming X_i i.i.d., and uses this information to estimate the unknown risk premium $\mu(\theta)$, through the Bayes premium μ^*, obtained by minimizing the Bayes risk, i.e. minimizing $E_{\pi_{\mathbf{x}}}\left[L(\mu(\theta), \theta)\right]$. Here, $\pi_{\mathbf{x}}$ is the posterior distribution of the risk parameter, θ, given the sample information $\mathbf{x}$.

Thus, in our case, if $L(x, a) = (x - a)^2$, the net risk, collective and Bayes premiums are given by

$$\begin{aligned} \mu_C(\theta_1) &= \frac{1-\theta_1}{\theta_1}, \\ \mu_C &= \frac{\beta}{\alpha-\gamma-1}, \\ \mu_C^* &= \frac{\beta+k}{\alpha-\gamma+t-1} = Z(t)\bar{n} + (1 - Z(t))\mu_C, \end{aligned} \tag{4.2}$$

where the credibility factor is given by $Z(t) = t/(\alpha - \gamma + t - 1)$ and $\bar{n}$ is the sample mean based only on the claim frequency observed. The subscript C indicates that the premiums are based on the number of claims.

Now suppose that the practitioner chooses to compute the premium according to the individual claim size, assuming that this follows an exponential distribution with parameter $\theta_2 > 0$ and that the prior distribution on θ_2 is the distribution given in (**??**) with $\beta \to 0$ and $\alpha = 1$. In this case, (**??**) reduces to the gamma distribution with shape parameter γ and scale parameter σ. Given a sample information $x_1, \ldots, x_t$ the posterior distribution is a gamma distribution with parameters $\gamma + t$ and $\sigma + t\bar{x}$. Again, it is simple to see that the unconditional distribution of the individual claim size is given by

$$f(x) = \frac{\gamma\sigma^\gamma}{(\sigma + x)^{\gamma+1}},$$

and that the risk, collective and Bayes premiums are as follows:

$$\begin{aligned} \mu_{CC}(\theta_2) &= \frac{1}{\theta_2}, \\ \mu_{CC} &= \frac{\sigma}{\gamma}, \\ \mu_{CC}^* &= \frac{\sigma+t\bar{x}}{\gamma+t} = Z(t)\bar{x} + (1 - Z(t))\mu, \end{aligned} \tag{4.3}$$

where the credibility factor is given by $Z(t) = t/(\gamma + t)$ and $\bar{x}$ is the sample mean based only on the size observed. The subscript CC denotes that the premiums are based on the individual claim size.

Frangos and Vrontos (2001) and Mert and Saykan (2005) introduced a model where the number of claims and the individual claim size were used jointly to compute the bonus-malus premiums. According to the independence assumption assumed in the collective risk model between the two random variables, they computed the premium by multiplying the bonus-malus premiums based only on the number of claims by the bonus-malus premiums based only on the individual claim size, i.e. multiplying (**??**) by (**??**).

The most reasonable model consists in working with both random variables but not in a separate way. To do so, let x_i, $i = 1, 2, \ldots, t$ be independent and identically distributed random variables following the probability density function (**??**), i.e.

$$f(x_1, \ldots, x_t) = \theta_1^t \theta_2^t (1 - \theta_1)^t \exp(-t\bar{x}\theta_1\theta_2),$$

provided that $x_i > 0,\ i = 1, 2, \ldots, t$. Let us suppose that (θ_1, θ_2) follows the prior distribution $\pi(\theta_1, \theta_2)$ given in (**??**), then the posterior distribution of (θ_1, θ_2) given the sample information $(x_1, \ldots, x_t)$ is of the same form as in (**??**) with the updated parameters $(\alpha^*, \beta^*, \gamma^*, \sigma^*)$ given by

$$\begin{aligned} \alpha^* &= \alpha + t, \\ \beta^* &= \beta + t, \\ \gamma^* &= \gamma + t, \\ \sigma^* &= \sigma + \kappa, \end{aligned}$$

where $\kappa = \sum_{i=1}^t x_i$.

When $x_i = 0,\ i = 1, 2, \ldots, t$ then the posterior distribution has the following updated parameters

$$\begin{aligned} \alpha^* &= \alpha + t, \\ \beta^* &= \beta, \\ \gamma^* &= \gamma, \\ \sigma^* &= \sigma, \end{aligned}$$

Now, denoting the unknown risk premium by $\mu(\theta_1, \theta_2) = \mu(\Theta)$, and again using the square–error loss function, the net risk, collective and Bayes premiums are given by

$$\mu(\Theta) = \int x f(x|\Theta) dx, \tag{4.4}$$

$$\mu = \int \mu(\Theta)\pi(\Theta) d\Theta, \tag{4.5}$$

$$\mu^* = \int \mu(\Theta)\pi(\Theta|\underline{x}) d\Theta, \tag{4.6}$$

respectively.

As mentioned above, it is clear that under the model Assumed, the net risk premium (**??**) is given by

$$\mu(\Theta) = E(X|\Theta) = \frac{1 - \theta_1}{\theta_1\theta_2}$$

while the net collective premium in (**??**) is described by

$$\mu = \int_0^\infty \int_0^1 \mu(\Theta)\pi(\theta_1, \theta_2)\, d\theta_1\, d\theta_2 = \frac{\beta\sigma}{(\alpha + \beta - \gamma)(\gamma - 1)}, \quad \gamma > 1. \tag{4.7}$$

Finally, the net Bayes premium in (**??**) is given by

$$\mu^* = \frac{(\beta+t)(\sigma+t\bar{x})}{(\alpha+\beta-\gamma+t)(\gamma+t-1)}, \quad \gamma > 1, \tag{4.8}$$

for $x_i > 0,\ i = 1, 2, \ldots, t$. When $x_i = 0,\ i = 1, 2, \ldots, t$, the net Bayes premium is

$$\mu^* = \frac{\beta\sigma}{(\alpha+\beta-\gamma+t)(\gamma-1)}, \quad \gamma > 1, \tag{4.9}$$

Observe that (**??**) can be rewritten in the two following ways. Firstly, some simple computations provide that

$$\mu^* = \mathcal{H}(\alpha,\beta,\gamma,t)\mu^*_{CC},$$

where

$$\mathcal{H}(\alpha,\beta,\gamma,t) = \frac{(\beta+t)(\gamma+t)}{(\alpha+\beta-\gamma+t)(\gamma+t-1)}.$$

And secondly,

$$\mu^* = Z(t)h_1(\bar{x}) + (1-Z(t))h_2(\mu),$$

where the credibility factor $Z(t)$ is given by

$$Z(t) = \frac{t}{t+\alpha+\beta-\gamma} \tag{4.10}$$

and the functions $h_1(\cdot)$ and $h_2(\cdot)$ are given by

$$\begin{aligned} h_1(x) &= \frac{(\beta+t)x+\sigma}{\gamma+t-1}, \\ h_2(x) &= \frac{\gamma-1}{\gamma+t-1}x. \end{aligned}$$

When $t \to \infty$ we have that, since $h_1(\bar{x}) \to \bar{x}$ and $Z(t) \to 1$, then $\mu^* \to \bar{x}$ and when $t \to 0$ it is easy to see that $\mu* \to \mu$. Thus, it is reasonable to assume that when the sample size tends to infinity the Bayes premium converges to the sample mean, and it converges to the collective premium when the sample size tends to zero.

The credibility factor for expression (**??**) is as in (**??**) where now $h_1(x) = 0$ and $h_2(x) = (\gamma-1)x$.

5. Inference

Moment estimators can be obtained by equating the sample moments to the population moments in (**??**). Furthermore, the parameters of the unconditional distribution of the total claim amount can be estimated via maximum likelihood. To do so, consider a random sample $\{x_1, x_2, \ldots, x_t\}$. The likelihood function can be written as

$$f(x_1,\ldots,x_t|\alpha,\beta,\gamma,\sigma) = \left(\frac{\alpha-\gamma}{\alpha+\beta-\gamma}\right)^{t_0}\left(\frac{\beta\gamma\sigma^\gamma}{\alpha+\beta-\gamma}\right)^{t^*}\prod_{x_i>0}\frac{1}{(x_i+\sigma)^{\gamma+1}}, \tag{5.1}$$

where t_0 is the number of zero-observations and $t^* = t - t_0$ is the number of non-zero sample observations, where t is the sample size. Finally $\prod_{x_i>0}$ denotes the product over the t^* non-zero observations.

Taking logarithms of (**??**) we have the log-likelihood equation given by

$$\ell \equiv \ell(\alpha,\beta,\gamma,\sigma|x_1,\dots,x_t) = t_0 \log\left(\frac{\alpha-\gamma}{\alpha+\beta-\gamma}\right) + t^* \log\left(\frac{\beta\gamma\sigma^\gamma}{\alpha+\beta-\gamma}\right) - (\gamma+1)\sum_{x_i>0}\log(x_i+\sigma). \tag{5.2}$$

Now, differentiating (**??**) with respect to the four parameters in turn, and equating to zero, we obtain the maximum likelihood estimating equations given by

$$\frac{\partial \ell}{\partial \alpha} = \frac{t_0}{\alpha-\gamma} - \frac{t}{\alpha+\beta-\gamma} = 0, \tag{5.3}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{t^*}{\beta} - \frac{t}{\alpha+\beta-\gamma} = 0, \tag{5.4}$$

$$\frac{\partial \ell}{\partial \gamma} = \frac{t_0}{\alpha-\gamma} - \frac{t^*}{\gamma} + \frac{t}{\alpha+\beta-\gamma} - t^*\log\sigma + \sum_{x_i>0}\log(x_i+\sigma) = 0, \tag{5.5}$$

$$\frac{\partial \ell}{\partial \sigma} = \frac{t^*}{\sigma} - (\gamma+1)\sum_{x_i>0}\frac{1}{x_i+\sigma} = 0. \tag{5.6}$$

The second partial derivatives are as follows:

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha^2} &= \frac{t}{(\alpha+\beta-\gamma)^2} - \frac{t_0}{(\alpha-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \alpha\beta} = \frac{t}{(\alpha+\beta-\gamma)^2},\\
\frac{\partial^2 \ell}{\partial \alpha \partial \gamma} &= \frac{t_0}{(\alpha-\gamma)^2} - \frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \alpha \partial \sigma} = 0,\\
\frac{\partial^2 \ell}{\partial \beta^2} &= \frac{t}{(\alpha+\beta-\gamma)^2} - \frac{t^*}{\beta^2}, \quad \frac{\partial^2 \ell}{\partial \beta\gamma} = -\frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \beta \partial \sigma} = 0,\\
\frac{\partial^2 \ell}{\partial \gamma^2} &= \frac{t_0}{(\alpha-\gamma)^2} + \frac{t^*}{\gamma^2} + \frac{t}{(\alpha+\beta-\gamma)^2}, \quad \frac{\partial^2 \ell}{\partial \gamma \partial \sigma} = \sum_{x_i>0}\frac{1}{x_i+\sigma},\\
\frac{\partial^2 \ell}{\partial \sigma^2} &= -\frac{t^*}{\sigma^2} + (\gamma+1)\sum_{x_i>0}\frac{1}{(x_i+\sigma)^2}.
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
E\left(\frac{1}{X+\sigma}\right) &= \frac{\alpha+(-1+\alpha+\beta)\gamma-\gamma^2}{(\alpha+\beta-\gamma)(1+\gamma)\sigma},\\
E\left[\frac{1}{(X+\sigma)^2}\right] &= \frac{(-2+\beta-\gamma)\gamma+\alpha(2+\gamma)}{(\alpha+\beta-\gamma)(2+\gamma)\sigma^2}.
\end{aligned}$$

Therefore, Fisher's information matrix (not reproduced here) can be obtained easily, in closed form expression.

6. An application to a real data set

In order to compare the premiums based only on the number of claims with the premiums obtained when the total claim amount distribution is used, we examined a data set based on one-year vehicle insurance policies taken out in 2004 or 2005. This data set is available on the website of the Faculty of Business and Economics, Macquarie University (Sydney, Australia), see also Jong and Heller (2008). The first 100 observations of this data set are shown in Table **??**, with the following elements: from left to right, the policy number, the number of claims and the size of the claims. The total portfolio contains 67856 policies of which 4624 have at least one claim. Some descriptive statistics for this data set are shown in Table **??**. It can be seen that the standard deviation is very large for the size of the claims, which means that a premium based only on the mean size of the claims is not adequate for computing the bonus-malus premiums. The covariance between the claims and sizes is positive and takes the value 141.574.

Table 1. First 100 observations of the data set

1	0	0	21	0	0	41	2	1811.71	61	0	0	81	0	0
2	0	0	22	0	0	42	0	0	62	0	0	82	0	0
3	0	0	23	0	0	43	0	0	63	0	0	83	0	0
4	0	0	24	0	0	44	0	0	64	0	0	84	0	0
5	0	0	25	0	0	45	0	0	65	1	5434.44	85	0	0
6	0	0	26	0	0	46	0	0	66	1	865.79	86	0	0
7	0	0	27	0	0	47	0	0	67	0	0	87	0	0
8	0	0	28	0	0	48	0	0	68	0	0	88	0	0
9	0	0	29	0	0	49	0	0	69	0	0	89	0	0
10	0	0	30	0	0	50	0	0	70	0	0	90	0	0
11	0	0	31	0	0	51	0	0	71	0	0	91	0	0
12	0	0	32	0	0	52	0	0	72	0	0	92	0	0
13	0	0	33	0	0	53	0	0	73	0	0	93	0	0
14	0	0	34	0	0	54	0	0	74	0	0	94	0	0
15	1	669.51	35	0	0	55	0	0	75	0	0	95	0	0
16	0	0	36	0	0	56	0	0	76	0	0	96	1	1105.77
17	1	806.61	37	0	0	57	0	0	77	0	0	97	0	0
18	1	401.80	38	0	0	58	0	0	78	0	0	98	0	0
19	0	0	39	0	0	59	0	0	79	0	0	99	1	200
20	0	0	40	0	0	60	0	0	80	0	0	100	0	0

Figure **??** shows the complete number of claims and the total claim amount concerning these claims. It can be seen that the larger claim values appear in the case of single claims and that these values fall with larger numbers of claims. It is probable that a first severe accident encourages the driver to be more careful, which tends to reduce the size of the claims in future accidents. For this reason, we believe the bonus-malus premiums should not be based only on the number of claims but also on their size.

LISTPLOT-eps-converted-to.pdf

Figure 1. Number of claims and their costs

We used (**??**) to estimate the α and β parameters of this distribution when only the number of claims was used, and assumed that $\alpha \equiv \alpha - \gamma$. The maximum likelihood method does not provide a solution in this case, and so $\alpha = 1/\beta$ was assumed, which produced $\widehat{\beta} = 0.2528$ and -18684.10 for the value of the maximum of the log-likelihood function. The bonus-malus premiums (BMP) are computed according to the expression

$$\mathrm{BMP} = \frac{\beta + k}{\alpha + t - 1}\,\frac{\alpha - 1}{\beta},$$

Table 2. Some descriptive data of claims and claim size for the data set

	Number of claims	Total claim amount
Mean	0.072	137.27
Standard deviation	0.278	1056.30
min	0	0
max	4	55922.10

where $k = t\bar{n}$. The resulting bonus-malus premiums are shown in Table **??**.

Now, using the expressions given in (**??**), (**??**), (**??**) and (**??**) we computed the maximum likelihood estimates of the parameters when the claim amount distribution is used to compute the bonus-malus premiums. In order to simplify the computations the values of the total claim amounts have been divided by 1000. These are given by $\widehat{\alpha} = 2.4282$, $\widehat{\beta} = 0.0299$, $\widehat{\gamma} = 2.0465$ and $\widehat{\sigma} = 2.2051$. Now, the value of the maximum of the log-likelihood function is -24111.80 and the estimated value of the covariance, using expression (**??**), is 0.1072. Observe that for these estimates the net collective premium based on both, the number of claims and the individual claim size, which is given in (**??**), is provided by $1000 \times 0.153068 = 153.068$, the latter value is close to the sample mean appearing in Table **??**.

Let us now compute the bonus-malus premiums using the expression

$$\mathrm{BMP} = \frac{(\beta + t)(\sigma + \kappa)}{(\alpha + \beta - \gamma + t)(\gamma + t - 1)} \frac{(\alpha + \beta - \gamma)(\gamma - 1)}{\beta\sigma},$$

for $\kappa = t\bar{x} > 0$. When $\kappa = 0$ the bonus-malus premiums are given by

$$\mathrm{BMP} = \frac{\beta\sigma}{(\alpha + \beta - \gamma + t)(\gamma - 1)} \frac{(\alpha + \beta - \gamma)(\gamma - 1)}{\beta\sigma}.$$

Table **??** shows the bonus-malus premiums obtained with the aggregate model, taking into account the number of claims and the individual claim size.

Observe that the premiums based only on the number of claims and on the total claim amount have several levels of premiums, but these levels have different meanings. Although the first is based on k and the second on κ we can consider both levels used by the insurance firm to move a policyholder from one column to another, i.e. to move the policyholder from one class to another. The first column is usually termed the bonus class and the other, the malus class.

Tables **??** and **??** show that the bonus-malus premiums under the model based on the total claim amount distribution are slightly lower for the bonus class and larger for the classes $k = 1, 2$ and 3, in comparison with the bonus-malus premiums based only on the number of claims. It seems reasonable that policyholders with no claims should pay less, taking into account that those reporting claims are now going to pay more. It could be said that the new bonus-malus system is very generous to drivers in the bonus class and very strict with those in the malus classes. The drivers in the bonus class, for the first claim free year, will receive 70.94% of the basic premium, while drivers who report one accident in the first year will have to pay a malus of 695.400% of the basic premium. It might be thought that this is dangerous for the insurance firm, because most policyholders would look for another company with more competitive prices, but it should be recalled that most of the policyholders in the portfolio do not make a claim.

Table 3. Bonus-malus premiums based only on the frequency component and the net premium principle

Year	Number of claims, k					
	0	1	2	3	4	5
0	1.00000					
1	0.74712	3.70161	6.65609	9.61058	12.5651	15.5196
2	0.59632	2.95449	5.31265	7.67081	10.0290	12.3871
3	0.49617	2.45831	4.42044	6.38257	8.34470	10.3068
4	0.42483	2.10482	3.78481	5.46480	7.14480	8.82479
5	0.37142	1.84021	3.30900	4.77780	6.24659	7.71538
6	0.32994	1.63471	2.93947	4.24423	5.54899	6.85376
7	0.29679	1.47049	2.64418	3.81787	4.99156	6.16525
8	0.26970	1.33625	2.40280	3.46935	4.53589	5.60244
9	0.24714	1.22447	2.20180	3.17913	4.15646	5.13379
10	0.22806	1.12995	2.03184	2.93373	3.83561	4.73750

Table 4. Bonus-malus premiums based on the severity component and the net premium principle

Year	Total claim amount, κ					
	0	1	2	3	4	5
0	1000.00					
1	290.57	7953.99	10435.70	12917.40	15399.10	17880.80
2	169.98	6166.55	8090.54	10014.50	11938.5	13862.50
3	120.12	4898.92	6427.41	7955.90	9484.40	11012.90
4	92.88	4040.46	5301.11	6561.76	7822.41	9083.06
5	75.71	3431.31	4501.90	5572.48	6643.07	7713.66
6	63.90	2979.23	3908.77	4838.31	5767.84	6697.38
7	55.27	2631.28	3452.25	4273.22	5094.20	5915.17
8	48.70	2355.53	3090.47	3825.41	4560.35	5295.29
9	43.52	2131.79	2796.92	3462.05	4127.18	4792.30
10	39.34	1946.68	2554.05	3161.43	3768.80	4376.18

In fact, for the policyholder studied here, 93.18% of the policyholders did not report any claim.

Conditional distributions form the theoretical basis of all regression analysis and therefore it is important to examine them. The conditional distribution of $X|Y=y$ is given by

(6.1) $$f_{X|Y}(x|y)=\frac{x^{\alpha-1}(1-x)^{\beta-1}\exp(-\sigma xy)}{B(\alpha,\beta)\,{}_1F_1(\alpha,\alpha+\beta,-\sigma y)},\quad 0<x<1,$$

which is the confluent hypergeometric distribution with parameters α, β and σy according to Gordy (1998).

The conditional distribution of $Y|X = x$ is given by

$$f_{Y|X}(y|x) = \frac{(\sigma x)^\gamma}{\Gamma(\gamma)} y^{\gamma-1} \exp(-\sigma x y), \quad y > 0, \tag{6.2}$$

which is a gamma distribution with shape parameter γ and scale parameter σx.

Furthermore, some algebra on (**??**) and (**??**) provides the conditional expectations (the regression of x on y and the regression of y on x), which are given by

$$E(X|Y = y) = \frac{\alpha}{\alpha+\beta} \frac{{}_1F_1(\alpha+1, \alpha+\beta+1, -\sigma y)}{{}_1F_1(\alpha, \alpha+\beta, -\sigma y)},$$

$$E(Y|X = x) = \frac{\gamma}{\alpha x}. \tag{6.3}$$

7. Conclusions

This paper presents an optimal BMS based on both random variables, i.e., the number of claims and the individual claim size. This model was constructed using a bivariate prior distribution for the two risk profiles on which the total claim amount depends. In consequence, we obtained premiums which can be written as credibility formula and are suitable for the computation of bonus-malus premiums. These premiums appear to be of most benefit to policyholders who report claims with a low individual claim size, while a larger premium is charged to those who produce a high individual claim size. It is concluded that it is fairer to charge policyholders premiums which not only take into account the number of claims, but also the total claim amount (which depends on both the number of claims and the individual claim size).

Acknowledgements. The authors are grateful to the referees for a very careful reading of the manuscript and suggestions which improved the paper. EGD was partially supported by ECO2009–14152 (MICINN, Spain).

References

[1] Bhattacharya, S. K., (1966). Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness. *Bull. Calcutta Statist. Assoc.*, 15, 20–31.

[2] Bühlmann, H. and Gisler, A. (2005). *A Course in Credibility Theory and its Applications.* Springer, New York.

[3] Denuit, M., Marechal, X., Pitrebois, S. and Walhin, J-F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems.* Wiley.

[4] Frangos, N. and Vrontos, S. (2001). Design of optimal Bonus-Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *Astin Bulletin*, 31, 1, 1–22.

[5] Gómez-Déniz, E., Pérez, J.M., Hernández, A., Vázquez, F. (2002). Measuring sensitivity in a bonus-malus system. *Insurance: Mathematics and Economics*, 31, 105–113.

[6] Gómez-Déniz, E.; Pérez, J.; and Vázquez, F. (2006). On the use of posterior regret Γ-minimax actions to obtain credibility premiums. *Insurance: Mathematics and Economics*, 42, 39–49.

[7] Gómez-Déniz, E.; Sarabia, J.M. and Calderín, E. (2008). Univariate and multivariate versions of the negative binomial inverse Gaussian distributions with applications. *Insurance: Mathematics and Economics*, 39, 115–121.

[8] Heilmann, W. (1989) Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics* 8, 1, 77–95.
[9] Gerber, H.U. (1979). *An Introduction to Mathematical Risk Theory.* Huebner Foundation Monograph 8.
[10] Gordy, M.B. (1998). Computationally convenient distributional assumptions for common-value auctions. *Computational Economics*, 12, 61–78.
[11] Jong, P. and Heller, G. (2008). *Generalized Linear Models for Insurance Data*. Cambridge University Press.
[12] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2008). *Loss Models: From Data to Decisions. Third Edition.* Wiley.
[13] Lemaire, J. (1979). How to define a bonus-malus system with an exponential utility function. *Astin Bulletin*, **10**, 274–282.
[14] Lemaire, J. (1985). *Automobile Insurance. Actuarial Models.* Kluwer-Nijhoff Publishing, Dordrecht.
[15] Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, London.
[16] Lemaire, L. (2004). Bonus-Malus systems. In *Encyclopedia of Actuarial Science*. Wiley.
[17] Mert, M. and Saykan, Y. (2005). Bonus-Malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto. *Hacettepe Journal of Mathematics and Statistics*, 34, 75–81.
[18] Rolski, T., Schmidli, H. Schmidt, V. and Teugel, J. (1999). *Stochastic Processes for Insurance and Finance.* John Wiley & Sons.
[19] Sarabia, J.M., Gómez-Déniz, E. and Vázquez, F.J. (2004). On the use of conditional specification models in claim count distributions: An application to bonus-malus systems. *Astin Bulletin*, 34, 85–98.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1063 – 1077

The Lindley-Poisson distribution in lifetime analysis and its properties

Wenhao Gui* , Shangli Zhang† and Xinman Lu‡

Abstract

In this paper, we introduce a new compounding distribution, named the Lindley-Poisson distribution. We investigate its characterization and statistical properties. The maximum likelihood inference using EM algorithm is developed. Asymptotic properties of the MLEs are discussed and simulation studies are performed to assess the performance of parameter estimation. We illustrate the proposed model with two real applications and it shows that the new distribution is appropriate for lifetime analyses.

2000 AMS Classification:

Keywords: Lindley distribution, Poisson distribution, Hazard function, Maximum likelihood estimation, EM algorithm, Fisher information matrix.

Received 08 : 08 : 2013 : Accepted 11 : 01 : 2014 Doi : 10.15672/HJMS.201427453

1. Introduction

The Lindley distribution was originally introduced by [16] to illustrate a difference between fiducial distribution and posterior distribution. It has attracted a wide applicability in survival and reliability. Its density function is given by

$$f(t) = \frac{\theta^2}{1+\theta}(1+t)e^{-\theta t}, \quad t, \theta > 0. \tag{1.1}$$

We denoted this by writing $LD(\theta)$. The density in (1.1) indicates that the Lindley distribution is a mixture of an exponential distribution with scale θ and a gamma distribution with shape 2 and scale θ, where the mixing proportion is $\theta/(1+\theta)$.

*Department of Mathematics and Statistics, University of Minnesota Duluth, Duluth, MN 55812, USA.
Email: wgui@d.umn.edu
†Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China.
‡Business School,Nankai University, Tianjin 300071, China.

[11] provided a comprehensive treatment of the statistical properties of the Lindley distribution and showed that in many ways it performs better than the well-known exponential distribution. [20] discussed the discrete Poisson–Lindley distribution by compounding the Poission distribution and the Lindley distribution. [10] investigated the properties of the zero-truncated Poisson—Lindley distribution. [3] extended the Lindley distribution by exponentiation. [22] introduced and analyzed a three-parameter generalization of the Lindley distribution, which was used by [17] to derive an extended version of the compound Poisson distribution. [21] introduced a two-parameter Lindley distribution of which the one-parameter $LD(\theta)$ is a particular case, for modeling waiting and survival times data. [9] introduced a two-parameter power Lindley distribution (PL) and discussed its properties. [18] proposed a generalized Lindley distribution (GL) and provided comprehensive account of the mathematical properties of the distribution.

On the other hand, the studies and analysis of lifetime data play a central role in a wide variety of scientific and technological fields. There have been developed several distributions by compounding some useful life distributions. [1] introduced a two-parameter exponential-geometric (EG) distribution with decreasing failure rate by compounding an exponential with a geometric distribution. [15] proposed an exponential-Poisson (EP) distribution by mixing an exponential and zero truncated Poisson distribution and discussed its various properties. [5] introduced a new two-parameter distribution family with decreasing failure rate by mixing power-series distribution and exponential distribution.

The aim of this paper is to propose an extension of the Lindley distribution which offers a more flexible distribution for modeling lifetime data. In this paper, we introduce an extension of the Lindley distribution by mixing Lindley and zero truncated Poisson distribution. It differs from the discrete Poisson–Lindley distribution proposed by [20]. Since the Lindley distribution is not a generalization of exponential distribution, the model EP in [15] can not be obtained as a particular case of the new model in this paper. An interpretation of the proposed model is as follows: a situation where failure occurs due to the presence of an unknown number, Z, of initial defects of same kind. Z is a zero truncated Poisson variable. Their lifetimes, Y's, follow a Lindley distribution. Then for modeling the first failure X, the distribution leads to the Lindley–Poisson distribution. We aim to discuss some properties of the proposed distribution.

The rest of this paper is organized as follows: in Section 2, we present the new Lindley-Poisson distribution and investigate its basic properties, including the shape properties of its density function and the hazard rate function, stochastic orderings and representation, moments and measurements based on the moments. Section 3 discusses the distributions of some extreme order statistics. The maximum likelihood inference using EM algorithm and asymtotical properties of the estimates are discussed in Section 4. Simulation studies are also conducted in this Section. Section 5 gives a real illustrative application and reports the results. Our work is concluded in Section 6.

2. Lindley-Poisson Distribution and its Properties

2.1. Density and hazard function. The new distribution can be constructed as follows. Suppose that the failure of a device occurs due to the presence of Z (unknown number) initial defects of some kind. Let $Y_1, Y_2, ..., Y_Z$ denote the failure times of the initial defects, then the failure time of this device is given by $X = \min(Y_1, ..., Y_Z)$.

Suppose the failure times of the initial defects $Y_1, Y_2, ..., Y_Z$ follow a Lindley distribution $LD(\theta)$ and Z has a zero truncated Poisson distribution with probability mass function as follows:

$$p(Z = z) = \frac{\lambda^z e^{-\lambda}}{z!(1 - e^{-\lambda})}, \quad \lambda > 0, z = 1, 2, ... \tag{2.1}$$

By assuming that the random variables Y_i and Z are independent, then the density of $X|Z = z$ is given by

$$f(x|z) = \frac{\theta^2 (x+1) z e^{-xz\theta} (\theta + \theta x + 1)^{z-1}}{(\theta+1)^z}, \quad x > 0,$$

and the marginal probability density function of X is

$$f(x) = \frac{\theta^2 \lambda (x+1) e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1} - \theta x}}{(\theta+1)\left(e^{\lambda}-1\right)}, \quad \theta > 0, \lambda > 0, x > 0. \tag{2.2}$$

In the sequel, the distribution of X will be refered to as the LP, which is customary for such a name given to the distribution arising via the operation of compounding in the literature.

2.1. Theorem. *Considering the LP distribution with the probability density function in (2.2), we have the following properties:*

(1) *As λ goes to zero, $LP(\theta, \lambda)$ leads to the Lindley distribution $LD(\theta)$.*
(2) *If $\theta^2(\lambda+1) \geq 1$, $f(x)$ is decreasing in x. If $\theta^2(\lambda+1) < 1$, $f(x)$ is a unimodal function at x_0, where x_0 is the solution of the equation $\theta^2\lambda(x+1)^2 + (\theta+1)e^{\theta x}(\theta + \theta x - 1) = 0$.*

Proof. 1. As λ goes to zero, then

$$\begin{aligned} \lim_{\lambda \to 0} f(x) &= \lim_{\lambda \to 0} \frac{\theta^2 \lambda (x+1) e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1} - \theta x}}{(\theta+1)\left(e^{\lambda}-1\right)} \\ &= \frac{\theta^2 (x+1) e^{-\theta x}}{\theta+1}, \end{aligned}$$

which is the probability density distribution of $LD(\theta)$.

2. $f(0) = \frac{\theta^2 e^{\lambda} \lambda}{(\theta+1)\left(e^{\lambda}-1\right)}$ and $f(\infty) = 0$. The first derivative of $\log f(x)$ is

$$\frac{d \log f(x)}{dx} = -\frac{e^{-\theta x}\left[\theta^2\lambda(x+1)^2 + (\theta+1)e^{\theta x}(\theta+\theta x - 1)\right]}{(\theta+1)(x+1)}.$$

Let $s(x) = \theta^2\lambda(x+1)^2 + (\theta+1)e^{\theta x}(\theta+\theta x-1)$, then $s(0) = \theta^2(\lambda+1) - 1$ and $s(\infty) = \infty$, $s'(x) = \theta^2(x+1)\left[2\lambda + (\theta+1)e^{\theta x}\right] > 0$.

If $\theta^2(\lambda+1) \geq 1$, then $s(x) \geq 0$, $\frac{d \log f(x)}{dx} \leq 0$, i.e., $f(x)$ is decreasing in x. If $\theta^2(\lambda+1) < 1$, $f(x)$ is a unimodal function at x_0, where x_0 is the solution of the equation $s(x) = 0$.

□

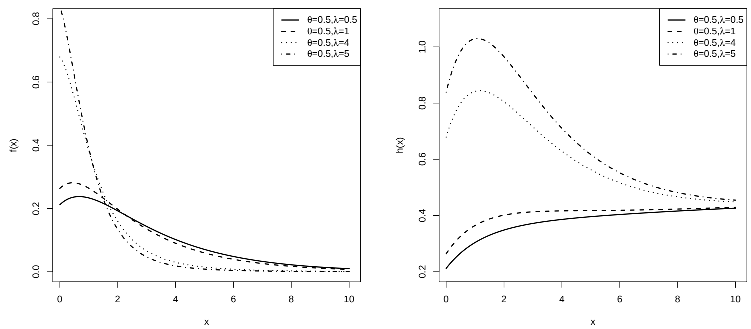
The cumulative distribution of the LP distribution is given by

$$F(x) = \frac{e^{\lambda} - e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1}}}{e^{\lambda}-1}, \quad x > 0. \tag{2.3}$$

The hazard rate function of the $LP(\theta, \lambda)$ distribution is given by

$$h(x) = \frac{\theta^2 \lambda (x+1) e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1} - \theta x}}{(\theta+1)\left[e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1}} - 1\right]}, \quad x > 0. \tag{2.4}$$

2.2. Theorem. *Considering the hazard function of the LP distribution, we have the following properties:*



(a) density function **(b)** hazard function

Figure 1. Plots of the LP density and hazard function for some parameter values.

(1) *If* $-\theta^3\lambda+\theta^2\lambda+\theta+1>0$ *and the equation* $(\theta+1)e^{\theta x}-\theta^2\lambda(x+1)^2(\theta+\theta x-1)=0$ *has no real roots, then the hazard function is increasing.*
(2) *If* $-\theta^3\lambda+\theta^2\lambda+\theta+1<0$ *and the equation* $(\theta+1)e^{\theta x}-\theta^2\lambda(x+1)^2(\theta+\theta x-1)=0$ *has one real roots, then the hazard function is bathtub shaped.*

Proof. $h(0)=\frac{\theta^2 e^{\lambda}\lambda}{(\theta+1)(e^{\lambda}-1)}$. For the LP distribution, we have

$$\eta(x)=-\frac{f'(x)}{f(x)}=\frac{e^{-\theta x}\left[\theta^2\lambda(x+1)^2+(\theta+1)e^{\theta x}(\theta+\theta x-1)\right]}{(\theta+1)(x+1)},$$

and its first derivative is

$$\eta'(x)=\frac{e^{-\theta x}\left[(\theta+1)e^{\theta x}-\theta^2\lambda(x+1)^2(\theta+\theta x-1)\right]}{(\theta+1)(x+1)^2}.$$

Let $t(x)=(\theta+1)e^{\theta x}-\theta^2\lambda(x+1)^2(\theta+\theta x-1)$, then $t(0)=-\theta^3\lambda+\theta^2\lambda+\theta+1$ and $t(\infty)=\infty$, the sign of $\eta'(x)$ is the sign of $t(x)$ and $\eta'(x)=0$ if $t(x)=0$. The properties follow from the results in [12]. □

For the Lindley distribution $LD(\theta)$, its hazard function $h(x)=\frac{\theta^2(1+x)}{\theta+1+\theta x}$ which is increasing. For the exponential distribution, its hazard function $h(x)=\theta$ which is a constant. (2.4) shows the flexibility of the LP distribution over the Lindley and exponential distribution.

Figure 1a shows some density functions of the $LP(\theta,\lambda)$ distribution with various parameters. Figure 1b shows some shapes of the $LP(\theta,\lambda)$ hazard function with various parameters.

2.2. Stochastic Ordering. In probability theory and statistics, a stochastic order quantifies the concept of one random variable being “bigger" than another. A random variable X is less than Y in the ususal stochastic order (denoted by $X\prec_{st}Y$) if $F_X(x)\geq F_Y(x)$ for all real x. X is less than Y in the hazard rate order (denoted by $X\prec_{hr}Y$) if $h_X(x)\geq h_Y(x)$, for all $x\geq 0$. X is less than Y in the likelihood ratio order (denoted by $X\prec_{lr}Y$) if $f_X(x)/f_Y(x)$ increases in x over the union of the supports of X and Y. It is known that $X\prec_{lr}Y\Rightarrow X\prec_{hr}\Rightarrow X\prec_{st}Y$, see [19].

2.3. Theorem. *If $X \sim LP(\theta, \lambda_1)$ and $Y \sim LP(\theta, \lambda_2)$, and $\lambda_1 < \lambda_2$, then $Y \prec_{lr} X$, $Y \prec_{hr} X$ and $Y \prec_{st} X$.*

Proof. The density ratio is given by

$$U(x) = \frac{f_X(x)}{f_Y(x)} = \frac{\left(e^{\lambda_2} - 1\right)\lambda_1 \exp\left(\frac{\lambda_1 e^{-\theta x}(\theta+\theta x+1)}{\theta+1} - \frac{\lambda_2 e^{-\theta x}(\theta+\theta x+1)}{\theta+1}\right)}{\left(e^{\lambda_1} - 1\right)\lambda_2}.$$

Taking the derivative with respect to x,

$$U'(x) = -\frac{\theta^2\left(e^{\lambda_2} - 1\right)\lambda_1\left(\lambda_1 - \lambda_2\right)(x+1)\exp\left(-\frac{e^{-\theta x}\left(-\lambda_1(\theta+\theta x+1)+\lambda_2(\theta+\theta x+1)+\theta(\theta+1)xe^{\theta x}\right)}{\theta+1}\right)}{(\theta+1)\left(e^{\lambda_1} - 1\right)\lambda_2}.$$

If $\lambda_1 < \lambda_2$, $U'(x) > 0$, $U(x)$ is an increasing function of x. The results follow. $\square$

2.3. Moments and Measures based on moments. In this section, we consider the moments and measures of the LP distribution $X \sim LP(\theta, \lambda)$. The k-th raw moment of X is given by, for $k = 1, 2, ...$,

$$\mu_k = \mathbb{E}(X^k) = k\int_0^\infty x^{k-1}\bar{G}(x)dx = \int_0^\infty \frac{kx^{k-1}[e^{\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1}} - 1]}{e^\lambda - 1}\,dx.$$

$\mathbb{E}(X^k)$ cannot be expressed in a simple closed-form and need be calculated numerically. Using numerical integration, we can find some measures based on the moments such as mean, variance, skewness and kurtosis etc. For the skewness and kurtosis coefficients, $\sqrt{\beta_1} = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{3/2}}$ and $\beta_2 = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2}$.

The cumulative distribution of the LP distribution is given in (2.3). The qth ($0 \le q \le 1$) quantile $x_q = F^{-1}(q)$ of the $LP(\theta, \lambda)$ distribution is

$$x_q = \frac{-\theta - W\left(-\frac{e^{-\theta-1}(\theta+1)\log\left(e^\lambda - e^\lambda q + q\right)}{\lambda}\right) - 1}{\theta},$$

where $W(a)$ giving the principal solution for w in $a = we^w$ is pronounced as Lambert W function, see [14].

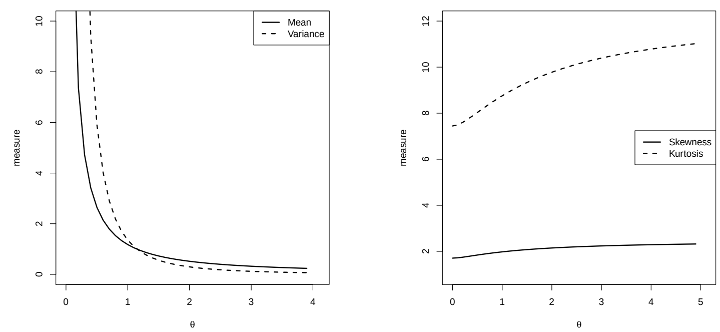
In particular, the median of the $LP(\theta, \lambda)$ distribution is given by

$$x_m = \frac{-\theta - W\left(-\frac{e^{-\theta-1}(\theta+1)\log\left(\frac{1}{2}\left(e^\lambda + 1\right)\right)}{\lambda}\right) - 1}{\theta}. \tag{2.5}$$

Figure 2a displays the mean and variance of the $LP(\theta, \lambda = 1)$ distribution. Figure 2b shows the skewness and kurtosis coefficients of the $LP(\theta, \lambda = 1)$ distribution. From the figures, it is found that the $LP(\theta, \lambda = 1)$ distribution has positive skewness and kurtosis coefficients. The coefficients are increasing functions of θ.

3. Distributions of Order Statistics

Let $X_1, X_2, ..., X_n$ be a random sample of size n from the $LP(\theta, \lambda)$ distribution. By the usual central limit theorem, the same mean $(X_1 + ... + X_n)/n$ approaches the normal distribution as $n \to \infty$. Sometimes one would be interested in the asymptotics of the sample minima $X_{1:n} = \min(X_1, ..., X_n)$ and the sample maxima $X_{n:n} = \max(X_1, ..., X_n)$. These extreme order statistics represent the life of series and parallel system and have important applications in probability and statistics.



(a) **(b)**

Figure 2. (a) Plot of mean and variance of the $LP(\theta, \lambda = 1)$ distribution; (b) Plot of skewness and kurtosis coefficients of the $LP(\theta, \lambda = 1)$ distribution.

3.1. Theorem. *Let $X_{1:n}$ and $X_{n:n}$ be the smallest and largest order statistics from the $LP(\theta, \lambda)$ distribution. Then*
(1) $\lim_{n\to\infty} P(X_{1:n} \le b_n^* t) = 1 - e^{-t}, t > 0,$ *where* $b_n^* = F^{-1}(1/n)$.
(2) $\lim_{n\to\infty} P(X_{n:n} \le b_n t) = e^{-t^{-1}}, t > 0,$ *where* $b_n = F^{-1}(1 - 1/n)$.

Proof. We apply the following asymptotical results for $X_{1:n}$ and $X_{n:n}$ ([2]).
(1) For the smallest order statistic $X_{1:n}$, we have

$$\lim_{n\to\infty} P(X_{1:n} \le a_n^* + b_n^* t) = 1 - e^{-t^c}, \quad t > 0, c > 0,$$

(of the Weibull type) where $a_n^* = F^{-1}(0)$ and $b_n^* = F^{-1}(1/n) - F^{-1}(0)$ if and only if $F^{-1}(0)$ is finite and for all $t > 0$ and $c > 0$,

$$\lim_{\epsilon\to 0^+} \frac{F(F^{-1}(0) + \epsilon t)}{F(F^{-1}(0) + \epsilon)} = t^c.$$

For the $LP(\theta, \lambda)$ distribution, its cumulative distribution function is

$$F(x) = \frac{e^{\lambda} - e^{\frac{\lambda e^{-\theta x}(\theta + \theta x + 1)}{\theta + 1}}}{e^{\lambda} - 1}, \quad \theta > 0, \lambda > 0, x > 0.$$

Let $F(x) = 0$, we have $\theta + \theta x + 1 = e^{\theta x}(\theta + 1) \ge (1 + \theta x)(\theta + 1), \theta x^2 \le 0$. Thus $F^{-1}(0) = 0$ is finite. Furthermore,

$$\lim_{\epsilon\to 0^+} \frac{F(0 + \epsilon t)}{F(0 + \epsilon)} = t \lim_{\epsilon\to 0^+} \frac{f(\epsilon t)}{f(\epsilon)} = t.$$

Therefore, we obtain that $c = 1$, $a_n^* = 0$ and $b_n^* = F^{-1}(1/n)$ which is the $\frac{1}{n}$th quantile.
(2) For the largest order statistic $X_{n:n}$, we have

$$\lim_{n\to\infty} P(X_{n:n} \le a_n + b_n t) = e^{-t^{-d}}, \quad t > 0, d > 0$$

(of the Fréchet type) where $a_n = 0$ and $b_n^* = F^{-1}(1 - 1/n)$ if and only if $F^{-1}(1) = \infty$ and there exists a constant $d > 0$ such that

$$\lim_{x\to\infty} \frac{1 - F(xt)}{1 - F(x)} = t^{-d}.$$

For the $LP(\theta, \lambda)$ distribution, let $F(x) = 1$, then $\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1} = 0$, we have the solution $x = \infty$. Thus $F^{-1}(1) = \infty$. Furthermore,

$$\lim_{x\to\infty} \frac{1 - F(xt)}{1 - F(x)} = t^{-1}.$$

Therefore, we obtain that $d = 1$, $a_n = 0$ and $b_n = F^{-1}(1-1/n)$ which is the the $(1-\frac{1}{n})$th quantile.

□

3.2. Remark. Let $Q^*(t)$ and $Q(t)$ denote the limiting distributions of the random variables $(X_{1:n} - a_n^*)/b_n^*$ and $(X_{n:n} - a_n)/b_n$ respectively, then for $k > 1$, the limiting distributions of $(X_{k:n} - a_n^*)/b_n^*$ and $(X_{n-k+1:n} - a_n)/b_n$ are given by, see [2],

$$\lim_{n\to\infty} P(X_{k:n} \le a_n^* + b_n^* t) = 1 - \sum_{j=0}^{k-1} (1 - Q^*(t)) \frac{[-\log(1 - Q^*(t))]^j}{j!},$$

$$\lim_{n\to\infty} P(X_{n-k+1:n} \le a_n + b_n t) = \sum_{j=0}^{k-1} Q(t) \frac{[-\log Q(t)]^j}{j!}.$$

4. Estimation and inference

4.1. Maximum likelihood estimation. Here, we consider the maximum likelihood estimation about the parameters (θ, λ) of the LP model. Suppose $y_{obs} = \{x_1, x_2, ..., x_n\}$ is a random sample of size n from the $LP(\theta, \lambda)$ distribution. Then the log-likelihood function is given by

$$\begin{aligned} l &= \log \prod_{i=1}^{n} f_X(x_i) \\ &= \lambda \sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta\lambda \sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta + 1} - \theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \log(x_i + 1) \\ &\quad +2n\log(\theta) - n\log(\theta+1) - n\log\left(e^{\lambda} - 1\right) + n\log(\lambda). \end{aligned} \tag{4.1}$$

The associated gradients are found to be

$$\frac{\partial l}{\partial \theta} = -\sum_{i=1}^{n} x_i + \frac{2n}{\theta} - \frac{n}{\theta+1} - \frac{\theta(\theta+2)\lambda \sum_{i=1}^{n} x_i e^{-\theta x_i}}{(\theta+1)^2} - \frac{\theta\lambda \sum_{i=1}^{n} x_i^2 e^{-\theta x_i}}{\theta+1}, \tag{4.2}$$

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta \sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta+1} - \frac{n e^{\lambda}}{e^{\lambda} - 1} + \frac{n}{\lambda}. \tag{4.3}$$

The estimates of the parameters maximize the likelihood function. Equalizing the obtained gradients expressions to zero yield the likelihood equations. However, they do not lead to explicit analytical solutions for the parameters. Thus, the estimates can be obtained by means of numerical procedures such as Newton-Raphson method. The program R provides the nonlinear optimization routine *optim* for solving such problems.

The equation $\frac{\partial l}{\partial \theta} = 0$ could be solved exactly for λ, namely

$$(4.4)\hat{\lambda} = \frac{(\hat{\theta}+1)\left[\hat{\theta}(\hat{\theta}+1)\sum_{i=1}^{n} x_i - (\hat{\theta}+2)n\right]}{\hat{\theta}\left[-(\hat{\theta}+1)^2\sum_{i=1}^{n} x_i e^{-\hat{\theta}x_i} - \hat{\theta}(\hat{\theta}+1)\sum_{i=1}^{n} x_i^2 e^{-\hat{\theta}x_i} + \sum_{i=1}^{n} x_i e^{-\hat{\theta}x_i}\right]},$$

conditional on the value of $\hat{\theta}$, where $\hat{\theta}$ and $\hat{\lambda}$ are the maximum likelihood estimators for the parameters θ and λ, respectively.

In the following, Theorem 4.1 gives the condition for the existence and uniqueness of $\hat{\lambda}$ when θ is known.

4.1. Theorem. *For the MLEs, let $l_2(\lambda;\theta,y_{obs})$ denote the function on the RHS of the expression in (4.3), if θ is known, then the root of $l_2(\lambda;\theta,y_{obs}) = 0$, $\hat{\lambda}$, uniquely exists if* $\sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta\sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta+1} > \frac{n}{2}$.

Proof. Notice that $\lim_{\lambda\to 0} l_2(\lambda;\theta,y_{obs}) = \sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta\sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta+1} - \frac{n}{2} > 0$ when $\sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta\sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta+1} > \frac{n}{2}$. On the other hand, we can show that $\lim_{\lambda\to\infty} l_2(\lambda;\theta,y_{obs}) = \sum_{i=1}^{n} e^{-\theta x_i} + \frac{\theta\sum_{i=1}^{n} x_i e^{-\theta x_i}}{\theta+1} - n$. Consider $g(x) = e^{-\theta x} + \frac{\theta}{\theta+1} x e^{-\theta x} - 1$, $g(0) = 0$ and $g(\infty) = -1$, $g'(x) = -\frac{\theta^2(x+1)e^{\theta(-x)}}{\theta+1} < 0$, therefore, $\lim_{\lambda\to\infty} l_2(\lambda;\theta,y_{obs}) < 0$, there is at least one root of $l_2(\lambda;\theta,y_{obs}) = 0$. We need to prove that the function $l_2(\lambda;\theta,y_{obs})$ is decreasing in λ. Taking the first derivative

$$l_2'(\lambda;\theta,y_{obs}) = -\frac{\left[-e^{\lambda}\left(\lambda^2+2\right) + e^{2\lambda} + 1\right]n}{\left(e^{\lambda}-1\right)^2\lambda^2} = -\frac{e^{\lambda}\left[-\left(\lambda^2+2\right) + e^{\lambda} + e^{-\lambda}\right]n}{\left(e^{\lambda}-1\right)^2\lambda^2} < 0.$$

This completes the proof.

□

4.2. An EM algorithm. An expectation–maximization (EM) algorithm ([7]) is a powerful method for finding maximum likelihood estimates of parameters in statistical models, where the model depends on unobserved latent variables. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter estimates are then used to determine the distribution of the latent variables in the next E step. We propose the use of the EM algorithm in this section.

Assume that (X,Z) denotes a random vector, where X denotes the observed data and Z denotes the missing data. To implement the algorithm we define the hypothetical complete–data distribution with density function

$$f(x,z) = p(z)f(x|z) = \frac{\theta^2(x+1)ze^{-xz\theta}(\theta+\theta x+1)^{z-1}}{(\theta+1)^z}\frac{\lambda^z e^{-\lambda}}{z!(1-e^{-\lambda})}, x>0, z=1,2,...,$$

where $\theta > 0$ and $\lambda > 0$ are parameters. It is straightforward to verify that the computation of the conditional expectation of $(Z|X)$ using the pdf

$$p(z|x) = \frac{(\theta+1)^{1-z}\lambda^{z-1}(\theta+\theta x+1)^{z-1}\exp\left(-\frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1} + \theta x - \theta x z\right)}{(z-1)!}, z=1,2,...$$

Then we have

$$\mathbb{E}(Z|X) = 1 + \frac{\lambda e^{-\theta x}(\theta+\theta x+1)}{\theta+1}.$$

The cycle is completed with the M–step which is essentially-full data maximum likelihood over the parameters, with the missing $Z's$ replaced by their conditional expectations $\mathbb{E}(Z|X)$. Thus, an EM iteration is given by

$$\theta^{(t+1)} = 2n[\sum_{i=1}^{n} \frac{x_i+1}{\theta^{(t)}+\theta^{(t)}x_i+1} - \sum_{i=1}^{n} \frac{(x_i+1)\,w_i^{(t)}}{\theta^{(t)}+\theta^{(t)}x_i+1} + \sum_{i=1}^{n} x_i w_i^{(t)} + \frac{\sum_{i=1}^{n} w_i^{(t)}}{\theta^{(t)}+1}]^{-1},$$

$$\lambda^{(t+1)} = n^{-1}[1-e^{-\lambda^{(t)}}]\sum_{i=1}^{n} w_i^{(t)},$$

where $w_i^{(t)} = 1 + \frac{\lambda^{(t)} e^{-\theta^{(t)} x_i}(\theta^{(t)}+\theta^{(t)}x_i+1)}{\theta^{(t)}+1}$.

4.3. Asymtotic variance and covariance of MLEs. It is known that under some regular conditions, as the sample size increases, the distribution of the MLE tends to the bivariate normal distribution with mean (θ, λ) and covariance matrix equal to the inverse of the Fisher information matrix, see [6]. The bivariate normal distribution can be used to construct approximate confidence intervals for the parameters θ and λ.

Let $I = I(\theta, \lambda; y_{obs})$ be the observed matrix with elements I_{ij} with $i, j = 1, 2$. The elements of the observed information matrix are found as follows:

$$I_{11} = -\frac{\left((\theta+1)^2-2\lambda\right)\sum_{i=1}^{n} x_i^2 e^{-\theta x_i}}{(\theta+1)^2} - \frac{\theta\lambda\sum_{i=1}^{n} x_i^3 e^{-\theta x_i}}{\theta+1} + \frac{2\lambda\sum_{i=1}^{n} x_i e^{-\theta x_i}}{(\theta+1)^3} + \frac{2n}{\theta^2} - \frac{n}{(\theta+1)^2},$$

$$l_{12} = l_{21} = \frac{\theta(\theta+2)\sum_{i=1}^{n} x_i e^{-\theta x_i}}{(\theta+1)^2} + \frac{\theta\sum_{i=1}^{n} x_i^2 e^{-\theta x_i}}{\theta+1},$$

$$l_{22} = -\frac{e^{\lambda} n}{(e^{\lambda}-1)^2} + \frac{n}{\lambda^2}.$$

The expectation $J = \mathbb{E}(I(\theta, \lambda; y_{obs}))$ is taken with respect to the distribution of X. The Fisher information matrix is given by

$$J(\theta, \lambda) = n\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}$$

where

$$J_{11} = -\frac{\left((\theta+1)^2-2\lambda\right)\mathbb{E}(X^2 e^{-\theta X})}{(\theta+1)^2} - \frac{\theta\lambda\mathbb{E}(X^3 e^{-\theta X})}{\theta+1} + \frac{2\lambda\mathbb{E}(X e^{-\theta X})}{(\theta+1)^3} + \frac{2}{\theta^2} - \frac{1}{(\theta+1)^2},$$

$$J_{12} = J_{21} = \frac{\theta(\theta+2)\mathbb{E}(X e^{-\theta X})}{(\theta+1)^2} + \frac{\theta\mathbb{E}(X^2 e^{-\theta X})}{\theta+1},$$

$$J_{22} = \frac{1}{\lambda^2} - \frac{e^{\lambda}}{(e^{\lambda}-1)^2}.$$

The inverse of $J(\theta, \lambda)$, evaluated at $\hat{\theta}$ and $\hat{\lambda}$ provides the asymptotic variance–covariance matrix of the MLEs. Alternative estimates can be obtained from the inverse of the observed information matrix since it is a consistent estimator of J^{-1}.

4.4. Simulation study. The random data X from the proposed distribution can be generated as follows:

(1) Generate $Z \sim$ zero truncated Poisson (λ).
(2) Generate $U_i \sim$ Uniform$(0, 1)$, $i = 1, ..., Z$.
(3) Generate $V_i \sim$ Exponential(θ), $i = 1, ..., Z$.
(4) Generate $W_i \sim$ Gamma$(2, \theta)$, $i = 1, ..., Z$.
(5) If $U_i \leq \theta/(1+\theta)$, then set $Y_i = V_i$, otherwise, set $Y_i = W_i$, $i = 1, ..., Z$.
(6) Set $X = \min(Y_1, ..., Y_Z)$.

In order to assess the performance of the approximation of the variances and covariances of the MLEs determined from the information matrix, a simulation study (based on 10000 simulations) has been conducted.

For each value of (θ, λ), the parameter estimates have been obtained by the EM iteration in Section 4.2 with different initial values. The convergence is assumed when the absolute differences between successive estimates are less than 10^{-5}.

The simulated values of $Var(\hat{\theta})$, $Var(\hat{\lambda})$ and $Cov(\hat{\theta}, \hat{\lambda})$ as well as the approximate values determined by averaging the corresponding values obtained from the expected and observed information matrices are given in Table 1. We can see that for large values of n, the approximate values determined from expected and observed information matrices are quite close to the corresponding simulated values. The approximation becomes quite accurate as n increases. As expected, variances and covariances of the MLEs obtained from the observed information matrix are quite close to that of the expected information matrix for large values of n.

Table 1. Variances and covariances of the MLEs.

n	(θ, λ)	Simulated			From expected information			From observed information		
		$Var(\hat{\theta})$	$Var(\hat{\lambda})$	$Cov(\hat{\theta}, \hat{\lambda})$	$Var(\hat{\theta})$	$Var(\hat{\lambda})$	$Cov(\hat{\theta}, \hat{\lambda})$	$Var(\hat{\theta})$	$Var(\hat{\lambda})$	$Cov(\hat{\theta}, \hat{\lambda})$
50	(0.5, 1.0)	0.1263	5.1246	-0.5943	0.0669	4.6146	-0.5249	0.0675	5.5077	-0.5827
50	(1.0, 0.5)	0.1809	2.3515	-0.7432	0.2009	2.6355	-0.8070	0.1112	1.9026	-0.7148
50	(0.5, 2.0)	0.0854	3.0085	-0.4022	0.0755	3.4085	-0.4615	0.0503	2.6381	-0.3235
50	(2.0, 0.5)	0.7783	3.3421	-1.5915	0.7578	3.0401	-1.3959	0.8382	3.6288	-1.6234
50	(2.0, 2.0)	0.7069	2.5474	-1.1334	0.7001	2.0854	-1.0336	0.7149	3.4743	-1.2402
100	(0.5, 1.0)	0.0365	2.9019	-0.3419	0.0476	2.9411	-0.3599	0.0334	2.3195	-0.3281
100	(1.0, 0.5)	0.0901	1.7915	-0.3829	0.0996	1.9011	-0.4122	0.0925	1.643	-0.3645
100	(0.5, 2.0)	0.0234	1.4168	-0.1738	0.0289	1.4896	-0.1882	0.0252	1.2935	-0.162
100	(2.0, 0.5)	0.2743	1.2773	-0.5513	0.2824	1.2676	-0.5510	0.2605	1.2929	-0.511
100	(2.0, 2.0)	0.3602	1.0218	-0.5014	0.3588	1.0148	-0.5218	0.349	0.9358	-0.4904
500	(0.5, 1.0)	0.0064	0.4256	-0.0506	0.0063	0.4238	-0.0496	0.0065	0.4462	-0.052
500	(1.0, 0.5)	0.0545	0.943	-0.2255	0.0522	0.9426	-0.2201	0.0567	0.9446	-0.2278
500	(0.5, 2.0)	0.0028	0.2001	-0.0211	0.0027	0.2009	-0.0209	0.0029	0.1998	-0.0213
500	(2.0, 0.5)	0.0899	0.3562	-0.1753	0.0888	0.3596	-0.1761	0.0938	0.3548	-0.1749
500	(2.0, 2.0)	0.0419	0.1672	-0.0723	0.0418	0.1672	-0.0733	0.0416	0.1673	-0.0723

In addition, simulations have been conduced to investigate the convergence of the proposed EM algorithm in Section 4.2. Ten thousand samples of size 100 and 500 of which are randomly sampled from the LP distribution for each of the five values of (θ, λ) are generated.

The results are presented in Table 2, which gives the averages of the 10000 MLEs, $av(\hat{\theta}), av(\hat{\lambda})$, and average number of iterations to convergence, $av(h)$, together with their standard errors, where

$$av(\hat{\theta}) = \frac{1}{10000}\sum_{i=1}^{10000}\hat{\theta}_i, \quad se(\hat{\theta}) = \sqrt{\frac{1}{10000}\sum_{i=1}^{10000}(\hat{\theta}_i - av(\hat{\theta}))^2},$$

$$av(\hat{\lambda}) = \frac{1}{10000}\sum_{i=1}^{10000}\hat{\lambda}_i, \quad se(\hat{\lambda}) = \sqrt{\frac{1}{10000}\sum_{i=1}^{10000}(\hat{\lambda}_i - av(\hat{\lambda}))^2},$$

$$av(\hat{h}) = \frac{1}{10000}\sum_{i=1}^{10000}\hat{h}_i, \quad se(\hat{h}) = \sqrt{\frac{1}{10000}\sum_{i=1}^{10000}(\hat{h}_i - av(\hat{h}))^2}.$$

From Table 2, it is observed that convergence has been achieved in all cases, even when the initial values are far from the true values and this endorses the numerical stability of the proposed EM algorithm. The EM estimates performed consistently. Standard errors of the MLEs decrease when sample size n increases.

Table 2. The means and standard errors of the EM estimator and iterations to convergence with initial values $(\theta^{(0)}, \lambda^{(0)})$ from 10000 samples.

n	θ	λ	$\theta^{(0)}$	$\lambda^{(0)}$	$av(\hat{\theta})$	$av(\hat{\lambda})$	$se(\hat{\theta})$	$se(\hat{\lambda})$	$av(h)$	$se(h)$
100	0.5	1	0.5	1	0.470	1.493	0.103	1.206	481.949	423.532
100	1	0.5	1	0.5	0.897	0.733	0.171	1.070	435.405	363.209
100	0.5	2	0.5	2	0.525	2.061	0.142	1.469	516.551	318.547
100	2	0.5	2	0.5	1.840	0.854	0.364	1.078	404.442	412.249
100	2	2	2	2	2.093	2.123	0.593	1.346	484.928	489.149
100	0.5	1	0.1	0.1	0.481	1.406	0.107	1.249	537.204	452.071
100	1	0.5	0.1	0.1	0.920	0.807	0.179	1.086	453.290	382.990
100	0.5	2	0.1	0.1	0.523	2.011	0.133	1.288	589.371	498.996
100	2	0.5	0.1	0.1	1.780	0.724	0.366	1.143	445.348	379.776
100	2	2	0.1	0.1	2.130	1.981	0.583	1.271	534.251	462.154
500	0.5	1	0.5	1	0.496	1.106	0.068	0.781	443.485	405.746
500	1	0.5	1	0.5	0.977	0.631	0.085	0.415	327.897	145.757
500	0.5	2	0.5	2	0.507	2.061	0.094	0.979	592.532	380.115
500	2	0.5	2	0.5	1.970	0.576	0.165	0.341	293.798	112.133
500	2	2	2	2	2.020	2.087	0.387	0.954	560.947	576.358
500	0.5	1	0.1	0.1	0.495	1.097	0.066	0.705	572.473	428.584
500	1	0.5	0.1	0.1	0.989	0.586	0.083	0.453	377.760	171.738
500	0.5	2	0.1	0.1	0.508	2.057	0.096	0.952	823.717	605.764
500	2	0.5	0.1	0.1	1.969	0.591	0.167	0.383	347.611	175.877
500	2	2	0.1	0.1	2.041	2.053	0.401	0.962	736.315	735.316

5. Illustrative Examples

In this section, we consider two numerical applications to test the performance of the new distribution. First, we consider the time intervals of the successive earthquakes taken from University of Bosphoros, Kandilli Observatory and Earthquake Research Institute-National Earthquake Monitoring Center. The data set has been previously studied by

[15]. The second dataset originally due to [4], which has also been analyzed previously by [13]. The data represent the survival times of guinea pigs injected with different doses of tubercle bacilli.

Table 3. Maximum likelihood parameter estimates(with (SE)) of the LP, LD, PL and GL models for the two datasets.

Example	Model	Estimations		loglik	AIC	K-S statistic	p-value
1 ($n = 24$)	LP	0.6515	2.7778	−32.0766	68.1532	0.1667	0.9024
		(0.2112)	(0.1578)				
	LD	1.0420	–	−34.5092	71.0184	0.2500	0.4490
		(0.1612)	–				
	PL	0.6215	1.0898	−32.6134	69.2268	0.2083	0.6860
		(0.1026)	(0.1745)				
	GL	0.5940	0.7701	−32.3633	68.7266	0.1667	0.9024
		(0.1567)	(0.1895)				
2 ($n = 72$)	LP	0.0112	2.9545	−392.4274	788.8548	0.1111	0.7658
		(0.0033)	(0.1496)				
	LD	0.0198	–	−394.5197	791.0394	0.1528	0.3701
		(0.0016)	–				
	PL	0.8451	0.0387	−396.8082	797.6164	0.1667	0.2700
		(0.0503)	(0.1745)				
	GL	1.1389	0.0212	−394.2822	792.5644	0.1528	0.3701
		(0.2101)	(0.0026)				

We fit the data sets with the Lindley–Poisson distribution $LP(\theta, \lambda)$, Lindley distribution $LD(\theta)$, Power Lindley distribution $PL(\alpha, \beta)$ and generalized Lindley distribution $GL(\alpha, \lambda)$ and examine the performances of the distributions.

Those probability density functions are given below:

$$PL: \qquad f(x|\Theta_1) = \frac{\alpha\beta^2}{\beta+1}(1+x^\alpha)x^{\alpha-1}e^{-\beta x^\alpha}, \quad \Theta_1 = (\alpha, \beta), \quad x > 0,$$

$$GL: \qquad f(x|\Theta_2) = \frac{\alpha\lambda^2}{1+\lambda}(1+x)[1 - \frac{1+\lambda+\lambda x}{1+\lambda}e^{-\lambda x}]^{\alpha-1}, \quad \Theta_2 = (\alpha, \lambda), \quad x > 0.$$

The maximum likelihood estimates of the parameters are obtained and the results are reported in Table 3. The Akaike information criterion (AIC) is computed to measure the goodness of fit of the models. $AIC = 2k - 2\log L$, where k is the number of parameters in the model and L is the maximized value of the likelihood function for the estimated model. Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. The Kolmogorov-Smirnov (K-S) statistics and the p-values for these models are also presented. The K-S test compares an empirical and a theoretical model by computing the maximum absolute difference between the empirical and theoretical distribution functions: $D = \max_x |F_n(x) - F(x)|$. The associated the p-value is the chance that the value of the Komogorov-Smirnov D statistic would be as large or larger than observed. The computation of p-value can be found in [8].

For the first dataset, the K-S statistics for the LP and GL models are same and smaller than those for the LD and PL models. For the LP model, AIC=68.1532 is smaller than that obtained for the GL model. Log-likelihood value=−32.0766 is larger than those for the GL model. It indicates that the LP model performs a best fit for this dataset. The good performance of the LP model can also be supported by the second dataset.

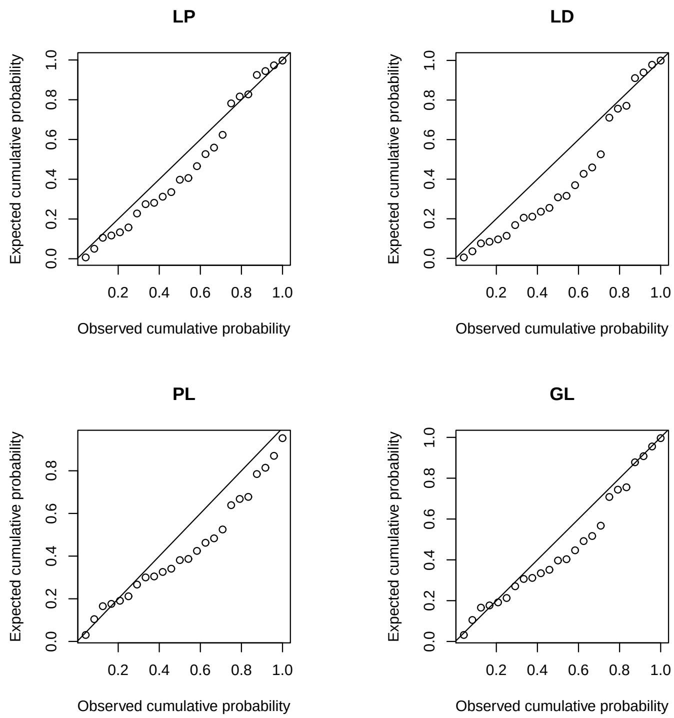


Figure 3. P-P plots for the first dataset.

Figure 3 and 4 display the probability-probability (P-P) plot for the two datasets.

6. Concluding Remarks

In this article, we have introduced a continuous Lindley-Poisson distribution by compounding the Lindley distribution and zero truncated Poisson distribution. The properties, including the shape properties of its density function and the hazard rate function, stochastic orderings, moments and measurements based on the moments are investigated. The distributions of some extreme order statistics are also derived. Maximum likelihood estimation method using EM algorithm is developed for estimating the parameters. Asymptotic properties of the MLEs are studied. We conduct intensive simulations and the results show that the estimation performance is satisfied as expected. We apply the model to two real datasets and the results demonstrate that the proposed model is appropriate for the datasets.

7. Acknowledgements

We thank the editor and two anonymous reviewers for their very valuable guidance and comments that helped us improve the manuscript. The first author's work was partially conducted while he was a postdoctorate researcher at Cornell University, supported

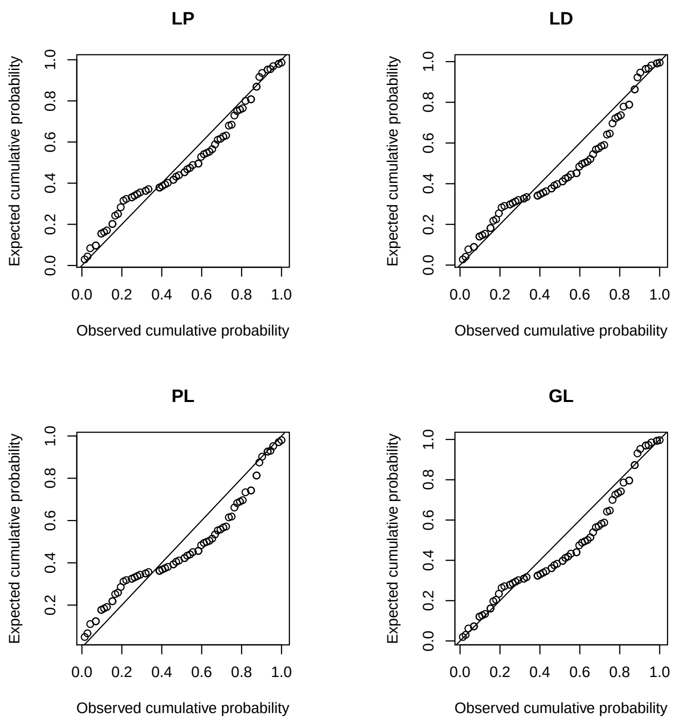


Figure 4. P-P plots for the second dataset.

by NSF-DMS-0808864, ONR-YIP-N000140910911, and a grant from Microsoft. The second author's work was partially supported by the Fundamental Research Funds for the Central Universities (2014JBM125).

References

[1] K Adamidis and S Loukas. A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, 39(1):35–42, 1998.

[2] B.C. Arnold, N. Balakrishnan, and H.N. Nagaraja. *A first course in order statistics*, volume 54. Society for Industrial Mathematics, 2008.

[3] H.S. Bakouch, B.M. Al-Zahrani, A.A. Al-Shomrani, V.A.A. Marchi, and F. Louzada. An extended lindley distribution. *Journal of the Korean Statistical Society*, 41(1):75–85, 2012.

[4] T. Bjerkedal. Acquisition of resistance in guinea pies infected with different doses of virulent tubercle bacilli. *American Journal of Hygiene*, 72(1):130–148, 1960.

[5] M Chahkandi and M Ganjali. On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53(12):4433–4440, 2009.

[6] D.R. Cox and D.V. Hinkley. *Theoretical statistics*. Chapman & Hall, London., 1979.

[7] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[8] William Feller. On the kolmogorov-smirnov limit theorems for empirical distributions. *The annals of mathematical statistics*, 19(2):177–189, 1948.

[9] ME Ghitany, DK Al-Mutairi, N Balakrishnan, and LJ Al-Enezi. Power lindley distribution and associated inference. *Computational Statistics & Data Analysis*, 64:20–33, 2013.

[10] ME Ghitany, DK Al-Mutairi, and S. Nadarajah. Zero-truncated poisson–lindley distribution and its application. *Mathematics and Computers in Simulation*, 79(3):279–287, 2008.

[11] ME Ghitany, B. Atieh, and S. Nadarajah. Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4):493–506, 2008.

[12] Ronald E Glaser. Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371):667–672, 1980.

[13] Ramesh C Gupta, Nandini Kannan, and Aparna Raychaudhuri. Analysis of lognormal survival data. *Mathematical biosciences*, 139(2):103–115, 1997.

[14] P Jodrá. Computer generation of random variables with lindley or poisson–lindley distribution via the lambert W function. *Mathematics and Computers in Simulation*, 81(4):851–859, 2010.

[15] Coşkun Kuş. A new lifetime distribution. *Computational Statistics & Data Analysis*, 51(9):4497–4509, 2007.

[16] D.V. Lindley. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(1):102–107, 1958.

[17] E. Mahmoudi and H. Zakerzadeh. Generalized poisson–lindley distribution. *Communications in Statistics Theory and Methods*, 39(10):1785–1798, 2010.

[18] Saralees Nadarajah, Hassan S Bakouch, and Rasool Tahmasbi. A generalized lindley distribution. *Sankhya B*, 73(2):331–359, 2011.

[19] C.G. Ramesh and S. Kirmani. On order relations between reliability measures. *Stochastic Models*, 3(1):149–156, 1987.

[20] M. Sankaran. The discrete poisson-lindley distribution. *Biometrics*, 26(1):145–149, 1970.

[21] Rama Shanker, Shambhu Sharma, and Ravi Shanker. A two-parameter lindley distribution for modeling waiting and survival times data. *Applied Mathematics*, 4:363–368, 2013.

[22] H. Zakerzadeh and A. Dolati. Generalized lindley distribution. *Journal of Mathematical Extension*, 3(2):13–25, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1079 – 1093

Improved exponential type estimators of finite population mean under complete and partial auxiliary information

Abdul Haq* and Javid Shabbir†

Abstract

This paper proposes some improved exponential type estimators of finite population mean under simple random sampling and double sampling. Expressions for biases and mean squared errors of the proposed estimators are derived up to the first order of approximation. Theoretical and numerical comparisons are made to investigate the performances of the estimators. The proposed estimators always perform better than the difference estimator of the population mean. They also perform better than the estimators suggested by Gupta and Shabbir [3] and Grover and Kaur [2].

2000 AMS Classification: 62D05, 62G05.

Keywords: Auxiliary variable, Bias, Mean squared error, Difference estimator, Two-phase sampling.

Received 01 : 08 : 2013 : Accepted 13 : 01 : 2014 Doi : 10.15672/HJMS.201437454

1. Introduction

The auxiliary information is frequently used to increase precision of the population estimates by taking advantage of the correlation between the study variable and the auxiliary variable. Several authors including Kadilar and Cingi [4], Kadilar and Cingi [5], Kadilar and Cingi [6], Kadilar and Cingi [7] and Gupta and Shabbir [3] have proposed different estimators by utilizing information on the auxiliary variable for estimation of the population mean.

In this paper, we propose some improved exponential type estimators for estimating finite population mean using complete and partial auxiliary information. Explicit expressions for biases and mean squared errors (*MSE*s) of the proposed estimators are derived up to the first order of approximation. An empirical study is conducted to assess the

* Department of Statistics, Quaid-i-Azam University, Islamabad, 45320, Pakistan
†Department of Statistics, Quaid-i-Azam University, Islamabad, 45320, Pakistan
Email: javidshabbir@gmail.com

performance of the proposed estimators. It is observed that the proposed estimators are more precise than the existing estimators of the finite population mean.

Consider a finite population comprises of N units. We draw a sample of size n from this population by using simple random sampling without replacement (SRSWOR). Let y and x be the study and the auxiliary variables of the characteristics y_i and x_i, respectively, for the ith unit. Let $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample means corresponding to the population means $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ and $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$, respectively. Let $s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2$ and $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2$ be the sample variances corresponding to the population variances $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N} \left(y_i - \bar{Y}\right)^2$ and $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N} \left(x_i - \bar{X}\right)^2$, respectively. Let ρ be the correlation coefficient between y and x. Let $C_y = \frac{S_y}{\bar{Y}}$ and $C_x = \frac{S_x}{\bar{X}}$ be the coefficients of variation of y and x, respectively.

The rest of the paper is organized as follows: Section 2 includes the estimators adopted by several authors when using complete auxiliary information. In Section 3, the proposed estimators based on complete information are discussed in detail. Theoretical comparisons of the proposed estimators with the existing estimators are given in Section 4. Section 5 contains some suggested estimators when partial auxiliary information is available. The work on the proposed estimators is extended to two-phase sampling in Section 6. Section 7 contains theoretical comparisons of the suggested estimators and existing estimators. For numerical comparisons of estimators, we consider three real data sets in Section 8, and concluding remarks are given in Section 9.

2. Estimators based on complete auxiliary information

In the following subsequent sections, we discuss the properties of the difference, difference-ratio-type and exponential-type estimators of finite population mean suggested by several authors.

2.1. Usual difference estimator of population mean. The unbiased difference estimator of population mean is

$$\hat{\bar{Y}}_D = \bar{y} + k\left(\bar{X} - \bar{x}\right), \tag{2.1}$$

where k is an unknown constant.
The minimum variance of $\hat{\bar{Y}}_D$, at optimum value of k, i.e., $k_{(opt)} = \frac{\bar{Y}\rho C_y}{\bar{X}C_x}$, is given by

$$Var_{\min}\left(\hat{\bar{Y}}_D\right) \cong \bar{Y}^2\lambda\left(1-\rho^2\right)C_y^2, \tag{2.2}$$

where $\lambda = \frac{1-f}{n}$ and $f = \frac{n}{N}$.

2.2. Gupta and Shabbir [3] family of estimators. Gupta and Shabbir [3] introduced the following family of estimators for estimating finite population mean:

$$\hat{\bar{Y}}_{GS} = \left\{s_1\bar{y} + s_2\left(\bar{X} - \bar{x}\right)\right\}\left(\frac{a\bar{X}+b}{a\bar{x}+b}\right), \tag{2.3}$$

where s_1 and s_2 are two unknown constants. Here a and b are the known population parameters which may be coefficient of skewness (β_{1x}), coefficient of kurtosis (β_{2x}), coefficient of variation (CV) and correlation coefficient (ρ).

Expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{GS}$, to first order of approximation, are given by

$$\text{(2.4)} \qquad Bias\left(\hat{\bar{Y}}_{GS}\right) \cong -\bar{Y} + \bar{Y}\left\{1 + \lambda\tau C_x\left(\tau C_x - \rho C_y\right)\right\} s_1 + \bar{X}\lambda\tau C_x^2 s_2$$

and

$$\text{(2.5)} \qquad \begin{aligned} MSE\left(\hat{\bar{Y}}_{GS}\right) &\cong \bar{Y}^2 + \bar{Y}^2\left\{1 + \lambda\left(3\tau^2 C_x^2 - 4\rho\tau C_x C_y + C_y^2\right)\right\} s_1^2 + \bar{X}\lambda C_x^2 s_2\left(-2\bar{Y}\tau + \bar{X}s_2\right) \\ &\quad -2\bar{Y}s_1\left[\bar{Y} + \lambda C_x\left\{\tau C_x\left(\bar{Y}\tau - 2\bar{X}s_2\right) + \rho C_y\left(-\bar{Y}\tau + \bar{X}s_2\right)\right\}\right], \end{aligned}$$

where $\tau = \frac{a\bar{X}}{a\bar{X}+b}$.

The optimum values of s_1 and s_2, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{GS}$, are given by $s_{1(opt)} = \frac{-1+\lambda\tau^2 C_x^2}{-1+\lambda\tau^2 C_x^2+\lambda\left(-1+\rho^2\right)C_y^2}$ and $s_{2(opt)} = \frac{\bar{Y}\left[-\rho C_y+\tau C_x\left\{1-\lambda\tau^2 C_x^2+\lambda\rho\tau C_x C_y+\lambda\left(-1+\rho^2\right)C_y^2\right\}\right]}{\bar{X}C_x\left\{-1+\lambda\tau^2 C_x^2+\lambda\left(-1+\rho^2\right)C_y^2\right\}}$.

The minimum *MSE* of $\hat{\bar{Y}}_{GS}$, at optimum values of s_1 and s_2, is given by

$$\text{(2.6)} \qquad MSE_{\min}\left(\hat{\bar{Y}}_{GS}\right) \cong \frac{\bar{Y}^2\lambda\left(1-\rho^2\right)\left(-1+\lambda\tau^2 C_x^2\right)C_y^2}{-1+\lambda\tau^2 C_x^2+\lambda\left(-1+\rho^2\right)C_y^2}.$$

Gupta and Shabbir [3] estimator $\hat{\bar{Y}}_{GS}$ will perform better than the difference estimator $\hat{\bar{Y}}_D$, if

$$\frac{\bar{Y}^2\lambda^2\left(-1+\rho^2\right)^2 C_y^4}{1-\lambda\tau^2 C_x^2+\lambda\left(1-\rho^2\right)C_y^2} > 0.$$

2.3. Grover and Kaur [2] **estimator.** Grover and Kaur [2] proposed the following estimator of finite population mean:

$$\text{(2.7)} \qquad \hat{\bar{Y}}_{GK} = \left\{t_1\bar{y} + t_2\left(\bar{X} - \bar{x}\right)\right\}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right),$$

where t_1 and t_2 are two unknown constants, whose values are to be determined later on.
Expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{GK}$, to first order of approximation, are given by

$$\text{(2.8)} \qquad Bias\left(\hat{\bar{Y}}_{GK}\right) \cong \frac{1}{8}\left[-8\bar{Y} + \bar{Y}\left\{8 + \lambda C_x\left(3C_x - 4\rho C_y\right)\right\}t_1 + 4\bar{X}\lambda C_x^2 t_2\right]$$

and

$$\text{(2.9)} \qquad \begin{aligned} MSE\left(\hat{\bar{Y}}_{GK}\right) &\cong \bar{Y}^2 + \bar{Y}^2\left\{1 + \lambda\left(C_x^2 - 2\rho C_x C_y + C_y^2\right)\right\}t_1^2 + \bar{X}\lambda C_x^2 t_2\left(-\bar{Y} + \bar{X}t_2\right) \\ &\quad +\frac{1}{4}\bar{Y}t_1\left[-8\bar{Y} + \lambda C_x\left\{4\rho C_y\left(\bar{Y} - 2\bar{X}t_2\right) + C_x\left(-3\bar{Y} + 8\bar{X}t_2\right)\right\}\right]. \end{aligned}$$

The optimum values of t_1 and t_2, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{GK}$, are given by $t_{1(opt)} = \frac{-8+\lambda C_x^2}{-8+8\lambda\left(-1+\rho^2\right)C_y^2}$ and $t_{2(opt)} = \frac{\bar{Y}\left[-8\rho C_y+C_x\left\{4-\lambda C_x^2+\lambda\rho C_x C_y+4\lambda\left(-1+\rho^2\right)C_y^2\right\}\right]}{8\bar{X}C_x\left\{-1+\lambda\left(-1+\rho^2\right)C_y^2\right\}}$.

The minimum *MSE* of $\hat{\bar{Y}}_{GK}$, at optimum values of t_1 and t_2, is given by

$$\text{(2.10)} \qquad MSE_{\min}\left(\hat{\bar{Y}}_{GK}\right) \cong \frac{\bar{Y}^2\lambda\left\{\lambda C_x^4 - 16\left(-1+\rho^2\right)\left(-4+\lambda C_x^2\right)C_y^2\right\}}{64\left\{-1+\lambda\left(-1+\rho^2\right)C_y^2\right\}}.$$

Grover and Kaur [2] estimator $\hat{\bar{Y}}_{GK}$ will perform better than the difference estimator $\hat{\bar{Y}}_D$, if

$$\frac{\bar{Y}^2\lambda^2\left\{C_x^2 - 8\left(-1+\rho^2\right)C_y^2\right\}^2}{64\left\{1+\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

Gupta and Shabbir [3] estimator $\hat{\bar{Y}}_{GS}$ will perform better than the Grover and Kaur [2] estimator $\hat{\bar{Y}}_{GK}$, if

$$\bar{Y}^2\lambda\left(\frac{\left(-1+\rho^2\right)\left(-1+\lambda\tau^2C_x^2\right)C_y^2}{-1+\lambda\tau^2C_x^2+\lambda\left(-1+\rho^2\right)C_y^2}+\frac{\lambda C_x^4-16\left(-1+\rho^2\right)\left(-4+\lambda C_x^2\right)C_y^2}{64\left\{-1+\lambda\left(-1+\rho^2\right)C_y^2\right\}}\right)>0.$$

3. Proposed estimators

In this section, we propose some improved exponential type estimators for estimating finite population mean when complete auxiliary information is available.

3.1. First proposed estimator. On the lines of Singh and Espejo [8], the average ratio-product estimator is given by

$$\hat{\bar{Y}}_{SE}=\frac{1}{2}\bar{y}\left(\frac{\bar{X}}{\bar{x}}+\frac{\bar{x}}{\bar{X}}\right). \tag{3.1}$$

By replacing $\hat{\bar{Y}}_{SE}$ in place of $\bar{y}$ in (2.7), the proposed estimator becomes

$$\hat{\bar{Y}}_{P1}=\left\{u_1\hat{\bar{Y}}_{SE}+u_2\left(\bar{X}-\bar{x}\right)\right\}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right), \tag{3.2}$$

where u_1 and u_2 are two unknown constants, whose values are determined for optimality. Expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{P1}$, to first order of approximation, are given by

$$Bias\left(\hat{\bar{Y}}_{P1}\right)\cong\frac{1}{8}\left[-8\bar{Y}+\bar{Y}\left\{8+\lambda C_x\left(7C_x-4\rho C_y\right)\right\}u_1+4\bar{X}\lambda C_x^2u_2\right] \tag{3.3}$$

and

$$MSE\left(\hat{\bar{Y}}_{P1}\right)\cong\bar{Y}^2+\bar{Y}^2\left\{1+\lambda\left(2C_x^2-2\rho C_xC_y+C_y^2\right)\right\}u_1^2+\bar{X}\lambda C_x^2u_2\left(-\bar{Y}+\bar{X}u_2\right)$$

$$+\frac{1}{4}\bar{Y}u_1\left[-8\bar{Y}+\lambda C_x\left\{4\rho C_y\left(\bar{Y}-2\bar{X}u_2\right)+C_x\left(-7\bar{Y}+8\bar{X}u_2\right)\right\}\right]. \tag{3.4}$$

The optimum values of u_1 and u_2, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{P1}$, are given by $u_{1(opt)}=\frac{8+3\lambda C_x^2}{8\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}}$ and $u_{2(opt)}=\frac{\bar{Y}\left[8\rho C_y+C_x\left\{-4+\lambda\left(C_x^2+3\rho C_xC_y-4\left(-1+\rho^2\right)C_y^2\right)\right\}\right]}{8\bar{X}C_x\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}}$.
The minimum *MSE* of $\hat{\bar{Y}}_{P1}$, at optimum values of u_1 and u_2, is given by

$$MSE_{\min}\left(\hat{\bar{Y}}_{P1}\right)\cong\frac{\bar{Y}^2\lambda\left\{-25\lambda C_x^4+16\left(-1+\rho^2\right)\left(-4+\lambda C_x^2\right)C_y^2\right\}}{64\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}}. \tag{3.5}$$

3.2. Second proposed estimator. On the line of Bahl and Tuteja [1], we can define the average exponential ratio-product type estimator, given by

$$\hat{\bar{Y}}_{BTW}=\frac{1}{2}\bar{y}\left\{\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)+\exp\left(\frac{\bar{x}-\bar{X}}{\bar{X}+\bar{x}}\right)\right\}, \tag{3.6}$$

By replacing $\hat{\bar{Y}}_{BTW}$ in place of $\bar{y}$ in (2.7), the proposed estimator becomes

$$\hat{\bar{Y}}_{P2}=\left\{v_1\hat{\bar{Y}}_{BTW}+v_2\left(\bar{X}-\bar{x}\right)\right\}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right), \tag{3.7}$$

where v_1 and v_2 are two unknown constants.
Expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{P2}$, to first order of approximation, are given by

$$Bias\left(\hat{\bar{Y}}_{P2}\right)\cong\frac{1}{2}\left[-2\bar{Y}+\bar{Y}\left\{2+\lambda C_x\left(C_x-\rho C_y\right)\right\}v_1+\bar{X}\lambda C_x^2v_2\right] \tag{3.8}$$

and

$$MSE\left(\hat{\hat{Y}}_{P2}\right) \cong \bar{Y}^2 + \frac{1}{4}\bar{Y}^2\left(4+5\lambda C_x^2 - 8\lambda\rho C_x C_y + 4\lambda C_y^2\right)v_1^2 + \bar{X}\lambda C_x^2 v_2\left(-\bar{Y}+\bar{X}v_2\right)$$
$$+\bar{Y}v_1\left\{-2\bar{Y} - \lambda C_x\left(C_x - \rho C_y\right)\left(\bar{Y} - 2\bar{X}v_2\right)\right\}. \qquad (3.9)$$

The optimum values of v_1 and v_2, obtained by minimizing the *MSE* of $\hat{\hat{Y}}_{P2}$, are given by $v_{1(opt)} = \frac{4}{4+\lambda C_x^2 - 4\lambda\left(-1+\rho^2\right)C_y^2}$ and $v_{2(opt)} = \frac{\bar{Y}}{2\bar{X}}\left(1 + \frac{-8C_x+8\rho C_y}{C_x\left\{4+\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2\right\}}\right)$.
The minimum *MSE* of $\hat{\hat{Y}}_{P2}$, at optimum values of v_1 and v_2, is given by

$$MSE_{\min}\left(\hat{\hat{Y}}_{P2}\right) \cong \frac{\bar{Y}^2\lambda\left\{-\lambda C_x^4 + 4\left(-1+\rho^2\right)\left(-4+\lambda C_x^2\right)C_y^2\right\}}{4\left\{4+\lambda C_x^2 - 4\lambda\left(-1+\rho^2\right)C_y^2\right\}}. \qquad (3.10)$$

3.3. Third proposed estimator. Replacing $\hat{\hat{Y}}_{SE}$ from (3.1) in place of $\bar{y}$ given in (3.6), the estimator becomes

$$\hat{\hat{Y}}_{BTSEW} = \bar{y}\frac{1}{4}\left(\frac{\bar{X}}{\bar{x}} + \frac{\bar{x}}{\bar{X}}\right)\left\{\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \exp\left(\frac{\bar{x}-\bar{X}}{\bar{X}+\bar{x}}\right)\right\}. \qquad (3.11)$$

Also replacing $\hat{\hat{Y}}_{BTSEW}$ in place of $\bar{y}$ in (2.7), the proposed estimator turns out to be

$$\hat{\hat{Y}}_{P3} = \left[w_1\hat{\hat{Y}}_{BTSEW} + w_2\left(\bar{X}-\bar{x}\right)\right]\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right), \qquad (3.12)$$

where w_1 and w_2 are two unknown constants.
Expressions for *Bias* and *MSE* of $\hat{\hat{Y}}_{P3}$, to first order of approximation, are given by

$$Bias\left(\hat{\hat{Y}}_{P3}\right) \cong \frac{1}{2}\left[-2\bar{Y} + \bar{Y}\left\{2+\lambda C_x\left(2C_x - \rho C_y\right)\right\}w_1 + \bar{X}\lambda C_x^2 w_2\right] \qquad (3.13)$$

and

$$MSE\left(\hat{\hat{Y}}_{P3}\right) \cong \bar{Y}^2 + \frac{1}{4}\bar{Y}^2\left(4+9\lambda C_x^2 - 8\lambda\rho C_x C_y + 4\lambda C_y^2\right)w_1^2 + \bar{X}\lambda C_x^2 w_2\left(-\bar{Y}+\bar{X}w_2\right)$$
$$+\bar{Y}w_1\left[-2\bar{Y} + \lambda C_x\left\{\rho C_y\left(\bar{Y}-2\bar{X}w_2\right) - 2C_x\left(\bar{Y}-\bar{X}w_2\right)\right\}\right]. \qquad (3.14)$$

The optimum values of w_1 and w_2, obtained by minimizing the *MSE* of $\hat{\hat{Y}}_{P3}$, are given by $w_{1(opt)} = \frac{4+2\lambda C_x^2}{4+5\lambda C_x^2 - 4\lambda\left(-1+\rho^2\right)C_y^2}$ and $w_{2(opt)} = \frac{\bar{Y}\left[8\rho C_y + C_x\left\{-4+\lambda\left(C_x^2+4\rho C_x C_y - 4\left(-1+\rho^2\right)C_y^2\right)\right\}\right]}{2\bar{X}C_x\left\{4+5\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2\right\}}$.
The minimum *MSE* of $\hat{\hat{Y}}_{P3}$, at optimum values of w_1 and w_2, is given by

$$MSE_{\min}\left(\hat{\hat{Y}}_{P3}\right) \cong \frac{\bar{Y}^2\lambda\left\{-9\lambda C_x^4 + 4\left(-1+\rho^2\right)\left(-4+\lambda C_x^2\right)C_y^2\right\}}{4\left\{4+5\lambda C_x^2 - 4\lambda\left(-1+\rho^2\right)C_y^2\right\}}. \qquad (3.15)$$

Remarks: Expressions given in (3.5), (3.10) and (3.15) contain unknown population parameters, which can be estimated either from the sample values or through repeated survey or by experience gathered in due course of time.

4. Efficiency comparisons under simple random sampling

In this section, we compare the proposed estimators with the existing estimators.

(**a**) Comparison with difference type estimator

(i) From (2.2) and (3.5), $MSE_{\min}\left(\hat{\hat{Y}}_{P1}\right) < Var_{\min}\left(\hat{\hat{Y}}_D\right)$, if

$$\frac{\bar{Y}^2\lambda^2\left\{5C_x^2 - 8\left(-1+\rho^2\right)C_y^2\right\}^2}{64\left\{1+\lambda C_x^2 + \lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

(ii) From (2.2) and (3.10), $MSE_{\min}\left(\hat{\bar{Y}}_{P2}\right) < Var_{\min}\left(\hat{\bar{Y}}_{D}\right)$, if

$$\frac{\bar{Y}^2\lambda^2\left\{C_x^2-4\left(-1+\rho^2\right)C_y^2\right\}^2}{4\left\{4+\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

(iii) From (2.2) and (3.15), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}\right) < Var_{\min}\left(\hat{\bar{Y}}_{D}\right)$, if

$$\frac{\bar{Y}^2\lambda^2\left\{3C_x^2-4\left(-1+\rho^2\right)C_y^2\right\}^2}{4\left\{4+5\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

Note: Conditions (i)-(iii) are always true.

(b) Comparison with Gupta and Shabbir [3] **estimator**

(iv) From (2.6) and (3.5), $MSE_{\min}\left(\hat{\bar{Y}}_{P1}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GS}\right)$, if

$$\frac{\bar{Y}^2\lambda^2C_x^2\left[25C_x^2\left(-1+\lambda\tau^2C_x^2\right)-5\left(-1+\rho^2\right)\left\{-16+\lambda\left(-5+16\tau^2\right)C_x^2\right\}C_y^2+16\lambda\left(-1+\rho^2\right)^2\left(-1+4\tau^2\right)C_y^4\right]}{64\left\{-1+\lambda\tau^2C_x^2+\lambda\left(-1+\rho^2\right)C_y^2\right\}\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

(v) From (2.6) and (3.10), $MSE_{\min}\left(\hat{\bar{Y}}_{P2}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GS}\right)$, if

$$\frac{1}{4}\bar{Y}^2\left\{\lambda\left(1-4\tau^2\right)C_x^2+\frac{16}{4+\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2}+\frac{4\left(-1+\lambda\tau^2C_x^2\right)^2}{-1+\lambda\tau^2C_x^2+\lambda\left(-1+\rho^2\right)C_y^2}\right\} > 0.$$

(vi) From (2.6) and (3.15), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GS}\right)$, if

$$\frac{\bar{Y}^2\lambda^2C_x^2\left[9C_x^2\left(-1+\lambda\tau^2C_x^2\right)-3\left(-1+\rho^2\right)\left\{-8+\lambda\left(-3+8\tau^2\right)C_x^2\right\}C_y^2+4\lambda\left(-1+\rho^2\right)^2\left(-1+4\tau^2\right)C_y^4\right]}{4\left\{4+5\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2\right\}\left\{-1+\lambda\tau^2C_x^2+\lambda\left(-1+\rho^2\right)C_y^2\right\}} > 0.$$

Note: The proposed estimators $\hat{\bar{Y}}_{Pi}(i=1,2,3)$ perform better than the Gupta and Shabbir [3] if conditions (iv)-(vi) are satisfied.

(c) Comparison with Grover and Kaur [2] **estimator**

(vii) From (2.10) and (3.5), $MSE_{\min}\left(\hat{\bar{Y}}_{P1}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GK}\right)$, if

$$\frac{\bar{Y}^2\lambda^2C_x^2\left\{C_x^2\left(-24+\lambda C_x^2\right)+8\left(-1+\rho^2\right)\left(8+\lambda C_x^2\right)C_y^2\right\}}{64\left\{-1+\lambda\left(-1+\rho^2\right)C_y^2\right\}\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

(viii) From (2.10) and (3.10), $MSE_{\min}\left(\hat{\bar{Y}}_{P2}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GK}\right)$, if

$$\bar{Y}^2\left(\frac{4}{4+\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2}+\frac{\left(-8+\lambda C_x^2\right)^2}{-64+64\lambda\left(-1+\rho^2\right)C_y^2}\right) > 0.$$

(ix) From (2.10) and (3.15), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{GK}\right)$, if

$$\frac{5\bar{Y}^2\lambda^2C_x^2\left\{C_x^2\left(-28+\lambda C_x^2\right)+4\left(-1+\rho^2\right)\left(16+3\lambda C_x^2\right)C_y^2\right\}}{64\left\{4+5\lambda C_x^2-4\lambda\left(-1+\rho^2\right)C_y^2\right\}\left\{-1+\lambda\left(-1+\rho^2\right)C_y^2\right\}} > 0.$$

Note: The proposed estimators $\hat{\bar{Y}}_{Pi}(i=1,2,3)$ perform better than the Grover and Kaur (2011) if conditions (vii)-(ix) are satisfied.

(d) Comparisons among proposed estimators

(x) From (3.5) and (3.10), $MSE_{\min}\left(\hat{\bar{Y}}_{P2}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{p1}\right)$, if

$$\bar{Y}^2\left(\frac{4}{4+\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2}-\frac{\left(8+3\lambda C_x^2\right)^2}{64\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}}\right) > 0.$$

(xi) From (3.5) and (3.15), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{p1}\right)$, if

$$\frac{\bar{Y}^2\lambda^2 C_x^2\left\{C_x^2\left(44+19\lambda C_x^2\right)+4\left(1-\rho^2\right)\left(16+7\lambda C_x^2\right)C_y^2\right\}}{64\left\{4+5\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2\right\}\left\{1+\lambda C_x^2+\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

(xii) From (3.10) and (3.15), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}\right) < MSE_{\min}\left(\hat{\bar{Y}}_{p2}\right)$, if

$$\frac{\bar{Y}^2\lambda^2 C_x^2\left\{C_x^2\left(8+\lambda C_x^2\right)+4\left(1-\rho^2\right)\left(4+\lambda C_x^2\right)C_y^2\right\}}{\left\{4+\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2\right\}\left\{4+5\lambda C_x^2+4\lambda\left(1-\rho^2\right)C_y^2\right\}} > 0.$$

Note: Conditions (xi) and (xii) are always true.

5. Estimators under two-phase sampling (partial information)

When the population mean of the auxiliary variable, x, is unknown, it is customary to apply the two-phase sampling procedure. The two-phase sampling scheme is explained as follows

(i) In first-phase, a sample of size $(n_1 < N)$ is selected from the population using SRSWOR to estimate $\bar{X}$.

(ii) In second-phase, a sample of size $(n < n_1)$ is selected to observe both y and x.

Let $\bar{x}_1$ be the sample mean based on first-phase sample of size n_1, and let $\bar{y}$ and $\bar{x}$ be the sample means based on second-phase sample of size n. Let $(\bar{x}_1, \bar{x})$ and $\bar{y}$ are the unbiased estimators of $\bar{X}$ and $\bar{Y}$, respectively. Now we discuss different estimators of finite population mean based on two-phase sampling.

5.1. Unbiased difference estimator. The unbiased difference estimator of population mean under two-phase sampling is

$$\hat{\bar{Y}}_D^* = \bar{y} + k^*\left(\bar{x}_1 - \bar{x}\right), \tag{5.1}$$

where k^* is an unknown constant.
The expression for variance of $\hat{\bar{Y}}_D^*$, at optimum value of k^*, i.e., $k^*_{(opt)} = \frac{\bar{Y}\rho C_y}{\bar{X}C_x}$ is given by

$$Var_{\min}\left(\hat{\bar{Y}}_D^*\right) \cong \bar{Y}^2\left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2, \tag{5.2}$$

where $\lambda_1 = \frac{1}{n} - \frac{1}{n_1}$.

5.2. Gupta and Shabbir [3] family of estimators. Under two-phase sampling, Gupta and Shabbir [3] family of estimators for estimating finite population mean, is given by

$$\hat{\bar{Y}}_{GS}^* = \left\{s_1^*\bar{y} + s_2^*\left(\bar{x}_1 - \bar{x}\right)\right\}\left(\frac{a\bar{x}_1 + b}{a\bar{x} + b}\right), \tag{5.3}$$

where s_1^* and s_2^* are two unknown constants.
The expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{GS}^*$, to first order of approximation, are given by

$$Bias\left(\hat{\bar{Y}}_{GS}^*\right) \cong -\bar{Y} + \bar{Y}\left\{1 + \left(\lambda - \lambda_1\right)\tau C_x\left(\tau C_x - \rho C_y\right)\right\}s_1^* + \bar{X}\lambda\tau C_x^2 s_2^* \tag{5.4}$$

and

$$MSE\left(\hat{\bar{Y}}_{GS}^*\right) \cong \bar{Y}^2 + \bar{Y}^2\left\{1 + 3\left(\lambda - \lambda_1\right)\tau^2 C_x^2 + 4\left(-\lambda + \lambda_1\right)\rho\tau C_x C_y + \lambda C_y^2\right\}s_1^{*2}$$
$$+\bar{X}\left(\lambda - \lambda_1\right)C_x^2 s_2^*\left(-2\bar{Y}\tau + \bar{X}s_2^*\right) - 2\bar{Y}s_1^*\left[\bar{Y} + \left(\lambda - \lambda_1\right)C_x\left\{\tau C_x\left(\bar{Y}\tau - 2\bar{X}s_2^*\right) + \rho C_y\left(-\bar{Y}\tau + \bar{X}s_2^*\right)\right\}\right], \tag{5.5}$$

where τ is defined earlier.
The optimum values of s_1^* and s_2^*, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{GS}^*$, are given by

$s^*_{1(opt)} = \frac{-1+(\lambda-\lambda_1)\tau^2 C_x^2}{-1+(\lambda-\lambda_1)\tau^2 C_x^2+\{-\lambda+(\lambda-\lambda_1)\}C_y^2}$ and
$s^*_{2(opt)} = \frac{\bar{Y}\left[-\rho C_y+\tau C_x\left\{1+(-\lambda+\lambda_1)\tau^2 C_x^2+(\lambda-\lambda_1)\rho\tau C_x C_y+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right\}\right]}{\bar{X}C_x\left\{-1+(\lambda-\lambda_1)\tau^2 C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right\}}$.
The minimum *MSE* of $\hat{\bar{Y}}^*_{GS}$, at optimum values of s_1^* and s_2^*, is given by

$$\text{(5.6)} \qquad MSE_{\min}\left(\hat{\bar{Y}}^*_{GS}\right) \cong \frac{\bar{Y}^2\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{1-(\lambda-\lambda_1)\tau^2 C_x^2\right\}C_y^2}{-1+(\lambda-\lambda_1)\tau^2 C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}.$$

Gupta and Shabbir [3] estimator $\hat{\bar{Y}}^*_{GS}$ will perform better than the difference estimator $\hat{\bar{Y}}^*_{D}$, if

$$\frac{\bar{Y}^2\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)^2 C_y^4}{1-(\lambda-\lambda_1)\tau^2 C_x^2-\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2} > 0.$$

5.3. Grover and Kaur [2] **estimator.** Grover and Kaur [2] estimator under double sampling for estimation of the population mean is given by

$$\text{(5.7)} \qquad \hat{\bar{Y}}^*_{GK} = \left\{t_1^*\bar{y} + t_2^*(\bar{x}_1-\bar{x})\right\}\exp\left(\frac{\bar{x}_1-\bar{x}}{\bar{x}_1+\bar{x}}\right),$$

where t_1^* and t_2^* are two unknown constants.
The expressions for *Bias* and *MSE* of $\hat{\bar{Y}}^*_{GK}$, to first order of approximation, are given by

(5.8)
$$Bias\left(\hat{\bar{Y}}^*_{GK}\right) \cong \frac{1}{8}\left[-8\bar{Y}+\bar{Y}\left\{8+(\lambda-\lambda_1)C_x(3C_x-4\rho C_y)\right\}t_1^*+4\bar{X}(\lambda-\lambda_1)C_x^2 t_2^*\right]$$

and

$$MSE\left(\hat{\bar{Y}}^*_{GK}\right) \cong \bar{Y}^2+\bar{Y}^2\left\{1+(\lambda-\lambda_1)C_x^2+2(-\lambda+\lambda_1)\rho C_x C_y+\lambda C_y^2\right\}t_1^{*2}$$
$$\text{(5.9)} \qquad +\bar{X}(\lambda-\lambda_1)C_x^2 t_2^*\left(-\bar{Y}+\bar{X}t_2^*\right)+\frac{1}{4}\bar{Y}t_1^*\left[-8\bar{Y}+(\lambda-\lambda_1)C_x\left\{4\rho C_y\left(\bar{Y}-2\bar{X}t_2^*\right)+C_x\left(-3\bar{Y}+8\bar{X}t_2^*\right)\right\}\right].$$

The optimum values of t_1^* and t_2^*, obtained by minimizing the *MSE* of $\hat{\bar{Y}}^*_{GK}$, are given by
$t^*_{1(opt)} = \frac{-8+(\lambda-\lambda_1)C_x^2}{-8+8\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}$ and $t^*_{2(opt)} = \frac{\bar{Y}\left[8\rho C_y+C_x\left\{-4+(\lambda-\lambda_1)C_x^2+(-\lambda+\lambda_1)\rho C_x C_y+4\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)C_y^2\right\}\right]}{8\bar{X}C_x\left\{1+\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)C_y^2\right\}}$.
The minimum *MSE* of $\hat{\bar{Y}}^*_{GK}$, at optimum values of t_1^* and t_2^*, is given by

(5.10)
$$MSE_{\min}\left(\hat{\bar{Y}}^*_{GK}\right) \cong \frac{\bar{Y}^2\left\{(\lambda-\lambda_1)^2 C_x^4-16\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left(-4+(\lambda-\lambda_1)C_x^2\right)C_y^2\right\}}{-64+64\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}.$$

Grover and Kaur [2] estimator $\hat{\bar{Y}}^*_{GK}$ will perform better than the difference estimator $\hat{\bar{Y}}^*_{D}$, if

$$\frac{\bar{Y}^2\left\{(\lambda-\lambda_1)C_x^2+8\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)C_y^2\right\}^2}{64\left\{1+\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)C_y^2\right\}} > 0.$$

Gupta and Shabbir [3] estimator $\hat{\bar{Y}}^*_{GS}$ will perform better than the Grover and Kaur [2] estimator $\hat{\bar{Y}}^*_{GK}$, if

$$\bar{Y}^2\left[\frac{\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left(-1+(\lambda-\lambda_1)\tau^2 C_x^2\right)C_y^2}{-1+(\lambda-\lambda_1)\tau^2 C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}\right.$$
$$\left.+\frac{(\lambda-\lambda_1)^2 C_x^4-16\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{-4+(\lambda-\lambda_1)C_x^2\right\}C_y^2}{64\left\{-1+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right\}}\right] > 0.$$

6. Proposed estimators under two-phase sampling

In this section, we derive the mathematical expressions of the biases and MSEs of the proposed estimators of finite population mean when partial auxiliary information is available.

6.1. First proposed estimator. Similar to (3.2), the proposed estimator under double sampling is given by

$$\hat{\bar{Y}}_{P1}^{*} = \left\{ u_1^{*} \frac{1}{2} \bar{y} \left(\frac{\bar{x}_1}{\bar{x}} + \frac{\bar{x}}{\bar{x}_1} \right) + u_2^{*} \left(\bar{x}_1 - \bar{x} \right) \right\} \exp\left(\frac{\bar{x}_1 - \bar{x}}{\bar{x}_1 + \bar{x}} \right), \tag{6.1}$$

where u_1^{*} and u_2^{*} are two unknown constants.
The expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{P1}^{*}$, to first order of approximation, are given by

$$Bias\left(\hat{\bar{Y}}_{P1}^{*}\right) \cong \frac{1}{8}\left[-8\bar{Y} + \bar{Y}\left\{ 8 + (\lambda - \lambda_1) C_x (7C_x - 4\rho C_y) \right\} u_1^{*} + 4\bar{X}(\lambda - \lambda_1) C_x^2 u_2^{*} \right] \tag{6.2}$$

and

$$MSE\left(\hat{\bar{Y}}_{P1}^{*}\right) \cong \bar{Y}^2 + \bar{Y}^2\left\{ 1 + 2(\lambda - \lambda_1) C_x^2 + 2(-\lambda + \lambda_1)\rho C_x C_y + \lambda C_y^2 \right\} u_1^{*2} + \bar{X}(\lambda - \lambda_1) C_x^2 u_2^{*}\left(-\bar{Y} + \bar{X}u_2^{*}\right) + \frac{1}{4}\bar{Y}u_1^{*}\left[-8\bar{Y} + (\lambda - \lambda_1) C_x \left\{ 4\rho C_y \left(\bar{Y} - 2\bar{X}u_2^{*}\right) + C_x\left(-7\bar{Y} + 8\bar{X}u_2^{*}\right) \right\} \right]. \tag{6.3}$$

The optimum values of u_1^{*} and u_2^{*}, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{P1}^{*}$, are given by
$u_{1(opt)}^{*} = \frac{8 + 3(\lambda - \lambda_1)C_x^2}{8\left\{1 + (\lambda - \lambda_1)C_x^2 + \left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2\right\}}$ and
$u_{2(opt)}^{*} = \frac{\bar{Y}\left[8\rho C_y + C_x\left\{-4 + (\lambda - \lambda_1)C_x^2 + 3(\lambda - \lambda_1)\rho C_x C_y + 4\left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2\right\}\right]}{8\bar{X}C_x\left\{1 + (\lambda - \lambda_1)C_x^2 + \left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2\right\}}$.
The minimum *MSE* of $\hat{\bar{Y}}_{P1}^{*}$, at optimum values of u_1^{*} and u_2^{*}, is given by

$$MSE_{\min}\left(\hat{\bar{Y}}_{P1}^{*}\right) \cong \frac{\bar{Y}^2\left\{ -25(\lambda - \lambda_1)^2 C_x^4 + 16\left\{ -\lambda + (\lambda - \lambda_1)\rho^2 \right\}\left\{ -4 + (\lambda - \lambda_1) C_x^2 \right\} C_y^2 \right\}}{64\left\{ 1 + (\lambda - \lambda_1) C_x^2 + (\lambda - \lambda\rho^2 + \lambda_1\rho^2) C_y^2 \right\}}. \tag{6.4}$$

6.2. Second proposed estimator. On the line of (3.7), the second proposed estimator under double sampling is given by

$$\hat{\bar{Y}}_{P2}^{*} = \left[v_1^{*}\frac{1}{2}\bar{y}\left\{ \exp\left(\frac{\bar{x}_1 - \bar{x}}{\bar{x}_1 + \bar{x}} \right) + \exp\left(\frac{\bar{x} - \bar{x}_1}{\bar{x}_1 + \bar{x}} \right) \right\} + v_2^{*}\left(\bar{x}_1 - \bar{x}\right) \right] \exp\left(\frac{\bar{x}_1 - \bar{x}}{\bar{x}_1 + \bar{x}} \right), \tag{6.5}$$

where v_1^{*} and v_2^{*} are two unknown constants.
The expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{P2}^{*}$, to first order of approximation, are given by

$$Bias\left(\hat{\bar{Y}}_{P2}^{*}\right) \cong \frac{1}{2}\left[-2\bar{Y} + \bar{Y}\left\{ 2 + (\lambda - \lambda_1) C_x (C_x - \rho C_y) \right\} v_1^{*} + \bar{X}(\lambda - \lambda_1) C_x^2 v_2^{*} \right] \tag{6.6}$$

and

$$MSE\left(\hat{\bar{Y}}_{P2}^{*}\right) \cong \bar{Y}^2 + \frac{1}{4}\bar{Y}^2\left\{ 4 + 5(\lambda - \lambda_1) C_x^2 + 8(-\lambda + \lambda_1)\rho C_x C_y + 4\lambda C_y^2 \right\} v_1^{*2} + \bar{X}(\lambda - \lambda_1) C_x^2 v_2^{*}\left(-\bar{Y} + \bar{X}v_2^{*}\right) + \bar{Y}v_1^{*}\left\{ -2\bar{Y} - (\lambda - \lambda_1) C_x (C_x - \rho C_y)\left(\bar{Y} - 2\bar{X}v_2^{*}\right) \right\}. \tag{6.7}$$

The optimum values of v_1^{*} and v_2^{*}, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{P2}^{*}$, are given by
$v_{1(opt)}^{*} = \frac{4}{4 + (\lambda - \lambda_1)C_x^2 + 4\left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2}$ and $v_{2(opt)}^{*} = \frac{\bar{Y}}{2\bar{X}}\left(1 + \frac{-8C_x + 8\rho C_y}{C_x\left\{4 + (\lambda - \lambda_1)C_x^2 + 4\left(\lambda - \lambda\rho^2 + \lambda_1\rho^2\right)C_y^2\right\}}\right)$.
The minimum *MSE* of $\hat{\bar{Y}}_{P2}^{*}$, at optimum values of v_1^{*} and v_2^{*}, is given by

(6.8)

$$MSE_{\min}\left(\hat{\bar{Y}}_{P2}^{*}\right)\cong\frac{1}{4}\bar{Y}^{2}\left\{4+\left(-\lambda+\lambda_{1}\right)C_{x}^{2}-\frac{16}{4+\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}}\right\}.$$

6.3. Third proposed estimator. On the line of (3.12), the third proposed estimator of the population mean under double sampling is given by

(6.9)

$$\hat{\bar{Y}}_{P3}^{*}=\left[w_{1}^{*}\frac{1}{4}\bar{y}\left(\frac{\bar{x}_{1}}{\bar{x}}+\frac{\bar{x}}{\bar{x}_{1}}\right)\left\{\exp\left(\frac{\bar{x}_{1}-\bar{x}}{\bar{x}_{1}+\bar{x}}\right)+\exp\left(\frac{\bar{x}-\bar{x}_{1}}{\bar{x}_{1}+\bar{x}}\right)\right\}+w_{2}^{*}\left(\bar{x}_{1}-\bar{x}\right)\right]\exp\left(\frac{\bar{x}_{1}-\bar{x}}{\bar{x}_{1}+\bar{x}}\right),$$

where w_1^* and w_2^* are two unknown constants.
The expressions for *Bias* and *MSE* of $\hat{\bar{Y}}_{P3}^{*}$, to first order of approximation, are given by

(6.10)

$$Bias\left(\hat{\bar{Y}}_{P3}^{*}\right)\cong\frac{1}{2}\left[-2\bar{Y}+\bar{Y}\left\{2+\left(\lambda-\lambda_{1}\right)C_{x}\left(2C_{x}-\rho C_{y}\right)\right\}w_{1}^{*}+\bar{X}\left(\lambda-\lambda_{1}\right)C_{x}^{2}w_{2}^{*}\right]$$

and

$$MSE\left(\hat{\bar{Y}}_{P3}^{*}\right)\cong\bar{Y}^{2}+\frac{1}{4}\bar{Y}^{2}\left(4+9\left(\lambda-\lambda_{1}\right)C_{x}^{2}+8\left(-\lambda+\lambda_{1}\right)\rho C_{x}C_{y}+4\lambda C_{y}^{2}\right)w_{1}^{*2}$$

(6.11) $$+\bar{X}\left(\lambda-\lambda_{1}\right)C_{x}^{2}w_{2}^{*}\left(-\bar{Y}+\bar{X}w_{2}^{*}\right)+\bar{Y}w_{1}^{*}\left[-2\bar{Y}+\left(\lambda-\lambda_{1}\right)C_{x}\left\{\rho C_{y}\left(\bar{Y}-2\bar{X}w_{2}^{*}\right)-2C_{x}\left(\bar{Y}-\bar{X}w_{2}^{*}\right)\right\}\right].$$

The optimum values of w_1^* and w_2^*, obtained by minimizing the *MSE* of $\hat{\bar{Y}}_{P3}^{*}$, are given by
$w_{1(opt)}^{*}=\frac{4+2(\lambda-\lambda_{1})C_{x}^{2}}{4+5(\lambda-\lambda_{1})C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}}$ and $w_{2(opt)}^{*}=\frac{\bar{Y}\left[8\rho C_{y}+C_{x}\left\{-4+(\lambda-\lambda_{1})C_{x}^{2}+4(\lambda-\lambda_{1})\rho C_{x}C_{y}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}\right]}{2\bar{X}C_{x}\left\{4+5(\lambda-\lambda_{1})C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}}$.
The minimum *MSE* of $\hat{\bar{Y}}_{P3}^{*}$, at optimum values of w_1^* and w_2^*, is given by

(6.12)

$$MSE_{\min}\left(\hat{\bar{Y}}_{P3}^{*}\right)\cong\frac{\bar{Y}^{2}\left\{-9\left(\lambda-\lambda_{1}\right)^{2}C_{x}^{4}+4\left\{-\lambda+\left(\lambda-\lambda_{1}\right)\rho^{2}\right\}\left(-4+\left(\lambda-\lambda_{1}\right)C_{x}^{2}\right)C_{y}^{2}\right\}}{4\left\{4+5\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}}.$$

Remarks: Expressions given in (6.4), (6.8) and (6.12) contain the unknown population parameters, which can be estimated either from the sample values or through repeated survey or by experience gathered in due course of time.

7. Efficiency comparisons under two-phase sampling

In this section, we compare the proposed estimators with the existing estimators of population mean based on double sampling scheme.

(a) **Comparison with difference type estimator**

(i) From (5.2) and (6.4), $MSE_{\min}\left(\hat{\bar{Y}}_{P1}^{*}\right)<Var_{\min}\left(\hat{\bar{Y}}_{D}^{*}\right)$, if

$$\frac{\left\{5\bar{Y}\left(\lambda-\lambda_{1}\right)C_{x}^{2}+8\bar{Y}\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}^{2}}{64\left\{1+\left(\lambda-\lambda_{1}\right)C_{x}^{2}+\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}}>0.$$

(ii) From (5.2) and (6.8), $MSE_{\min}\left(\hat{\bar{Y}}_{P2}^{*}\right)<Var_{\min}\left(\hat{\bar{Y}}_{D}^{*}\right)$, if

$$\frac{\bar{Y}^{2}}{4}\left[-4+\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}+\frac{16}{4+\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}}\right]>0.$$

(iii) From (5.2) and (6.12), $MSE_{\min}\left(\hat{\bar{Y}}_{P3}^{*}\right)<Var_{\min}\left(\hat{\bar{Y}}_{D}^{*}\right)$, if

$$\frac{\left\{3\bar{Y}\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\bar{Y}\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}^{2}}{4\left\{4+5\left(\lambda-\lambda_{1}\right)C_{x}^{2}+4\left(\lambda-\lambda\rho^{2}+\lambda_{1}\rho^{2}\right)C_{y}^{2}\right\}}>0.$$

Note: When conditions (i)-(iii) are satisfied, the proposed estimators $\hat{\hat{Y}}_{Pi}^{*}(i=1,2,3)$ perform better than difference type estimator $\hat{\hat{Y}}_{D}^{*}$.

(b) Comparison with Gupta and Shabbir [3] **estimator**

(iv) From (5.6) and (6.4), $MSE_{\min}\left(\hat{\hat{Y}}_{P1}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GS}^{*}\right)$, if

$$\frac{\bar{Y}^2\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{-1+(\lambda-\lambda_1)\tau^2C_x^2\right\}C_y^2}{1-(\lambda-\lambda_1)\tau^2C_x^2+\left\{\lambda-(\lambda-\lambda_1)\rho^2\right\}C_y^2}$$
$$-\frac{\bar{Y}^2\left[-25(\lambda-\lambda_1)^2C_x^4+16\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{-4+(\lambda-\lambda_1)C_x^2\right\}C_y^2\right]}{64\left\{1+(\lambda-\lambda_1)C_x^2+(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2\right\}}>0.$$

(v) From (5.6) and (6.8), $MSE_{\min}\left(\hat{\hat{Y}}_{P2}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GS}^{*}\right)$, if

$$\frac{1}{4}\bar{Y}^2\left(-4+(\lambda-\lambda_1)C_x^2-\frac{4\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{-1+(\lambda-\lambda_1)\tau^2C_x^2\right\}C_y^2}{-1+(\lambda-\lambda)\tau^2C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}\right.$$
$$\left.+\frac{16}{4+(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right)>0.$$

(vi) From (5.6) and (6.12), $MSE_{\min}\left(\hat{\hat{Y}}_{P3}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GS}^{*}\right)$, if

$$\frac{\bar{Y}^2}{4}\left(\frac{4\left\{\lambda-(\lambda-\lambda_1)\rho^2\right\}\left\{-1+(\lambda-\lambda_1)\tau^2C_x^2\right\}C_y^2}{-1+(\lambda-\lambda_1)\tau^2C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}\right.$$
$$\left.-\frac{-9(\lambda-\lambda_1)^2C_x^4+4\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{-4+(\lambda-\lambda_1)C_x^2\right\}C_y^2}{4+5(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right)>0,$$

Note: When conditions (iv)-(vi) are satisfied, the proposed estimators $\hat{\hat{Y}}_{Pi}^{*}(i=1,2,3)$ perform better than the Gupta and Shabbir [3] estimator $\hat{\hat{Y}}_{GS}^{*}$.

(c) Comparison with Grover and Kaur [2] **estimator**

(vii) From (5.10) and (6.4), $MSE_{\min}\left(\hat{\hat{Y}}_{P1}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GK}^{*}\right)$, if

$$\frac{\bar{Y}^2(\lambda-\lambda_1)^2C_x^2\left[(\lambda-\lambda_1)C_x^2\left\{24+(-\lambda+\lambda_1)C_x^2\right\}-8\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{8+(\lambda-\lambda_1)C_x^2\right\}C_y^2\right]}{64\left[-1+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right]\left[-1+(-\lambda+\lambda_1)C_x^2+\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right]}>0,$$

when above condition is satisfied, the estimator $\hat{\hat{Y}}_{P1}^{*}$ is more efficient than $\hat{\hat{Y}}_{GK}^{*}$.

(viii) From (5.10) and (6.8), $MSE_{\min}\left(\hat{\hat{Y}}_{P2}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GK}^{*}\right)$, if

$$\frac{\bar{Y}^2}{4}\left(-4+\frac{16(-\lambda+\lambda_1)C_x^2+(\lambda-\lambda_1)^2C_x^4+64\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}{-16+16\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2}\right.$$
$$\left.+\frac{16}{4+(\lambda-\lambda_1)C_x^2+4\left\{\lambda-\lambda\rho^2+\lambda_1\rho^2\right\}C_y^2}\right)>0,$$

when above condition is satisfied, the estimator $\hat{\hat{Y}}_{P2}^{*}$ is more efficient than $\hat{\hat{Y}}_{GK}^{*}$.

(ix) From (5.10) and (6.12), $MSE_{\min}\left(\hat{\hat{Y}}_{P3}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{GK}^{*}\right)$, if

$$\frac{5\bar{Y}^2(\lambda-\lambda_1)C_x^2\left[(\lambda-\lambda_1)C_x^2\left\{28+(-\lambda+\lambda_1)C_x^2\right\}-4\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{16+3(\lambda-\lambda_1)C_x^2\right\}C_y^2\right]}{64\left[1-\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right]\left[4+5(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2\right]}>0,$$

Note: When conditions (vii)-(ix) are satisfied, the proposed estimators $\hat{\hat{Y}}_{Pi}^{*}(i=1,2,3)$ perform better than the Grover and Kaur [2] estimator $\hat{\hat{Y}}_{GK}^{*}$.

(d) **Comparisons among proposed estimators**

(x) From (6.4) and (6.8), $MSE_{\min}\left(\hat{\hat{Y}}_{P2}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{p1}^{*}\right)$, if

$$\frac{\bar{Y}^2}{64}\left(-64+\frac{256}{4+(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right.$$
$$\left.+\frac{(\lambda-\lambda_1)C_x^2\left\{16+9(-\lambda+\lambda_1)C_x^2\right\}+64\left(\lambda-\lambda\rho^2+\lambda_1\rho^2\right)C_y^2}{1+(\lambda-\lambda_1)C_x^2+(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right)>0,$$

when above condition is satisfied, the estimator $\hat{\hat{Y}}_{P2}^{*}$ is more efficient than $\hat{\hat{Y}}_{P1}^{*}$.

(xi) From (6.4) and (6.12), $MSE_{\min}\left(\hat{\hat{Y}}_{P3}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{P1}^{*}\right)$, if

$$\frac{\bar{Y}^2(\lambda-\lambda_1)C_x^2\left[(\lambda-\lambda_1)C_x^2\left\{44+19(\lambda-\lambda_1)C_x^2\right\}-4\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}\left\{16+7(\lambda-\lambda_1)C_x^2\right\}C_y^2\right]}{64\left[1+(\lambda-\lambda_1)C_x^2+\left\{\lambda-\lambda\rho^2+\lambda_1\rho^2\right\}C_y^2\right]\left[4+5(\lambda-\lambda_1)C_x^2+4\left\{\lambda-\lambda\rho^2+\lambda_1\rho^2\right\}C_y^2\right]}>0,$$

when above condition is satisfied, the estimator $\hat{\hat{Y}}_{P3}^{*}$ is more efficient than $\hat{\hat{Y}}_{P1}^{*}$.

(xii) From (6.8) and (6.12), $MSE_{\min}\left(\hat{\hat{Y}}_{p3}^{*}\right) < MSE_{\min}\left(\hat{\hat{Y}}_{p2}^{*}\right)$, if

$$\frac{\bar{Y}^2}{4}\left(4-\frac{16}{4+(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right.$$
$$\left.+\frac{4\left[(-\lambda+\lambda_1)C_x^2+(\lambda-\lambda_1)^2C_x^4+4\left\{-\lambda+(\lambda-\lambda_1)\rho^2\right\}C_y^2\right]}{4+5(\lambda-\lambda_1)C_x^2+4(\lambda-\lambda\rho^2+\lambda_1\rho^2)C_y^2}\right)>0,$$

when above condition is satisfied, the estimator $\hat{\hat{Y}}_{P3}^{*}$ is more efficient than $\hat{\hat{Y}}_{P2}^{*}$.

8. Empirical Study

The empirical study is based on three populations under: (i) complete information case and (ii) incomplete information case.

8.1. Complete auxiliary information. In this section, we compare the estimators numerically by using different real life data sets. The values of minimum MSEs of the estimators are given in Tables 1-3 based on the Populations I-III, respectively.

Population 1: [source: Kadilar and Cingi [5]].
The summary statistics are: $N=200$, $n=50$, $\bar{Y}=500$, $\bar{X}=25$, $C_y=15$, $C_x=2$, $\rho=0.90$, $\beta_{2x}=50$, $\lambda=0.015$.

Population 2: [source: Kadilar and Cingi [6]].
Let y =level of apple production (1 unit = 100 tones) and x =number of trees (1 unit = 100 trees). The data statistics are: $N=106$, $n=20$, $\bar{Y}=2212.59$, $\bar{X}=27421.70$, $C_y=5.22$, $C_x=2.10$, $\rho=0.86$, $\beta_{2x}=34.57$, $\lambda=0.040566$.

Population 3: [source: Kadilar and Cingi [7]].
Let y =level of apple production (1 unit = 100 tones) and x = number of trees. The data statistics are: $N=104$, $n=20$, $\bar{Y}=6.254$, $\bar{X}=13931.683$, $C_y=1.866$, $C_x=1.653$,

$\rho = 0.865$, $\beta_{2x} = 17.516$, $\lambda = 0.040385$.

Under complete information case, the minimum MSE values of the proposed and existing estimators are given in Table 1.
For $\hat{\bar{Y}}_{GS(1)}$ with $(a = 1, b = \rho)$, $\hat{\bar{Y}}_{GS(2)}$ with $(a = 1, b = C_x)$, $\hat{\bar{Y}}_{GS(3)}$ with $(a = 1, \beta_{2x})$,

Table 1. Minimum *MSE* values of different estimators (complete information).

Estimator	Population-I	Population-II	Population-III
$\hat{\bar{Y}}_D$	160313.00	1409112.00	1.38
$\hat{\bar{Y}}_{GS(1)}$	95468.40	1043370.00	1.33
$\hat{\bar{Y}}_{GS(2)}$	95650.43	1043380.00	1.33
$\hat{\bar{Y}}_{GS(3)}$	97421.62	1043510.00	1.33
$\hat{\bar{Y}}_{GS(4)}$	95308.27	1043370.00	1.33
$\hat{\bar{Y}}_{GS(5)}$	97099.35	1043440.00	1.33
$\hat{\bar{Y}}_{GK}$	96203.40	1043340.00	1.29
$\hat{\bar{Y}}_{P1}$	92612.00	876024.00	1.01
$\hat{\bar{Y}}_{P2}$	95306.60	1002810.00	1.24
$\hat{\bar{Y}}_{P3}$	91712.50	832286.00	0.92

$\hat{\bar{Y}}_{GS(4)}$ with $(a = \beta_{2x}, b = C_x)$, and $\hat{\bar{Y}}_{GS(5)}$ with $(a = C_x, b = \beta_{2x})$.

8.2. Summary statistics under two-phase sampling (partial information). Population 1: [source: Kadilar and Cingi [5]].
The summary statistics are: $N = 200$, $n_1 = 90$, $n = 50$, $\bar{Y} = 500$, $\bar{X} = 25$, $C_y = 15$, $C_x = 2$, $\rho = 0.90$, $\beta_{2x} = 50$, $\lambda = 0.015$.

Population 2: [source: Kadilar and Cingi [6]].
Let y =level of apple production (1 unit = 100 tones) and x =number of trees (1 unit = 100 trees). The summary statistics are: $N = 106$, $n_1 = 40$, $n = 20$, $\bar{Y} = 2212.59$, $\bar{X} = 27421.70$, $C_y = 5.22$, $C_x = 2.10$, $\rho = 0.86$, $\beta_{2x} = 34.57$, $\lambda = 0.040566$.

Population 3: [source: Kadilar and Cingi [7]].
Let y =level of apple production (1 unit = 100 tones) and x = number of trees. The summary statistics are: $N = 104$, $n_1 = 40$, $n = 20$, $\bar{Y} = 6.254$, $\bar{X} = 13931.683$, $C_y = 1.866$, $C_x = 1.653$, $\rho = 0.865$, $\beta_{2x} = 17.516$, $\lambda = 0.040385$.

The values of minimum MSEs of the proposed and existing estimators constructed under two-phase sampling for all populations are given in Table 2. For $\hat{\bar{Y}}^*_{GS(1)}$ with $(a = 1, b = \rho)$, $\hat{\bar{Y}}^*_{GS(2)}$ with $(a = 1, b = C_x)$, $\hat{\bar{Y}}^*_{GS(3)}$ with $(a = 1, \beta_{2x})$, $\hat{\bar{Y}}^*_{GS(4)}$ with $(a = \beta_{2x}, b = C_x)$, and $\hat{\bar{Y}}^*_{GS(5)}$ with $(a = C_x, b = \beta_{2x})$.
It is worth mentioning here that for each of the three populations, the proposed estimators $\hat{\bar{Y}}_{Pi}$ and $\hat{\bar{Y}}^*_{Pi}$ $(i = 1, 2, 3)$ perform better than the existing estimators. It is observed that the proposed estimator $\hat{\bar{Y}}_{P3}$ and $\hat{\bar{Y}}^*_{P3}$ are more efficient than their counterparts considered here.

Table 2. Minimum *MSE* values of different estimators in double sampling (partial information)

Estimator	Population-I	Population-II	Population-III
$\hat{\hat{Y}}_D^*$	438750.00	2944860.00	2.95
$\hat{\hat{Y}}_{GS(1)}^*$	155854.00	1757000.00	2.73
$\hat{\hat{Y}}_{GS(2)}^*$	156129.00	1757010.00	2.73
$\hat{\hat{Y}}_{GS(3)}^*$	158855.00	1757220.00	2.73
$\hat{\hat{Y}}_{GS(4)}^*$	155613.00	1757000.00	2.73
$\hat{\hat{Y}}_{GS(5)}^*$	158351.00	1757100.00	2.73
$\hat{\hat{Y}}_{GK}^*$	157838.00	1787510.00	2.67
$\hat{\hat{Y}}_{P1}^*$	155785.00	1659340.00	2.47
$\hat{\hat{Y}}_{P2}^*$	157326.00	1755550.00	2.65
$\hat{\hat{Y}}_{P3}^*$	155271.00	1627170.00	2.41

9. Conclusion

In this paper, we proposed some improved exponential type estimators of finite population mean when complete and partial auxiliary information is available. The proposed estimators perform better than all other competitor estimators considered here. It is to be noted the suggested estimators although biased but are always better than the unbiased difference type estimator of the finite population mean. Based on both theoretical and numerical comparisons, the proposed estimators are more precise than their counterparts. The work can easily be extended to improve the estimation of finite population mean using information on auxiliary attributes, stratified random sampling and other sampling designs. Finally, we recommend the use of $\hat{\hat{Y}}_{P3}$ and $\hat{\hat{Y}}_{P3}^*$ for efficient estimation of the population mean under simple and two-phase sampling schemes, respectively.

Acknowledgements

The authors would like to thank the Editor and the two anonymous referees for their valuable comments.

References

[1] Bahl S., Tuteja R.K. *Ratio and product type exponential estimator.* Journal of Information and Optimization Sciences, **12** (1), 159–164, 1991.

[2] Grover LK., Kaur P. *An improved estimator of the finite population mean in simple random sampling.* Model Assisted Statistics and Applications, **6**(1), 47–55, 2011.

[3] Gupta S., Shabbir J. *On improvement in estimating the population mean in simple random sampling.* Journal of Applied Statistics, **35** (5), 559–556, 2008.

[4] Kadilar, C., Cingi, H. *Ratio estimators in simple random sampling.* Applied Mathematics and Computation, **151** (3), 893–902, 2004.

[5] Kadilar, C., Cingi, H. *New ratio estimators using correlation coefficient.* Interstat **4**, 1–11, 2006a.

[6] Kadilar, C., Cingi, H. *Improvement in estimating the population mean in simple random sampling.* Applied Mathematics Letters, 19, 75–79, 2006b.

[7] Kadilar, C., Cingi, H. *Improvement in variance estimation using auxiliary information.* Hacettepe Journal of Mathematics and Statistics, **35** (1), 111–115, 2006c.

[8] Singh, H.P., Espejo, M.R. *On linear regression and ratio-product estimation of a finite population mean.* The Statistician, **52** (1), 59–67, 2003.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (6) (2014), 1095 – 1106

Jump-diffusion CIR model and its applications in credit risk

Yongfeng Wu *

Abstract

In this paper, the author discusses the distribution of the jump-diffusion CIR model (JCIR) and its applications in credit risk. Applying the piecewise deterministic Markov process theory and martingale theory, we first obtain the closed forms of the Laplace transforms for the distribution of the jump-diffusion CIR model and its integrated process. Based on the obtained Laplace transforms, we derive the pricing of the defaultable zero-coupon bond and the fair premium of a Credit Default Swap (CDS) in a reduced form model of credit risk. Some numerical calculations are also provided.

2000 AMS Classification: 97K60, 91G40.

Keywords: jump-diffusion CIR model; reduced form model of credit risk; Laplace transform; defaultable zero-coupon bond; Credit Default Swap.

Received 06 : 09 : 2013 : Accepted 31 : 01 : 2014 Doi : 10.15672/HJMS.201447455

1. Introduction

As we know, Cox et al. (1985) proposed the classical Cox-Ingersoll-Ross (CIR) process which is defined by an equation of the form

$$\mathrm{d}y_t = \lambda(\eta - y_t)\mathrm{d}t + \theta\sqrt{y_t}\,\mathrm{d}W_t, \tag{1.1}$$

where λ is the rate of mean reversion, η is the long-run level, θ is the volatility coefficient and W_t is a standard Brownian motion.

Compared with the Vasicek process (Vasicek, 1977), although the CIR equation (1.1) does not have a closed-form solution, the CIR process is always positive. If y_t reaches zero, the diffusion term $\mathrm{d}W_t$ disappears and the positive drift term pushes the process in the positive territory. The precise behavior of the CIR process near zero depends on the values of parameters. If $\theta^2 \le 2\lambda\eta$, the positive drift term will always drive the process y_t

*Center for Financial Engineering and School of Mathematical Sciences, Soochow University, Suzhou 215006, China.
College of Mathematics and Computer Science, Tongling University, Tongling 244000, China.
Email: wyfwyf@126.com

away from zero before it will come too close. If $\theta^2 > 2\lambda\eta$, the process y_t will occasionally touch zero and reflect. On the other hand, the CIR process also has the character of mean-reverting. Due to the characters of non-negativity and mean-reverting, the CIR process is better for modelling the interest rate or default intensity in credit risk than the Vasicek model.

Over the recent years, some authors put their attention to the CIR processes and their applications in finance and insurance. We can refer the reader to Chen and Scott (1992), Delbaen (1993), Heston (1993), Berardi (1995), Chou and Lin (2006), Guo (2008), Wu et al. (2009), Ewald and Wang (2010), Trutnau (2011), Song et al. (2012), Bao and Yuan (2013).

In practice, there are primary events such as the governments fiscal and monetary policies, the release of corporate financial reports, some natural disasters and terrorist attacks etc., that will possibly result in some positive jumps in a firm's default intensity process. As time passes, the default intensity process decreases as the firm tries its best to avoid being in bankruptcy after the arrival of a primary event. This decrease will continue until another event occurs, which will result in another positive jump in its intensity processes. In order to describe the appearance of positive jumps in the default intensity process, we consider the jump-diffusion CIR model which has the following structure

$$\mathrm{d}y_t = \lambda(\eta - y_t)\mathrm{d}t + \theta\sqrt{y_t}\,\mathrm{d}W_t + \mathrm{d}J_t, \tag{1.2}$$

where λ, η, θ and W_t are as in the previous model (1.1). We assume that $\lambda > 0$, $\eta \geq 0$ and $\theta \geq 0$. J_t is a compound Poisson process which is given by

$$J_t = \sum_{j=1}^{M_t} X_j, \tag{1.3}$$

where M_t is a Poisson process with frequency ρ and stands for the total number of jumps up to time t. $\{X_j,\, j \geq 1\}$ are the jump sizes and assumed to be independent and identically distributed random variables with distribution function $F(x)\,(x > 0)$.

Clearly (1.1) is a special case of (1.2) for $\rho = 0$. In addition, we can find that $\eta = \theta = 0$ would lead to shot noise processes for y_t. It is well known that shot noise models have been applied to diverse areas such as finance, insurance and electronics. Therefore, from an applied point of view, it is very significant to investigate the wider class of jump-diffusion CIR models.

Let $Y_t = \int_0^t y_u \mathrm{d}u$ be the integrated process of y_t. In this work, we will first study the Laplace transforms of the distributions of the processes y_t and Y_t. Then we will discuss the applications of these Laplace transforms in credit risk.

The rest of this article is organized as follows. In Section 2, we obtain the Laplace transforms for jump-diffusion CIR models and their integrated processes. In section 3, based on the result of the previous section, we derive the pricing of the defaultable zero-coupon bond and the fair premium of a Credit Default Swap (CDS) in a reduced form model of credit risk. Some numerical calculations and concluding remarks are presented in Section 4.

2. The Laplace transforms of the distribution of jump-diffusion CIR model

In this section, by applying the piecewise deterministic Markov process theory and martingale theory, we first derive the joint Laplace transform of the distribution of the vector process $(y_t,\, Y_t)$. Then we obtain the Laplace transforms of the distribution of the jump-diffusion CIR model. The piecewise deterministic Markov process theory was developed by Davis (1984) and has been proved to be a very powerful mathematical tool for examining non-diffusion models. More details on this theory can be found in Davis (1984).

The (infinitesimal) generator $\mathcal{A}$ of the unique solution to SDE (1.1), is given by

$$\mathcal{A}f(y) = \lambda(\eta - y)\,\frac{\partial f}{\partial y} + \frac{1}{2}\,\theta^2\,y\,\frac{\partial^2 f}{\partial y^2}, \tag{2.1}$$

where f is an arbitrary twice continuously differentiable function. We assume that y_t is a jump-diffusion CIR model which is a solution of the SDE (1.2). With the aid of piecewise deterministic Markov process theory and using Theorem 5.5 in Davis (1984), one can see that the (infinitesimal) generator of the process $(Y_t,\, y_t,\, t)$ acting on a function $f(Y,\, y,\, t)$ is given by

$$\begin{aligned}\mathcal{A}f(Y,\, y,\, t) \;&=\; \frac{\partial f}{\partial t} + y\frac{\partial f}{\partial Y} + \lambda(\eta - y)\,\frac{\partial f}{\partial y} + \frac{1}{2}\,\theta^2\,y\,\frac{\partial^2 f}{\partial y^2} \\ &\quad +\rho\left\{\int_0^\infty f(Y,\, y + x,\, t)\mathrm{d}F(x) - f(Y,\, y,\, t)\right\},\end{aligned} \tag{2.2}$$

where $f:\,(0,\,\infty)\times(0,\,\infty)\times\mathrm{R}^+\to(0,\,\infty)$ satisfies:

(1) $f(Y,\, y,\, t)$ is bounded on arbitrary finite time intervals;

(2) $f(Y,\, y,\, t)$ is differentiable with respect to all $t,\, y,\, Y$;

(3)

$$\left|\int_0^\infty f(Y,\, y + x,\, t)\mathrm{d}F(x) - f(Y,\, y,\, t)\right| < \infty.$$

For the sake of simplicity in the presentations throughout the rest of this article, we will use the following functions which are given by

$$\coth x = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{\mathrm{e}^x - \mathrm{e}^{-x}},\;\; \sinh x = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2},\;\; \cosh x = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}.$$

In order to obtain the joint Laplace transform for the distribution of the vector process$(y_t,\, Y_t)$, we first present the following lemma.

Lemma 2.1. Assume that $m,\, k$ are two constants such that $m > 0$ and $k \geq 0$. Then for $0 \leq t < m/\sqrt{\lambda^2 + 2k\theta^2}$,

$$\exp\left\{-A(t)y_t - kY_t + \rho\int_0^t\Big[1 - h(A(v))\Big]\mathrm{d}v + \lambda\eta\int_0^t A(v)\mathrm{d}v\right\} \tag{2.3}$$

is a martingale where

$$h(\xi) = \int_0^\infty \mathrm{e}^{-\xi x}\mathrm{d}F(x), \tag{2.4}$$

$$A(t) = -\frac{\lambda}{\theta^2} - \frac{\sqrt{\lambda^2 + 2k\theta^2}}{\theta^2}\coth\left(\frac{\sqrt{\lambda^2 + 2k\theta^2}t - m}{2}\right). \tag{2.5}$$

Proof. Let

$$f(Y,\, y,\, t) = \exp\big\{-A(t)y - kY + R(t)\big\}. \tag{2.6}$$

From Theorem 7.6.1 in Jacobsen (2006), $f(Y,\, y,\, t)$ has to satisfy $\mathcal{A}f(Y,\, y,\, t) = 0$ for it to be a martingale. Hence by (2.2), it should hold

$$-A^{'}(t)y + R^{'}(t) - ky - \lambda(\eta - y)A(t) + \frac{1}{2}\theta^2 yA^2(t) + \rho\big[h(A(t)) - 1\big] = 0. \tag{2.7}$$

Solving the equation (2.7), we get

$$A(t)=\frac{\big(\sqrt{\lambda^2+2k\theta^2}-\lambda\big)+\big(\sqrt{\lambda^2+2k\theta^2}+\lambda\big)\exp\big(\sqrt{\lambda^2+2k\theta^2}t-m\big)}{\theta^2\big(1-\exp\big(\sqrt{\lambda^2+2k\theta^2}t-m\big)\big)} \tag{2.8}$$

and

$$R(t)=\rho\int_0^t\Big[1-h(A(v))\Big]\mathrm{d}v+\lambda\eta\int_0^t A(v)\mathrm{d}v. \tag{2.9}$$

By the definition of $\coth x$, it holds

$$\coth\Big(\frac{\sqrt{\lambda^2+2k\theta^2}t-m}{2}\Big)=\frac{\exp\big(\sqrt{\lambda^2+2k\theta^2}t-m\big)+1}{\exp\big(\sqrt{\lambda^2+2k\theta^2}t-m\big)-1}. \tag{2.10}$$

Combining (2.8) and (2.10), we get (2.5). Plugging (2.5) and (2.9) into (2.6), (2.3) follows immediately. The proof is completed.

Let $\mathcal{G}_t=\sigma(y_s,\,0\le s\le t)$. Now by means of Lemma 2.1, we give the joint Laplace transform of the distribution of the vector process $(y_t,\,Y_t)$.

Theorem 2.1. Assume that μ, k are two constants such that $\mu\ge 0$, $k\ge 0$. Then the joint Laplace transform of the distribution of $(y_t,\,Y_t)$ is given by

$$\begin{aligned}&E\big\{\mathrm{e}^{-\mu y_t}\mathrm{e}^{-k(Y_t-Y_s)}\big|\,\mathcal{G}_s\big\}\\ =\;&\exp\big(-B_{\mu,k}(s,t)y_s\big)\exp\big(\lambda^2\eta(t-s)/\theta^2\big)\\ &\big(C_{\mu,k}(s,t)\big)^{-\frac{2\lambda\eta}{\theta^2}}\exp\Big(-\rho\int_0^{t-s}\Big[1-h(B_{\mu,k}(0,u))\Big]\mathrm{d}u\Big),\end{aligned} \tag{2.11}$$

where

$$B_{\mu,k}(s,t)=\frac{(2k-\lambda\mu)+\mu\sqrt{\lambda^2+2k\theta^2}\coth\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big)}{(\theta^2\mu+\lambda)+\sqrt{\lambda^2+2k\theta^2}\coth\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big)}, \tag{2.12}$$

$$\begin{aligned}C_{\mu,k}(s,t)\;=\;&\cosh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big)\\ &+(\theta^2\mu+\lambda)(\lambda^2+2k\theta^2)^{-1/2}\sinh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big).\end{aligned} \tag{2.13}$$

Proof. By Lemma 2.1, for an arbitrary fixed time t^* $(0\le s\le t^*<m/\sqrt{\lambda^2+2k\theta^2})$, we have

$$\begin{aligned}&E\Big\{\exp\Big\{-A(t^*)y_{t^*}-kY_{t^*}+\rho\int_0^{t^*}\Big[1-h(A(v))\Big]\mathrm{d}v+\lambda\eta\int_0^{t^*}A(v)\mathrm{d}v\Big\}\Big|\,\mathcal{G}_s\Big\}\\ =\;&\exp\Big\{-A(s)y_s-kY_s+\rho\int_0^{s}\Big[1-h(A(v))\Big]\mathrm{d}v+\lambda\eta\int_0^{s}A(v)\mathrm{d}v\Big\}.\end{aligned} \tag{2.14}$$

Then

$$\begin{aligned}&E\Big\{\exp\Big\{-A(t^*)y_{t^*}-k(Y_{t^*}-Y_s)\Big\}\,\Big|\,\mathcal{G}_s\Big\}\\ =\;&\exp\big(-A(s)y_s\big)\exp\Big(-\rho\int_s^{t^*}\Big[1-h(A(v))\Big]\mathrm{d}v\Big)\exp\Big(-\lambda\eta\int_s^{t^*}A(v)\mathrm{d}v\Big).\end{aligned} \tag{2.15}$$

Set $A(t^*)=\mu\ge 0$. By (2.8), we get

$$m=\sqrt{\lambda^2+2k\theta^2}t^*-\ln\frac{\mu\theta^2+\lambda-\sqrt{\lambda^2+2k\theta^2}}{\mu\theta^2+\lambda+\sqrt{\lambda^2+2k\theta^2}}. \tag{2.16}$$

Clearly $m > \sqrt{\lambda^2 + 2k\theta^2}t^*$, i.e., $t^* < m/\sqrt{\lambda^2 + 2k\theta^2}$. Plugging (2.16) into $A(s)$ and $A(v)$ respectively, by direct computation, we have

$$\begin{aligned} A(s) &= \frac{(\sqrt{\lambda^2+2k\theta^2}-\lambda)+(\sqrt{\lambda^2+2k\theta^2}+\lambda)\exp(\sqrt{\lambda^2+2k\theta^2}s-m)}{\theta^2\big(1-\exp(\sqrt{\lambda^2+2k\theta^2}s-m)\big)} \\ &= \frac{(2k-\lambda\mu+\mu\sqrt{\lambda^2+2k\theta^2})+(\mu\sqrt{\lambda^2+2k\theta^2}+\lambda\mu-2k)\ \exp\big(\sqrt{\lambda^2+2k\theta^2}(s-t^*)\big)}{(\theta^2\mu+\lambda)\big(1-\exp\big(\sqrt{\lambda^2+2k\theta^2}(s-t^*)\big)\big)+\sqrt{\lambda^2+2k\theta^2}\big(1+\exp\big(\sqrt{\lambda^2+2k\theta^2}(s-t^*)\big)\big)} \\ &= \frac{(2k-\lambda\mu)+\mu\sqrt{\lambda^2+2k\theta^2}\coth(\sqrt{\lambda^2+2k\theta^2}\,(t^*-s)/2)}{(\theta^2\mu+\lambda)+\sqrt{\lambda^2+2k\theta^2}\coth(\sqrt{\lambda^2+2k\theta^2}\,(t^*-s)/2)} \\ &= B_{\mu,k}(s,t^*) \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} A(v) &= \frac{(\sqrt{\lambda^2+2k\theta^2}-\lambda)+(\sqrt{\lambda^2+2k\theta^2}+\lambda)\exp(\sqrt{\lambda^2+2k\theta^2}v-m)}{\theta^2\big(1-\exp(\sqrt{\lambda^2+2k\theta^2}v-m)\big)} \\ &= B_{\mu,k}(v,t^*). \end{aligned} \tag{2.18}$$

From (2.15), (2.17) and (2.18), we have

$$\begin{aligned} & E\Big\{\exp\Big\{-\mu y_{t^*}-k(Y_{t^*}-Y_s)\Big\}\,\Big|\,\mathcal{G}_s\Big\} \\ = \;& \exp\big(-B_{\mu,k}(s,t^*)y_s\big)\exp\Big(-\rho\int_s^{t^*}\Big[1-h(B_{\mu,k}(v,t^*))\Big]\mathrm{d}v\Big)\exp\Big(-\lambda\eta\int_s^{t^*}B_{\mu,k}(v,t^*)\mathrm{d}v\Big). \end{aligned} \tag{2.19}$$

Let $u = t^* - v$ in the integral of (2.19), then

$$\begin{aligned} & E\Big\{\exp\Big\{-\mu y_{t^*}-k(Y_{t^*}-Y_s)\Big\}\,\Big|\,\mathcal{G}_s\Big\} \\ = \;& \exp\big(-B_{\mu,k}(s,t^*)y_s\big)\exp\Big(-\rho\int_0^{t^*-s}\Big[1-h(B_{\mu,k}(0,u))\Big]\mathrm{d}u\Big)\exp\Big(-\lambda\eta\int_0^{t^*-s}B_{\mu,k}(0,u)\mathrm{d}u\Big). \end{aligned} \tag{2.20}$$

Since t^* is arbitrary, (2.20) remains true for all $0 \le s \le t < m/\sqrt{\lambda^2 + 2k\theta^2}$, then

$$\begin{aligned} & E\Big\{\exp\Big\{-\mu y_t-k(Y_t-Y_s)\Big\}\,\Big|\,\mathcal{G}_s\Big\} \\ = \;& \exp\big(-B_{\mu,k}(s,t)y_s\big)\exp\Big(-\rho\int_0^{t-s}\Big[1-h(B_{\mu,k}(0,u))\Big]\mathrm{d}u\Big)\exp\Big(-\lambda\eta\int_0^{t-s}B_{\mu,k}(0,u)\mathrm{d}u\Big). \end{aligned} \tag{2.21}$$

By standard integral calculation, then

$$\begin{aligned} & \exp\Big(-\lambda\eta\int_0^{t-s}B_{\mu,k}(0,u)\mathrm{d}u\Big) \\ = \;& \exp\left(-\lambda\eta\int_0^{t-s}\frac{(2k-\lambda\mu)+\mu\sqrt{\lambda^2+2k\theta^2}\coth(\sqrt{\lambda^2+2k\theta^2}\,u/2)}{(\theta^2\mu+\lambda)+\sqrt{\lambda^2+2k\theta^2}\coth(\sqrt{\lambda^2+2k\theta^2}\,u/2)}\mathrm{d}u\right) \\ = \;& \exp\big(\lambda^2\eta(t-s)/\theta^2\big)\Big(\cosh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big) \\ & +(\theta^2\mu+\lambda)(\lambda^2+2k\theta^2)^{-1/2}\sinh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big)\Big)^{-\frac{2\lambda\eta}{\theta^2}}. \end{aligned} \tag{2.22}$$

Plugging (2.22) into (2.21), (2.11) follows immediately. The proof is completed.

Setting $k = 0$ and $\mu = 0$ in (2.11) respectively, we obtain the following corollaries.

Corollary 2.1. Assume that μ, k are two constants such that $\mu \geq 0$, $k \geq 0$. Then the Laplace transforms of the distributions of y_t and Y_t are respectively given by

$$\begin{aligned} & E\{\mathrm{e}^{-\mu y_t} \,|\, \mathcal{G}_s\} \\ = \; & \exp\big(-B_{\mu,0}(s,t)y_s\big)\exp\big(\lambda^2\eta(t-s)/\theta^2\big)\big(C_{\mu,0}(s,t)\big)^{-\frac{2\lambda\eta}{\theta^2}}\exp\Big(-\rho\int_0^{t-s}\Big[1-h(B_{\mu,0}(0,u))\Big]\mathrm{d}u\Big) \end{aligned} \tag{2.23}$$

and

$$\begin{aligned} & E\{\mathrm{e}^{-k(Y_t-Y_s)} \,|\, \mathcal{G}_s\} \\ = \; & \exp\big(-B_{0,k}(s,t)y_s\big)\exp\big(\lambda^2\eta(t-s)/\theta^2\big)\big(C_{0,k}(s,t)\big)^{-\frac{2\lambda\eta}{\theta^2}}\exp\Big(-\rho\int_0^{t-s}\Big[1-h(B_{0,k}(0,u))\Big]\mathrm{d}u\Big). \end{aligned} \tag{2.24}$$

To make later calculation somewhat easier, we assume that jumps size in (1.3) follows exponential distribution, i.e., $F(x) = 1 - \mathrm{e}^{-\alpha x}$ $(x > 0,\ \alpha > 0)$. Then from Corollary 2.1, we get the following result.

Corollary 2.2. Assume that μ, k are two constants such that $\mu \geq 0$, $k \geq 0$ and that $F(x) = 1 - \mathrm{e}^{-\alpha x}$ $(x > 0,\ \alpha > 0)$. Then the Laplace transforms of the distributions of y_t and Y_t are respectively given by

$$\begin{aligned} & E\{\mathrm{e}^{-\mu y_t} \,|\, \mathcal{G}_s\} \\ = \; & \exp\big(-B_{\mu,0}(s,t)y_s + \lambda^2\eta(t-s)/\theta^2\big)\big(C_{\mu,0}(s,t)\big)^{-\frac{2\lambda\eta}{\theta^2}}\left(\frac{2\lambda(\alpha+\mu)\exp(\lambda(t-s))}{\alpha(\theta^2\mu+2\lambda)\exp(\lambda(t-s)) + (2\lambda\mu - \alpha\theta^2\mu)}\right)^{-\frac{2\rho}{2\lambda-\alpha\theta^2}} \end{aligned} \tag{2.25}$$

and

$$\begin{aligned} & E\{\mathrm{e}^{-k(Y_t-Y_s)} \,|\, \mathcal{G}_s\} \\ = \; & \exp\big(-B_{0,k}(s,t)y_s\big)\exp\Big\{M_1(k)\,(t-s) + M_2(k)\ln D_k(s,\,t) - M_3\ln C_{0,k}(s,t)\Big\}, \end{aligned} \tag{2.26}$$

where

$$\begin{aligned} D_k(s,\,t) &= \cosh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big) + \frac{\alpha\lambda+2k}{\alpha\sqrt{\lambda^2+2k\theta^2}}\sinh\big(\sqrt{\lambda^2+2k\theta^2}\,(t-s)/2\big), \\ M_1(k) &= \frac{\lambda^2\eta}{\theta^2} - \frac{2k\rho}{\alpha\big(\sqrt{\lambda^2+2k\theta^2}+\lambda\big)+2k} - \frac{\alpha\rho\sqrt{\lambda^2+2k\theta^2}}{2k+2\alpha\lambda-\alpha^2\theta^2}, \\ M_2(k) &= \frac{2\alpha\rho}{2k+2\alpha\lambda-\alpha^2\theta^2}, \quad M_3 = \frac{2\lambda\eta}{\theta^2}. \end{aligned}$$

3. Applications in credit risk

In this section, based on the results of previous section, we derive the pricing of the defaultable zero-coupon bond and the fair premium of a Credit Default Swap (CDS) in a reduced form model of credit risk. Reduced form models of credit risk were pioneered by Artzner and Delbaen (1995). For the literature on the reduced form model, we can refer to Jarrow and Turnbull (1995), Duffie and Singleton (1999), Bai, Hu, and Ye (2007), Liang and Wang (2012), Su and Wang (2013). In some literature on the reduced form model of credit risk, the default arrival time for the firm is defined as the first jump time of the Cox process. Due to some primary events which will possibly result in some positive jumps in a firm's default intensity process, we employ the jump-diffusion CIR model to describe the firm's default intensity.

We first state the definition of Cox process. Many alternative definitions of a Cox process can be found in the previous literature. We adopted the one used by Brémaud (1981).

Definition 3.1. Let $\{\Omega, \mathcal{F}, \mathrm{P}\}$ be a probability space with information structure given by $\{\mathcal{G}_t,\, t \in [0,T]\}$. Let N_t be a point process adapted to $\{\mathcal{G}_t,\, t \in [0,T]\}$. Let y_t be a nonnegative process adapted to $\{\mathcal{G}_t,\, t \in [0,T]\}$ such that

$$\int_0^t y_u \mathrm{d}u < \infty \quad \text{a. s. (no explosions)}.$$

If for all $u \in \mathbb{R}$ and $0 \le t_1 \le t_2$,

$$E\Big\{\mathrm{e}^{iu(N_{t_2}-N_{t_1})}\,\big|\mathcal{G}_{t_2}\Big\} = \exp\Big\{(\mathrm{e}^{iu}-1)\int_{t_1}^{t_2} y_u \mathrm{d}u\Big\}, \tag{3.1}$$

where $\mathcal{G}_t = \sigma(y_u,\, u \le t)$, then N_t is called a Cox process with intensity y_t.

From this definition, we can consider a Cox process as a two-step randomisation procedure. N_t is a Poisson process conditional to y_t and y_t is used to generate N_t by acting as its intensity. Therefore, a Cox process is also called a doubly stochastic Poisson process.

In the following we assume that y_t is a jump-diffusion CIR model satisfying (1.2) and $y_0 = 0$. Denote $\tau = \inf\{t \ge 0,\, N_t = 1\,|N_0 = 0\}$, where N_t is a Cox process with intensity y_t defined as (1.2). Then from (3.1), we have

$$P\big(N_t - N_s = k\,\big|\mathcal{G}_t\big) = \frac{1}{k!}\Big(\int_s^t y_u \mathrm{d}u\Big)^k \exp\Big(-\int_s^t y_u \mathrm{d}u\Big). \tag{3.2}$$

Let $\mathcal{H}_t := \sigma(\{\tau \le s\},\, s \le t)$, i.e. the σ-algebra generated by τ up to time t and $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$. By the definition of τ and (3.2), the conditional distributions of τ are given by

$$P\big(\tau > t\,\big|\mathcal{G}_t\big) = P\big(N_t - N_0 = 0\,\big|\mathcal{G}_t\big) = \mathrm{e}^{-Y_t}. \tag{3.3}$$

Now we present the survival probability of a firm which has a default intensity process y_t.

Theorem 3.1. Let y_t be a jump-diffusion CIR model satisfying $y_0 = 0$, and N_t be a Cox process with intensity y_t. Then the survival probability is given by

$$\begin{aligned} P\big(\tau > t\big) &= \exp\Big\{\Big(\frac{\lambda^2\eta}{\theta^2} - \rho\Big)t - \frac{2\lambda\eta}{\theta^2}\ln\big[C_{0,1}(0,t)\big] + \rho\int_0^t h(B_{0,1}(0,u))\mathrm{d}u\Big\} \\ &=: \exp\{\Phi(t)\}. \end{aligned} \tag{3.4}$$

Proof. By the definition of the default arrival time τ and (2.24), we have

$$\begin{aligned} & P\big(\tau > t\,\big|\mathcal{G}_s\big) \\ = \;& P\big(N_t - N_s = 0\,\big|\mathcal{G}_s\big) \;=\; E\big\{\mathrm{e}^{-(Y_t - Y_s)}\big|\,\mathcal{G}_s\big\} \\ = \;& \exp\big(-B_{0,1}(s,t)y_s\big)\exp\big(\lambda^2\eta(t-s)/\theta^2\big)\big(C_{0,1}(s,t)\big)^{-\frac{2\lambda\eta}{\theta^2}}\exp\Big(-\rho\int_0^{t-s}\Big[1 - h(B_{0,1}(0,u))\Big]\mathrm{d}u\Big). \end{aligned} \tag{3.5}$$

Note that $y_0 = 0$ and $P(\tau > t) = P(\tau > t \,|\, \mathcal{G}_0)$, we can easily get (3.4). The proof is completed.

Remark 3.1. Taking the derivative in (3.4), we can obtain the default probability density as

$$P(\tau \in \mathrm{d}t) = -\exp\{\Phi(t)\}\partial_t \Phi(t)\,\mathrm{d}t. \tag{3.6}$$

If we assume that $F(x) = 1 - \mathrm{e}^{-\alpha x}$ ($x > 0$, $\alpha > 0$), we can get the following corollary.

Corollary 3.1. Let y_t be a jump-diffusion CIR model satisfying $y_0 = 0$, and N_t be a Cox process with intensity y_t. Assume that $F(x) = 1 - \mathrm{e}^{-\alpha x}$ ($x > 0$, $\alpha > 0$). Then the survival probability is given by

$$\begin{aligned} P(\tau > t) &= \exp\Big\{M_1(1)\,t + M_2(1)\ln D_1(0,\,t) - M_3 \ln C_{0,1}(0,t)\Big\} \\ &=: \exp\{\Psi(t)\}. \end{aligned} \tag{3.7}$$

By some similar arguments as in the proof of Theorem 3.1, we can prove this corollary. Here we omit the details.

Next we will derive the pricing of the defaultable zero-coupon bond and the fair premium of a CDS. In recent years, the rapid expansion of market for credit derivatives has led to a growing interest in investigation of the pricing of the defaultable zero-coupon bond and the fair premium of CDS. A CDS is in fact a contract agreement between protection buyer and seller. Assume that firm A issues a defaultable zero-coupon bond and investor B holds the bond. Then B faces the credit risk arising from default of firm A. In order to protect from this credit risk, B buys a CDS contract which requires B to pay periodic premium to party C (CDS protection seller). In exchange, C will compensate B for his loss in the event of default of the bond.

The following definition of the price process of CDS can be found in Crépey et al. (2009).

Definition 3.2. The model price process of a CDS is given by $P_t = E\{p_T(t)\}$, where $p_T(t)$ corresponds to the CDS cumulative discounted cash flows on the time interval $(t,\,T]$ and satisfies

$$\beta(t)p_T(t) = (1-R)\beta(\tau)I_{(t<\tau<T)} - \kappa\int_t^{\tau\wedge T}\beta(v)\mathrm{d}v. \tag{3.8}$$

In equation (3.8), τ is the default arrival time of firm A, $\beta(t) = \mathrm{e}^{-\int_0^t r_u \mathrm{d}u}$ is the discount factor. Here we assume that the market interest rate r_t is a deterministic function of the time and that the recovery rate is R. (3.8) describes the change trend of cash flow for investor B. The first term on the right-hand side of (3.8) corresponds to the present value of the investor B's loss $(1-R)$ resulted by the default of firm A. The second term on the right-hand side of (3.8) corresponds to the present value of the premiums which B pays to C.

We first state the pricing of the defaultable zero-coupon bond and the fair premium of CDS based on the conclusion of Theorem 3.1.

Theorem 3.2. Let $B(0,\,T)$ be the present value of the defaultable zero-coupon bond at time 0 paying 1 at time T and κ be the fair premium of CDS. Then the following statements hold:

(1) The formula for calculating the value of $B(0,\,T)$ is given by

$$B(0,\,T) = \mathrm{e}^{-\int_0^T r_u \mathrm{d}u}\exp\{\Phi(T)\} - R\int_0^T \mathrm{e}^{-\int_0^t r_u \mathrm{d}u}\exp\{\Phi(t)\}\,\partial_t\Phi(t)\,\mathrm{d}t. \tag{3.9}$$

(2) The pricing of CDS at time t is given by

$$P_t = I_{(\tau>t)} E\left\{\int_t^T \big((1-R)y_v - \kappa\big)\mathrm{e}^{-\int_t^v (r_u+y_u)\mathrm{d}u}\mathrm{d}v\Big|\mathcal{G}_t\right\}. \tag{3.10}$$

(3) The fair premium of CDS is given by

$$\kappa = \frac{-(1-R)\int_0^T \mathrm{e}^{-\int_0^v r_u \mathrm{d}u}\exp\{\Phi(v)\}\,\partial_t\Phi(v)\,\mathrm{d}v}{\int_0^T \mathrm{e}^{-\int_0^v r_u \mathrm{d}u}\exp\{\Phi(v)\}\mathrm{d}v}. \tag{3.11}$$

Proof. (1) By (3.4), (3.6) and the definition of the defaultable zero-coupon bond, we obtain immediately

$$\begin{aligned} B(0,T) &= \mathrm{e}^{-\int_0^T r_u \mathrm{d}u} P(\tau>T) + R\int_0^T \mathrm{e}^{-\int_0^t r_u \mathrm{d}u}\mathrm{d}P(\tau\le t) \\ &= \mathrm{e}^{-\int_0^T r_u \mathrm{d}u}\exp\{\Phi(T)\} - R\int_0^T \mathrm{e}^{-\int_0^t r_u \mathrm{d}u}\exp\{\Phi(t)\}\,\partial_t\Phi(t)\,\mathrm{d}t. \end{aligned}$$

(2) By Definition 3.2, we get

$$\begin{aligned} P_t &= E\{p_T(t)\big|\mathcal{F}_t\} \\ &= E\left\{(1-R)\beta(\tau)\beta^{-1}(t)I_{(t<\tau<T)} - \kappa\int_t^T \beta(v)\beta^{-1}(t)I_{(\tau>v)}\mathrm{d}v\Big|\mathcal{F}_t\right\} \\ &= (1-R)E\left\{\mathrm{e}^{-\int_t^\tau r_u \mathrm{d}u}I_{(t<\tau<T)}\Big|\mathcal{F}_t\right\} - \kappa\int_t^T \mathrm{e}^{-\int_t^v r_u \mathrm{d}u}E\{I_{(\tau>v)}\big|\mathcal{F}_t\}\mathrm{d}v \\ &=: P_t' - P_t''. \end{aligned} \tag{3.12}$$

By Theorem 9.23 in McNeil et al. (2005), we have

$$P_t' = I_{(\tau>t)}(1-R)E\left\{\int_t^T y_v \mathrm{e}^{-\int_t^v (r_u+y_u)\mathrm{d}u}\mathrm{d}v\Big|\mathcal{G}_t\right\}. \tag{3.13}$$

By Lemma 7.4.1.1 in Jeanblanc et al. (2009) (taking $X\equiv 1$ and $T=v$) and (3.3), for $v\ge t$, we have

$$\begin{aligned} E\{I_{(\tau>v)}\big|\mathcal{F}_t\} &= I_{(\tau>t)}\frac{E\{I_{(\tau>v)}\big|\mathcal{G}_t\}}{E\{I_{(\tau>t)}\big|\mathcal{G}_t\}} = I_{(\tau>t)}\frac{E\{I_{(\tau>v)}\big|\mathcal{G}_t\}}{P\big(\tau>t\big|\mathcal{G}_t\big)} \\ &= I_{(\tau>t)}\mathrm{e}^{\int_0^t y_u \mathrm{d}u}E\{E\{I_{(\tau>v)}\big|\mathcal{G}_v\}\big|\mathcal{G}_t\} \\ &= I_{(\tau>t)}\mathrm{e}^{\int_0^t y_u \mathrm{d}u}E\{P\big(\tau>v\big|\mathcal{G}_v\big)\big|\mathcal{G}_t\} \\ &= I_{(\tau>t)}\mathrm{e}^{\int_0^t y_u \mathrm{d}u}E\{\mathrm{e}^{-\int_0^v y_u \mathrm{d}u}\big|\mathcal{G}_t\} \\ &= I_{(\tau>t)}E\{\mathrm{e}^{-\int_t^v y_u \mathrm{d}u}\big|\mathcal{G}_t\}. \end{aligned} \tag{3.14}$$

Plugging (3.14) into P_t'', we get

$$P_t'' = I_{(\tau>t)}\,\kappa\,E\left\{\int_t^T \mathrm{e}^{-\int_t^v (r_u+y_u)\mathrm{d}u}\mathrm{d}v\Big|\mathcal{G}_t\right\}. \tag{3.15}$$

Then (3.10) follows by (3.12), (3.13) and (3.15).

(3) Note that

$$P(\tau>v) = P(\tau>v|\mathcal{G}_0) = E\left\{\mathrm{e}^{-\int_0^v y_u \mathrm{d}u}\Big|\mathcal{G}_0\right\}$$

and

$$-\frac{\partial}{\partial v}P(\tau>v) = E\left\{y_v\mathrm{e}^{-\int_0^v y_u \mathrm{d}u}\Big|\mathcal{G}_0\right\}.$$

Then (3.11) follows by setting $P_0 = 0$ in (3.10). The proof is completed.

For the sake of the numerical calculations in next section, we present the following corollary based on the conclusion of Corollary 3.1.

Corollary 3.2. Let $B(0, T)$ be the present value of the defaultable zero-coupon bond at time 0 paying 1 at time T and κ be the fair premium of CDS. Then the following statements hold:

(1) The formula for calculating the value of $B(0, T)$ is given by

$$B(0,\,T) = \mathrm{e}^{-\int_0^T r_u \mathrm{d}u} \exp\{\Psi(T)\} - R\int_0^T \mathrm{e}^{-\int_0^t r_u \mathrm{d}u} \exp\{\Psi(t)\}\, \partial_t \Psi(t)\, \mathrm{d}t. \tag{3.16}$$

(2) The fair premium of CDS is given by

$$\kappa = \frac{-(1-R)\int_0^T \mathrm{e}^{-\int_0^v r_u \mathrm{d}u} \exp\{\Psi(v)\}\, \partial_t \Psi(v)\, \mathrm{d}v}{\int_0^T \mathrm{e}^{-\int_0^v r_u \mathrm{d}u} \exp\{\Psi(v)\} \mathrm{d}v}. \tag{3.17}$$

4. Numerical results and conclusions

In this section, using the conclusions of Corollary 3.2, let us illustrate the price calculations of the defaultable zero-coupon bond and the fair premium of CDS. We also analyse the dynamic relationships between $B(0, T)$, κ and the maturity date T respectively.

Example 4.1. The parameter values used to calculate the pricing of the defaultable zero-coupon bond and the fair premium of CDS are

$$\lambda = 0.1,\ \eta = 0,\ \theta = 0.2,\ \alpha = 15,\ \rho = 1,\ R = 0.4,\ r_t = 0.05.$$

Note that Corollary 3.2 is based on Corollary 3.1. The expressions of $\exp\{\Psi(T)\}$, $\exp\{\Psi(t)\}$, $\partial_t\Psi(t)$, $\exp\{\Psi(v)\}$ and $\partial_t\Psi(v)$ in Corollary 3.2 can be obtained from the conclusion of Corollary 3.1. Therefore, after substituting the above parameter values into $\exp\{\Psi(T)\}$, $\exp\{\Psi(t)\}$, $\partial_t\Psi(t)$, $\exp\{\Psi(v)\}$ and $\partial_t\Psi(v)$, one can obtain the numbers showed in the following Table 4.1 and 4.2 by means of the conclusions of Corollary 3.2 and MATLAB software.

Table 4.1. The dynamic relationship between $B(0, T)$ and T

T	2	4	6	8	10
$B(0, T)$	0.7613	0.5861	0.4733	0.4058	0.3669

Table 4.2. The dynamic relationship between κ and T

T	2	4	6	8	10
κ	0.0950	0.1040	0.1093	0.1123	0.1140

Table 4.1 shows that the dynamic relationship between the pricing of the defaultable zero-coupon bond and the maturity date T. Table 4.2 shows that the dynamic relationship between the fair premium of CDS and the maturity date T.

We can find from Table 4.1 that the price of the defaultable zero-coupon bond is monotonically decreasing function respect to the maturity date T. However, it is indicated from Table 4.2 that the price of the fair premium of CDS is monotonically increasing function respect to the maturity date T. The reason for this monotonically increasing trend is that the ruin probability of firm A increases with prolonged maturity date T.

In this paper, for the sake of simplifying calculation, we assume that the jump sizes are exponentially distributed. It is of interest and challenging to employ other heavy-tailed distributions for the jump sizes, such as Pareto distribution, Gumbel distribution and Fréchet distribution. However, since it is unlikely for us to obtain explicit expressions for the joint Laplace transform of the distribution of the vector process (y_t, Y_t), numerical methods need to be used to calculate the price of the defaultable zero-coupon and the fair premium of CDS.

Acknowledgments

The author is very grateful to the referees and the Editor Prof. Dr. Cem Kadilar for the careful reading, valuable suggestions and comments, which enabled him to greatly improve the paper. This work was supported by the National Natural Science Foundation of China (11371274), the Humanities and Social Sciences Foundation for the Youth Scholars of Ministry of Education of China (12YJCZH217) and the Natural Science Foundation of Anhui Province (1308085MA03).

References

[1] Artzner, P., Delbaen, F., 1995. Default risk insurance and incomplete markets. Math. Finance, 5: 187-195.

[2] Bai, Y., Hu, X., Ye, Z., 2007. A model for dependent default with hyperbolic attenuation effect and valuation of credit default swap. Appl. Math. Mechanics Engl. Ser., 28: 1643-1649.

[3] Bao, J. H., Yuan, C. G., 2013. Long-term behavior of stochastic interest rate models with jumps and memory. Insur. Math. Econ., 53: 266-272.

[4] Berardi, A., 1995. Estimating the Cox, ingersoll and Ross model of the term structure: a multivariate approach. Ric. Econ., 49: 51-74.

[5] Brémaud, P., 1981. Point Processes and Queues. Martingale Dynamics, Springer Series in Statistics, Springer, New York, NY, USA.

[6] Chen, R., Scott., L., 1992. Pricing interest rate options in a two-factor Cox CIngersoll CRoss model of the term structure. Rev. Fin. Studies, 5: 613-636.

[7] Chou, C. S., Lin, H. J., 2006. Some properties of CIR processes. Stochastic Anal. Appl., 24: 901-912.

[8] Cox, J., Ingersoll, J., Ross, S., 1985. A theory of term structure of interest rates. Econometrica, 53: 385 C407.

[9] Crépey, S., Jeanblanc, M., Zargari, B., 2009. Counterparty risk on a CDS in a Markov chain copula model with joint defaults. In Recent Advances In Financial Engineering 2009 (Kijima, M., Hara, C., Tanaka, K. and Muromachi, Y., Eds.), World Scientific Publishing Company, Singapore, 2010, 91-126.

[10] Davis, M. H. A., 1984. Piecewise deterministic Markov processes: A general class of non-diffusion stochastic models. J. R. Statist. Soc. Ser. B, 46: 353-388.

[11] Delbaen, F., 1993. Consols in the CIR model. Math. Finance, 3: 125-134.

[12] Duffie, D., Singleton, K. J., 1999. Modeling term structures of defaultable bonds. Rev. Fin. Studies, 12: 687-720.

[13] Ewald, C. O., Wang, W. K., 2010. Irreversible investment with Cox-Ingersoll-Ross type mean reversion. Math. Soc. Sci., 59: 314-318.

[14] Guo, Z. J., 2008. A note on the CIR process and the existence of equivalent martingale measures. Statist. Probab. Lett., 78: 481-487.

[15] Heston, S. L., 1993. A closed form solution for options with stochastic volatility with applications to bond and currency options. Rev. Fin. Studies, 6:327-343.

[16] Jacobsen, M., 2006. Point Process Theory and Applications. Marked Point and Piecewise Deterministic Processes. Birkhauser, Boston.

[17] Jarrow, R. A., Turnbull, S. M., 1995. Pricing derivatives on financial securities subject to credit risk. J. Finance, 50: 53-86.

[18] Jeanblanc, M., Yor, M., Chesney, M., 2009. Mathematical Methods for Financial Markets. Springer-Verlag London Limited 2009, New York.

[19] Liang, X., Wang, G. J., 2012. On a reduced form credit risk model with common shock and regime switching. Insur. Math. Econ., 51: 567-575.

[20] McNeil, A. J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, New Jersey.

[21] Song, N., Siu, T. K., Fard, F. A., Ching, W. K., Fung, E. S., 2012. Risk measures and behaviors for bonds under stochastic interest rate models. Math. Comput. Modelling, 56: 204-217.

[22] Su, X. N., Wang, W. S., 2013. Pricing options with credit risk in a reduced form model. J. Korean Stat. Soc., doi:10.1016/j.jkss.2012.01.006.

[23] Trutnau, G., 2011. Pathwise uniqueness of the squared Bessel and CIR processes with skew reflection on a deterministic time dependent curve. Stochastic Processes Appl., 121: 1845-1863.

[24] Vasicek, O., 1977. An equilibrium characterization of the term structure, Journal of financial economics, 5: 177-188.

[25] Wu, F. K., Mao, X. R., Chen, K., 2009. The Cox-Ingersoll-Ross model with delay and strong convergence of its Euler-Maruyama approximate solutions. Appl. Numer. Math., 59: 2641-2658.

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.

Copyright : No manuscript should be submitted which has previously been published, or which has been simultaneously submitted for publication elsewhere. The copyright in a published article rests solely with the Faculty of Science of Hacettepe University, and the paper may not be reproduced in whole in part by any means whatsoever without prior written permission.

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.