HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 43 Issue 1
2014

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 43 Issue 1
February 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 1 (2014)
ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2014 SAYI : 43 - 1 AY : Şubat
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 1 (2014)
ISSN 1303 – 5010

Published by Hacettepe University
Faculty of Science

CONTENTS

Mathematics

Statistics

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 1 – 14

GENERATING FUNCTIONS FOR THE BERNSTEIN TYPE POLYNOMIALS: A NEW APPROACH TO DERIVING IDENTITIES AND APPLICATIONS FOR THE POLYNOMIALS

Yılmaz Şimşek*

Received 14 : 02 : 2012 : Accepted 19 : 09 : 2012

Abstract

The main aim of this paper is to construct generating functions for the Bernstein type polynomials. Using these generating functions, various functional equations and differential equations can be derived. New proofs both for a recursive definition of the Bernstein type basis functions and for derivatives of the nth degree Bernstein type polynomials can be given using these equations. This paper presents a novel method for deriving various new identities and properties for the Bernstein type basis functions by using not only these generating functions but also these equations. By applying the Fourier transform and the Laplace transform to the generating functions, we derive interesting series representations for the Bernstein type basis functions. Furthermore, we discuss analytic representations for the generalized Bernstein polynomials through the binomial or Newton distribution and Poisson distribution with mean and variance. By using the mean and the variance, we generalize Szasz-Mirakjan type basis functions.

Keywords: Bernstein polynomials; Generating function; Szasz-Mirakjan basis functions; Bezier curves; Binomial distribution; Poisson distribution; Fourier transform; Laplace transform; Functional equation .

2000 AMS Classification: 14F10, 12D10, 26C05, 26C10, 30B40, 30C15, 42A38, 44A10.

*Department of Mathematics, Faculty of Science University of Akdeniz TR-07058 Antalya, Turkey, Email: ysimsek@akdeniz.edu.tr

1. Introduction and main definition

In the literature in *Bezier Curves and Surfaces*, one can find systematic and extensive investigations not only of the classical Bernstein polynomials and Bezier curves, but also of their various generalizations and q-extensions. According to Goldman [7], freeform curves and surfaces are smooth shapes often describing man-made objects. The hood of a car, the hull of a ship, and the fuselage of an airplane are all examples of freeform shapes which differ from the classical surfaces. The classical surfaces are easy to describe with a few parameters. But the hood of a car or the hull of a ship is not easy to describe with a few parameters. Thus recently many scientists and engineers have developed mathematical techniques for describing freeform curves and surfaces. It is also well-known that scientists and engineers use freeform curves and surfaces to interpolate data and to approximate shape. The Bezier curves, which are polynomials curves, have many practical applications, ranging from the design of new fonts to the creation of mechanical components and assemblies for large scale industrial design and manufacture. By using the Bernstein polynomials, one can easily find an explicit polynomial representation for Bezier curves. Therefore, the Bernstein polynomials have many applications in theory of freeform curves and surfaces, in approximations of functions, in statistics, in numerical analysis, in p-adic analysis and in the solution of differential equations. It is also well-known that in Computer Aided Geometric Design polynomials are often expressed in terms of the Bernstein basis functions. The goal of this paper is to develop some of properties underlying the Bernstein polynomials using their novel generating functions.

Many of the known identities for the Bernstein basis functions are currently derived in an *ad hoc* fashion, using either the binomial theorem, the binomial distribution, tricky algebraic manipulations or blossoming. The aim of this paper is to derive functional equations and differential equations using novel generating functions for the Bernstein polynomials. By using these equations, we provide a *new* approach to derive both for standard identities and for new identities for the Bernstein type basis functions.

The organization of the paper is as follows:

In Section 2; We define generating functions for the Bernstein type basis functions. We find many functional equations and differential equations of this novel generating function. Using these equations, many properties of the Bernstein type basis functions can be determined. For instance, we give sum and alternating sum of the Bernstein type basis functions, some well-known properties of the Bernstein type basis functions, subdivision property, a recursive definition of the Bernstein type basis functions, derivatives of the nth degree Bernstein basis functions. We also prove many other properties of the Bernstein basis functions via functional equations. In Section 3; we give some application of the Fourier transform and the Laplace transform to the generating functions for the Bernstein type basis functions. We derive series representations for the Bernstein type basis functions. In Section 4; by using novel generating functions and their functional equation, we give some new identities related to the Bernstein type basis function. In Section 5; we give relations between the Bernstein basis functions, the binomial distribution and the Poisson distribution. Using the Poisson distribution, we give generating functions for the Szasz-Mirakjan type basis functions. By using Abel and Li's method [1], and applying our generating functions to Proposition 5.1, we derive identities which give pointwise orthogonality relations for the Bernstein polynomials and the Szasz-Mirakjan type basis functions.

2. New approach to deriving new proofs of the identities and properties for the Bernstein type basis functions

In this section, we provide fundamental properties of the Bernstein basis functions and their generating functions. We introduce some functional equations and differential equations of the novel generating functions for the Bernstein basis functions. We also give new proofs of some well known properties of the Bernstein basis functions by using functional equations and differential equations.

2.1. Generating Functions. Recently the Bernstein polynomials have been defined and studied in many different ways, for example, by q-series, by complex functions, by p-adic Volkenborn integrals and many algorithms. Here, by using entire function, related to nonnegative real parameters, we construct generating functions for the Bernstein type basis functions.

The Bernstein type basis functions $\mathbb{Y}_k^n(x;a,b,m)$ are defined as follows:

2.1. Definition. Let a and b be nonnegative real parameters with $a \neq b$. Let m be a positive integer and let $x \in [a,b]$. Let n be non-negative integer. The Bernstein type basis functions $\mathbb{Y}_k^n(x;a,b,m)$ can be defined by

$$\mathbb{Y}_k^n(x;a,b,m) = \binom{n}{k} \frac{(x-a)^k (b-x)^{n-k}}{(b-a)^m}, \tag{2.1}$$

where

$$k = 0, 1, \ldots, n,$$

and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Remark 1. In the special case when $m = n$, Definition 2.1 immediately yields the corresponding well known results concerning the Bernstein basis functions $B_k^n(x,a,b)$ that appears, for example, in Goldman [7, p. 384, Eq.(24.6)] and cf. [3]:

$$\mathbb{Y}_k^n(x;a,b,n) = B_k^n(x;a,b) = \binom{n}{k} \left(\frac{x-a}{b-a}\right)^k \left(\frac{b-x}{b-a}\right)^{n-k},$$

where $k = 0,\ 1, \cdots,\ n$ and $x \in [a,b]$ (cf., see also [5]). One can easily see that

$$B_k^n(x) = \binom{n}{k} x^k (1-x)^{n-k}, \tag{2.2}$$

where $k = 0, 1, \cdots, n$ and $x \in [0,1]$ cf. [1]-[19]. In [7], Goldman gives many properties of the Bernstein polynomials $B_k^n(x,a,b)$. The functions $B_0^n(x,a,b), \cdots, B_n^n(x,a,b)$ are called the Bernstein basis functions. Goldman [7, Chapter 26], shows that the Bernstein basis functions form a basis for the polynomials of degree n.

Generating functions for the Bernstein type basis functions can be defined as follows:

2.2. Definition. Let a and b be nonnegative real parameters with $a \neq b$. Let $t \in \mathbb{C}$. Let m be a positive integer and let $x \in [a,b]$. The Bernstein type basis functions can be defined by means of the following generating function

$$f_{\mathbb{Y},k}(x,t;a,b,m) := \sum_{n=0}^{\infty} \mathbb{Y}_k^n(x;a,b,m) \frac{t^n}{n!}, \tag{2.3}$$

where $k = 0, 1, \ldots, n$.

We construct novel generating functions for the Bernstein type basis functions explicitly by the following theorem:

2.3. Theorem. *Let a and b be nonnegative real parameters with $a \neq b$. Let $t \in \mathbb{C}$. Let m be a positive integer and let $x \in [a,b]$. Then we have*

$$f_{\mathbb{Y},k}(x,t;a,b,m) = \frac{t^k (x-a)^k e^{(b-x)t}}{(b-a)^m k!}. \tag{2.4}$$

Proof. By using (2.1) and (2.3), we have

$$\sum_{n=0}^{\infty} \mathbb{Y}_k^n(x;a,b,m)\frac{t^n}{n!} = \sum_{n=0}^{\infty} \left(\begin{array}{c} n \\ k \end{array} \right) \frac{(x-a)^k (b-x)^{n-k}}{(b-a)^m} \frac{t^n}{n!}.$$

From this equation, we obtain

$$\sum_{n=0}^{\infty} \mathbb{Y}_k^n(x;a,b,m)\frac{t^n}{n!} = \frac{(x-a)^k t^k}{k!(b-a)^m} \sum_{n=k}^{\infty} \frac{(b-x)^{n-k} t^{n-k}}{(n-k)!}.$$

The series on the right hand side is the Taylor series for $e^{(b-x)t}$. Thus we are led to the formula (2.4) asserted by Theorem 2.3. $\square$

Alternative form of the generating functions for the Bernstein type basis functions can be given as follows

$$\frac{t^k (x-a)^k}{(b-a)^m k!} = f_{\mathbb{Y},k}(x,t;a,b,m)e^{(x-b)t}. \tag{2.5}$$

Substituting $m = n$ in (2.1), we now give *another well-known generating function* for the Bernstein basis functions:

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} B_k^n(x;a,b)t^k \right) \frac{z^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) t^k \left(\frac{x-a}{b-a} \right)^k \left(\frac{b-x}{b-a} \right)^{n-k} \right) \frac{z^n}{n!}.$$

By using the Cauchy product in the above equation, we have

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} B_k^n(x;a,b)t^k \right) \frac{z^n}{n!} = \sum_{n=0}^{\infty} \left(t\frac{x-a}{b-a} \right)^n \frac{z^n}{n!} \sum_{n=0}^{\infty} \left(\frac{b-x}{b-a} \right)^n \frac{z^n}{n!}.$$

From this equation, we find that

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} B_k^n(x;a,b)t^k \right) \frac{z^n}{n!} = e^{z\left(\frac{b-x}{b-a} + t\frac{x-a}{b-a} \right)}.$$

After some elementary calculations in the above relation, we arrive at the following *generating function* for the Bernstein basis functions:

$$\sum_{k=0}^{n} B_k^n(x;a,b)t^k = \left(\frac{b-x}{b-a} + t\frac{x-a}{b-a} \right)^n. \tag{2.6}$$

Remark 2. If we set $a = 0$ and $b = 1$ in (2.6), then we have

$$\sum_{k=0}^{n} B_k^n(x)t^k = ((1-x) + tx)^n. \tag{2.7}$$

This generating function is given by Goldman [9]-[8, Chapter 5, pp. 299-306]. Goldman [9]-[8, Chapter 5, pp. 299-306] also constructs the following generating function for the Bernstein basis functions:

$$\sum_{k=0}^{n} B_k^n(x)e^{ky} = ((1-x) + te^y)^n.$$

Remark 3. If we set $a = 0$ and $b = 1$ in (2.4), we obtain a result given by Simsek [18], Simsek et al. [19] and Acikgoz et al. [2]:

$$\frac{(xt)^k}{k!}e^{(1-x)t} = \sum_{n=0}^{\infty} B_k^n(x)\frac{t^n}{n!},$$

so that, obviously;

$$\mathbb{Y}_k^n(x;0,1,n) = B_k^n(x),$$

where $B_k^n(x)$ denote the Bernstein basis functions.

2.2. Bernstein type polynomials. A Bernstein type polynomial $\mathcal{P}(x,a,b,m)$ is a polynomial represented in the Bernstein basis functions:

$$\mathcal{P}(x,a,b,m) = \sum_{k=0}^{n} c_k^n \mathbb{Y}_k^n(x;a,b,m). \tag{2.8}$$

Remark 4. If we set $a = 0$, $b = 1$ and $m = n$ in (2.8), then we have

$$P(x) = \sum_{k=0}^{n} c_k^n B_k^n(x)$$

(cf. [4]).

2.3. Bezier type curve. We define the *Bezier type curve* $B(x,a,b)$ with control points

$$P_0, \ldots, P_n$$

as follows:

$$B(x,a,b;m) = \sum_{k=0}^{n} P_k \mathbb{Y}_k^n(x,a,b,m). \tag{2.9}$$

Remark 5. In the special case when $m = n$, Equation (2.9) yields the corresponding well known results concerning the Bezier curve $B(x,a,b)$ with control points $P_0, \ldots, P_n$ defined as follows (cf. [7]):

$$B(x,a,b) = \sum_{k=0}^{n} P_k B_k^n(x,a,b).$$

2.4. Some well-known properties of the Bernstein type basis functions. Below are some well-known properties of the Bernstein type basis functions:

Non-negative property:

$$\mathbb{Y}_k^n(x;a,b,m) \geq 0, \text{ for } 0 \leq a \leq x \leq b. \tag{2.10}$$

Symmetry property:

$$\mathbb{Y}_k^n(x;a,b,m) = \mathbb{Y}_{n-k}^n(b+a-x;a,b,m). \tag{2.11}$$

Corner values:

$$\mathbb{Y}_k^n(a;a,b,n) = \begin{cases} 0 & \text{if } k \neq 0, \\ 1 & \text{if } k = 0, \end{cases} \tag{2.12}$$

and

$$\mathbb{Y}_k^n(b;a,b,n) = \begin{cases} 0 & \text{if } k \neq n, \\ 1 & \text{if } k = n. \end{cases} \tag{2.13}$$

Remark 6. If we set $a = 0$, $b = 1$ and $m = n$, then (2.10)-(2.13) reduce to Goldman's results [9]-[8, Chapter 5, pp. 299-306]. In [9] and [8, Chapter 5, pp. 299-306], Goldman also gives many identities and properties for the univariate and bivariate Bernstein basis functions, for example boundary values, maximum values, partitions of unity, representation of monomials, representation in terms of monomials, conversion to monomial form, linear independence, Descartes' law of sign, discrete convolution, unimodality, subdivision, directional derivatives, integrals, Marsden identities, De Boor-Fix formulas, and the other properties.

In the next section, by using the same method in [18], we give some functional equations. By using this equations, we find sum and alternating sum of the Bernstein basis functions.

2.5. Sum of the Bernstein type basis functions. Using the same method proposed in [18], we get the following functional equation:

$$\sum_{k=0}^{\infty} f_{\mathbb{Y},k}(x,t;a,b,m) = \frac{e^{(b-a)t}}{(b-a)^m}.$$

From the above equation, we have the sum of the Bernstein basis functions:

$$\sum_{k=0}^{n} \mathbb{Y}_k^n(b;a,b,m) = (b-a)^{n-m}.$$

Observe that by substituting $n=m$ into the above equation, we obtain sum of the Bernstein basis function as follows:

$$\sum_{k=0}^{n} B_k^n(b;a,b) = 1.$$

2.6. Alternating sum of the Bernstein type basis functions. Using the same method proposed in [18], we get the following functional equation:

$$\sum_{k=0}^{\infty} (-1)^k f_{\mathbb{Y},k}(x,t;a,b,m) = \frac{e^{(b-a-2x)t}}{(b-a)^m}. \tag{2.14}$$

By using this equation, we easily arrive at the following alternating sum for the Bernstein type basis functions:

2.4. Theorem.

$$\sum_{k=0}^{n} (-1)^k \mathbb{Y}_k^n(b;a,b,m) = \frac{(b-a-2x)^n}{(b-a)^m}. \tag{2.15}$$

Remark 7. Substituting $m=n$ in (2.1), we get

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} (-1)^k B_k^n(x;a,b,n) \right) \frac{t^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{\left(\frac{a-x}{b-a}\right)^k \left(\frac{b-x}{b-a}\right)^{n-k}}{k!(n-k)!} \right) t^n.$$

By using the Cauchy product in the above equation, we have

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} (-1)^k B_k^n(x;a,b) \right) \frac{t^n}{n!} = e^{\left(\frac{a+b-2x}{b-a}\right)t}.$$

From this relation, we also arrive at the following alternating sum for the Bernstein basis functions:

$$\sum_{k=0}^{n} (-1)^k B_k^n(x;a,b) = \left(\frac{a+b-2x}{b-a} \right)^n.$$

2.7. Differentiating the generating function. Here, we give higher order derivatives of the Bernstein type basis functions by differentiating the generating function in (2.4) with respect to x. Using Leibnitz's formula for the lth derivative, with respect to x, of the product $f_{\mathbb{Y},k}(x,t;a,b,m)$ of two functions

$$g(t,x;a,b)=\frac{t^{k}\left(x-a\right)^{k}}{(b-a)^{m}k!}\quad(a\neq b)$$

and

$$h(t,x;b)=e^{(b-x)t},$$

we obtain the following *higher order partial derivative equation*:

$$\frac{\partial^{l}f_{\mathbb{Y},k}(x,t;a,b,m)}{\partial x^{l}}=\sum_{j=0}^{l}\left(\begin{array}{c}l\\j\end{array}\right)\left(\frac{\partial^{j}g(t,x;a,b)}{\partial x^{j}}\right)\left(\frac{\partial^{l-j}h(t,x;b)}{\partial x^{l-j}}\right).\tag{2.16}$$

By using induction on l, Equation (2.16) is easily obtained.

2.5. Theorem. *Let l be a non-negative integer. Then*

$$\frac{\partial^{l}f_{\mathbb{Y},k}(x,t;a,b,m)}{\partial x^{l}}=\sum_{j=0}^{l}\left(\begin{array}{c}l\\j\end{array}\right)(-1)^{l-j}\frac{t^{l}}{(b-a)^{j}}f_{\mathbb{Y},k-j}(x,t;a,b,m-j).$$

Proof. By using (2.16), we easily arrive at the desired result. □

By using Theorem 2.5, we obtain higher order derivatives of the Bernstein type basis functions by the following theorem:

2.6. Theorem. *Let a and b be nonnegative real parameters with $a\neq b$. Let m be a positive integer and let $x\in[a,b]$. Let k, l and n be nonnegative integers with $n\geq k$. Then*

$$\frac{d^{l}\mathbb{Y}_{k}^{n}(x;a,b,m)}{dx^{l}}=\frac{n!}{(n-l)!}\sum_{j=0}^{l}(-1)^{l-j}\left(\begin{array}{c}l\\j\end{array}\right)\frac{\mathbb{Y}_{k-j}^{n-l}(x;a,b,m-j)}{(b-a)^{j}}.$$

Remark 8. Substituting $a=0$, $b=1$ and $m=n$ into Theorem 2.6, we have

$$\frac{d^{l}B_{k}^{n}(x)}{dx^{l}}=\frac{n!}{(n-l)!}\sum_{j=0}^{l}(-1)^{l-j}\left(\begin{array}{c}l\\j\end{array}\right)B_{k-j}^{n-l}(x),$$

Substituting $l=1$ into the above equation, we have

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

(cf. [9], [8, Chapter 5, pp. 299-306], [18]) and (cf. [1]-[19]).

2.8. Recurrence Relation. Here, by using higher order derivatives of the novel generating function with respect to t, we derive a partial differential equation. Using this equation, we shall give a new proof of the recurrence relation for the Bernstein type basis functions.

Differentiating Equation (2.4) with respect to t, we prove a recurrence relation for the Bernstein type basis functions. This recurrence relation can also be obtained from Equation (2.1). By using Leibnitz's formula for the vth derivative, with respect to t, of the product $f_{\mathbb{Y},k}(x,t;a,b,m)$ of two function

$$g(t,x;a,b)=\frac{t^{k}\left(x-a\right)^{k}}{(b-a)^{m}k!}\quad(a\neq b)$$

and

$$h(t,x;b)=e^{(b-x)t},$$

we obtain another higher order partial differential equation as follows:

$$\frac{\partial^v f_{\mathbb{Y},k}(x,t;a,b,m)}{\partial t^v}=\sum_{j=0}^{v}\begin{pmatrix} v\\ j\end{pmatrix}\left(\frac{\partial^j g(t,x;a,b)}{\partial t^j}\right)\left(\frac{\partial^{v-j}h(t,x;b)}{\partial t^{v-j}}\right). \tag{2.17}$$

By using induction on v, Equation (2.17) is easily obtained.

2.7. Theorem. *Let v be an integer number. Then*

$$\frac{\partial^v f_{\mathbb{Y},k}(x,t;a,b,m)}{\partial t^v}=\sum_{j=0}^{v}(b-a)^{v-j}B_j^v(x;a,b)f_{\mathbb{Y},k-j}(x,t;a,b,m-j),$$

where $f_{\mathbb{Y},k}(x,t;a,b,m)$ and $B_j^v(x;a,b)$ are defined in (2.4) and (2.1), respectively.

Proof. Proof of Theorem 2.7 follows immediately from (2.17). □

Using definition (2.3), (2.1), and Theorem 2.7, we obtain a recurrence relation for the Bernstein type basis functions by the following theorem:

2.8. Theorem. *Let a and b be nonnegative real parameters with $a\neq b$. Let m be a positive integer and let $x\in[a,b]$. Let k, v and n be nonnegative integers with $n\geq k$. Then*

$$\mathbb{Y}_k^n(x;a,b,m)=\sum_{j=0}^{v}(b-a)^{v-j}B_j^v(x;a,b)\mathbb{Y}_{k-j}^{n-v}(x;a,b,m-j).$$

Remark 9. Substituting $a=0$ and $b=1$ into Theorem 2.8, we obtain the following result (cf. [18]):

$$B_k^n(x)=\sum_{j=0}^{v}B_j^v(x)B_{k-j}^{n-v}(x).$$

Substituting $v=1$ into above equation, we have (cf. [1]-[19])

$$B_k^n(x)=(1-x)B_k^{n-1}(x)+xB_{k-1}^{n-1}(x).$$

2.9. Multiplication and division by powers of $(\frac{x-a}{b-a})^d$ and $(\frac{b-x}{b-a})^d$. In [4], Buse and Goldman present much background material on computations with Bernstein polynomials. They provide formulas for multiplication and division of Bernstein polynomials by powers of x and $1-x$ and for degree elevation of Bernstein polynomials. Our method is similar to that of Buse and Goldman's [4]. Here, we find two functional equations. Using these equations, we also give new proofs of both the multiplication and division properties for the Bernstein polynomials.

By using the generating function in (2.4) , we provide formulas for multiplying Bernstein polynomials by powers of $\left(\frac{x-a}{b-a}\right)^d$ and $\left(\frac{b-x}{b-a}\right)^d$ and for degree elevation of the Bernstein polynomials.

Using (2.4), we obtain the following functional equation:

$$\left(\frac{x-a}{b-a}\right)^d f_{\mathbb{Y},k}(x,t;a,b,n)=\frac{(k+d)!}{k!t^d}f_{\mathbb{Y},k}(x,t;a,b,n).$$

After elementary manipulations in this equation, we get

$$\left(\frac{x-a}{b-a}\right)^d B_k^n(x;a,b)=\frac{n!(k+d)!}{k!(n+d)!}B_{k+d}^{n+d}(x;a,b). \tag{2.18}$$

Substituting $d = 1$, we have

$$\left(\frac{x-a}{b-a}\right) B_k^n(x;a,b) = \frac{k+1}{n+1} B_{k+1}^{n+1}(x;a,b). \tag{2.19}$$

Remark 10. Substituting $a = 0$ and $b = 1$ into (2.19), we have

$$xB_k^n(x) = \frac{k+1}{n+1} B_{k+1}^{n+1}(x).$$

The above relation can also be proved by (2.2) (cf. [4]).

Similarly, using (2.1), we obtain

$$\left(\frac{b-x}{b-a}\right)^d B_k^n(x;a,b) = \frac{n!(n+d-k)!}{(n+d)!(n-k)!} B_k^{n+d}(x;a,b).$$

Substituting $d = 1$ into the above equation, we have

$$\left(\frac{b-x}{b-a}\right) B_k^n(x;a,b) = \frac{n+1-k}{n+1} B_k^{n+1}(x;a,b). \tag{2.20}$$

Consequently, by the same method as in [4], if we have (2.8), then

$$\left(\frac{x-a}{b-a}\right)^d \mathcal{P}(x,a,b) = \sum_{k=0}^{n} c_k^n \frac{n!(k+d)!}{k!(n+d)!} B_{k+d}^{n+d}(x;a,b), \tag{2.21}$$

and

$$\left(\frac{b-x}{b-a}\right)^d \mathcal{P}(x,a,b) = \sum_{k=0}^{n} c_k^n \frac{n!(n+d-k)!}{(n+d)!(n-k)!} B_k^{n+d}(x;a,b). \tag{2.22}$$

We now consider *division* properties. We assume that (2.8) holds and that we are given an integer $j > 0$. Since $\left(\frac{x-a}{b-a}\right)^j$ divides $B_k^n(x;a,b)$ for all $k \geq j$, it follows that $\left(\frac{x-a}{b-a}\right)^j$ divides $\mathcal{P}(x,a,b)$. Similarly, using (2.4), we obtain the following functional equation:

$$\frac{f_{\mathbb{Y},k}(x,t;a,b,n)}{\left(\frac{x-a}{b-a}\right)^j} = \frac{(k-f)!t^j}{k!} f_{\mathbb{Y},k-j}(x,t;a,b,n-j).$$

For $k \geq j$, from the above equation, we have

$$\frac{B_k^n(x;a,b)}{\left(\frac{x-a}{b-a}\right)^j} = \frac{n!(k-j)!}{k!(n-j)!} B_{k-j}^{n-j}(x;a,b).$$

By a calculation similar to that in [4], for $j \leq n-k$, we have

$$\frac{B_k^n(x;a,b)}{\left(\frac{b-x}{b-a}\right)^j} = \frac{n!(n-j-k)!}{(n-k)!(n-j)!} B_k^{n-j}(x;a,b).$$

Therefore

$$\frac{\mathcal{P}(x,a,b)}{\left(\frac{x-a}{b-a}\right)^j} = \sum_{k=j}^{n} c_k^n \frac{n!(k-j)!}{k!(n-j)!} B_{k-j}^{n-j}(x;a,b), \tag{2.23}$$

and

$$\frac{\mathcal{P}(x,a,b)}{\left(\frac{b-x}{b-a}\right)^j} = \sum_{k=0}^{n-j} c_k^n \frac{n!(n-j-k)!}{(n-k)!(n-j)!} B_k^{n-j}(x;a,b). \tag{2.24}$$

2.10. Degree elevation. According to Buse and Goldman [4], given a polynomial represented in the univariate Bernstein basis of degree n, degree elevation computes representations of the same polynomial in the univariate Bernstein bases of degree greater than n. Degree elevation allows us to add two or more Bernstein polynomials which are not represented in the same degree Bernstein basis functions.

Adding (2.19) and (2.20), we obtain the degree elevation formula for the Bernstein basis functions:

$$B_k^n(x;a,b) = \frac{k+1}{n+1} B_{k+1}^{n+1}(x;a,b) + \frac{n+1-k}{n+1} B_k^{n+1}(x;a,b).$$

Substituting $d = 1$ into (2.22), and adding it with the latter equations gives the following degree elevation formula for the Bernstein polynomials:

$$\mathcal{P}(x,a,b) = \sum_{k=0}^{n} \left(\frac{k}{n+1} c_{k-1}^n + \frac{n+1-k}{(n+1)} c_k^n \right) B_k^{n+1}(x;a,b), \tag{2.25}$$

where

$$c_k^{n+1} = \frac{k}{n+1} c_{k-1}^n + \frac{n+1-k}{(n+1)} c_k^n.$$

Remark 11. If we set $a = 0$ and $b = 1$, then Equation (2.25) reduces to Equation (2.5) in [4, p. 853].

3. Application of the Fourier and the Laplace transforms to the generating functions

In this section, by applying the Fourier transform and the Laplace transform to the generating function for the Bernstein basis functions, we obtain some interesting series representations for the Bernstein basis functions.

In [18, p. 5, Eq. (11)], the following functional equation was derived:

$$f_{\mathbb{B},j}(xy,t) = f_{\mathbb{B},j}(x,ty)\, e^{t(1-y)}. \tag{3.1}$$

From this generating function, we obtain subdivision property for the Bernstein basis functions (see [18]):

$$B_j^n(xy) = \sum_{k=j}^{n} B_j^k(x) B_k^n(y)$$

cf. (see also [9]-[8, Chapter 5, pp. 299-306]).

By using (3.1), we obtain functional equation

$$f_{\mathbb{B},k}(xy,t)e^{-t} = f_{\mathbb{B},k}(x,ty)\, e^{-ty}.$$

For $a = 0$ and $b = 1$, combining (2.4) with the above equation, we get

$$\sum_{n=0}^{\infty} B_k^n(xy) \frac{t^n}{n!} e^{-t} = \sum_{n=0}^{\infty} B_k^n(x) y^n \frac{t^n}{n!} e^{-ty}. \tag{3.2}$$

Integrate this equation (by parts) with respect to t from 0 to ∞, we get

$$\sum_{n=0}^{\infty} \frac{B_k^n(xy)}{n!} \int_0^{\infty} t^n e^{-t} dt = \sum_{n=0}^{\infty} \frac{B_k^n(x) y^n}{n!} \int_0^{\infty} t^n e^{-ty} dt.$$

By using the Laplace transform in the above equation, we arrive at the following Theorem:

3.1. Theorem. *Let $x, y \in [0,1]$. The following relationship holds true:*

$$\sum_{n=0}^{\infty} B_k^n(xy) = \sum_{n=0}^{\infty} \frac{1}{y} B_k^n(x).$$

From (2.4), we define the following functional equation:

$$\frac{t^k\,(x-a)^k}{(b-a)^m k!} e^{-xt} = \sum_{n=0}^{\infty} \mathbb{Y}_k^n(x;a,b,m)\frac{t^n}{n!}e^{-bt}.$$

By applying the Fourier transform to the above equation,

$$\frac{(x-a)^k}{(b-a)^m k!}\int_0^{\infty} t^k e^{-xt} e^{-ist}dt = \sum_{n=0}^{\infty} \mathbb{Y}_k^n(x;a,b,m)\frac{1}{n!}\int_0^{\infty} t^n e^{-bt}e^{-ist}dt.$$

From this equation, we arrive at the following Theorem:

3.2. Theorem. *Let $x \in [a,b]$ and $s \in \mathbb{R}$. We have*

$$\sum_{n=0}^{\infty} \frac{\mathbb{Y}_k^n(x;a,b,m)}{(b+is)^{n+1}} = \frac{(x-a)^k}{(b-a)^m\,(x+is)^{k+1}},$$

where $\left|\frac{b-x}{b+is}\right| < 1$.

4. New Identities

By using novel generating functions, we derive some new identities related to the Bernstein type basis function.

4.1. Theorem.

$$\sum_{j=0}^{n}\sum_{k=0}^{j}(-1)^k \binom{n}{j} \mathbb{Y}_k^j(x;a,b,m)\,(2x)^{n-j} = (b-a)^{n-m}.$$

Proof. By using (2.14), we obtain

$$\sum_{k=0}^{\infty}(-1)^k f_{\mathbb{Y},k}(x,t;a,b,m)e^{2xt} = \frac{1}{(b-a)^m}e^{(b-a)t}. \tag{4.1}$$

From this equation, we get

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}(-1)^k\mathbb{Y}_k^n(x;a,b,m)\frac{t^n}{n!}\sum_{n=0}^{\infty}(2x)^n\,\frac{t^n}{n!} = \sum_{n=0}^{\infty}(b-a)^{n-m}\,\frac{t^n}{n!}.$$

Therefore

$$\sum_{n=0}^{\infty}\left(\sum_{j=0}^{n}\sum_{k=0}^{j}(-1)^k \binom{n}{j} \mathbb{Y}_k^j(x;a,b,m)\,(2x)^{n-j}\right)\frac{t^n}{n!} = \sum_{n=0}^{\infty}(b-a)^{n-m}\,\frac{t^n}{n!}.$$

Comparing the coefficients of $\frac{t^n}{n!}$ on the both sides of the above equation, we arrive at the the desired result. □

4.2. Theorem.

$$\sum_{k=j}^{n}(-1)^{n-k}\binom{n}{k}B_j^k(xy) = y^n\sum_{k=j}^{n}(-1)^{n-k}\binom{n}{k}B_j^k(x).$$

Proof. Using (3.2), we obtain

$$\sum_{n=0}^{\infty} B_k^n(xy)\frac{t^n}{n!}\sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!} = \sum_{n=0}^{\infty} B_k^n(x)y^n\frac{t^n}{n!}\sum_{n=0}^{\infty}(-y)^n\frac{t^n}{n!}.$$

From the above equation, we get

$$\sum_{n=j}^{\infty}\left(\sum_{k=j}^{n}(-1)^{n-k}\begin{pmatrix} n \\ k \end{pmatrix} B_j^k(xy)\right)\frac{t^n}{n!} = \sum_{n=j}^{\infty}\left(y^n\sum_{k=j}^{n}(-1)^{n-k}\begin{pmatrix} n \\ k \end{pmatrix} B_j^k(x)\right)\frac{t^n}{n!}.$$

Comparing the coefficients of $\frac{t^n}{n!}$ on the both sides of the above equation, we arrive at the the desired result. □

5. Further remarks and observations on the generating functions $f_{\mathbb{Y},k}(x,t;a,b,m)$, Poisson distribution and Szasz-Mirakjan type basis functions

The identity of Jetter and Stöckler represents a pointwise orthogonality relation for the multivariate Bernstein polynomials on a simplex. This identity give us a new representation for the dual basis which can be used to construct general quasi-interpolant operators (cf., see, for details, [10] and [1]). As an application of the generating functions for the basis functions to the identity of Jetter and Stöckler, Abel and Li [1] proved Proposition 5.1, which is given in this section. Applying our generating functions to Proposition 5.1, we give pointwise orthogonality relations for the Bernstein polynomials and the Szasz-Mirakjan basis functions.

In this section, we give relations between the Bernstein basis functions, the binomial distribution and the Poisson distribution. First we consider the generalized binomial or Newton distribution (probability function). Suppose that $0 \leq \frac{x-a}{b-a} \leq 1$ and $0 \leq \frac{b-x}{b-a} \leq 1$. Set

$$B_k^n(x;a,b) = \begin{pmatrix} n \\ k \end{pmatrix}\left(\frac{x-a}{b-a}\right)^k\left(\frac{b-x}{b-a}\right)^{n-k}. \tag{5.1}$$

Remark 12. If we set $a=0$ and $b=1$, then (5.1) reduces to

$$B_k^n(x) = \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k}$$

which is the binomial or Newton distribution (probabilities) function. If $0 \leq x \leq 1$ is the probability of an event E, then $B_k^n(x)$ is the probability that E will occur exactly k times in n independent trials (cf. [13]).

Expected value or mean and variance of $B_k^n(x;a,b)$ are given by

$$\mu = \sum_{k=0}^{n} kB_k^n(x;a,b) = n\left(\frac{x-a}{b-a}\right),$$

and

$$\sigma^2 = \sum_{k=0}^{n} k^2B_k^n(x;a,b) - \mu^2 = \frac{n\,(x-a)\,(b-x)}{(b-a)^2}.$$

If we let $n \to \infty$ in (5.1), then we arrive at the well-known Poisson distribution:

$$B_k^n\left(\frac{b-a}{n}\mu + a;a,b\right) \to \frac{\mu^k e^{-\mu}}{k!}. \tag{5.2}$$

The following proposition is proved by Abel and Li [1, p. 300, Proposition 3]:

5.1. Proposition. *Let the system $\{f_n(x)\}$ of functions be defined by the generating function*

$$A_t(x) = \sum_{n=0}^{\infty} f_n(x) t^n.$$

If there exists a sequence $w_k = w_k(x)$ such that

$$\sum_{k=0}^{\infty} w_k \mathcal{D}^k A_t(x) \mathcal{D}^k A_z(x) = A_{tz}(x)$$

with $\mathcal{D} = \frac{d}{dx}$, then we have

$$\sum_{k=0}^{\infty} w_k \mathcal{D}^k f_i(x) \mathcal{D}^k f_j(x) = \delta_{i,j} f_i(x), (i, j = 0, 1, \ldots).$$

As an application of Proposition 5.1, Abel and Li [1] use the generating function in Equation (2.7) for the Bernstein basis functions. They also use generating functions for the Szasz-Mirakjan basis functions and Baskakov basis functions.

In this section, we apply our novel generating functions to Proposition 5.1, which give pointwise orthogonality relations for the Bernstein polynomials and the Szasz-Mirakjan type basis functions, respectively.

As applications of Proposition 5.1, we give the following examples:

Example 1. For given n and k, the Bernstein basis functions

$$f_k(x, n; a, b) = B_k^n(x; a, b) = \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{x-a}{b-a} \right)^k \left(\frac{b-x}{b-a} \right)^{n-k}$$

are generated by the function in (2.4), that is

$$A_t(x) = \frac{t^k (x-a)^k e^{(b-x)t}}{(b-a)^n k!} = \sum_{k=0}^{\infty} \frac{f_k(x, n; a, b)}{k!} t^k.$$

It is easy to check that Proposition 5.1 holds with $w_k = w_k(x) = B_k^n(x; a, b)$.

Example 2. Using (5.2), for $j \geq 0$, we generalize the *Szasz-Mirakjan type basis functions* as follows

$$f_j(x, n; a, b) = \frac{(n\frac{x-a}{b-a})^j e^{-n\frac{x-a}{b-a}}}{j!},$$

where a and b are nonnegative real parameters with $a \neq b$, n is a positive integer and $x \in [a, b]$. The functions $f_j(x, n; a, b)$ are generated by

$$A_t(x) = \exp\left((t-1)n \left(\frac{x-a}{b-a} \right) \right) = \sum_{i=0}^{\infty} f_i(x, n; a, b) t^i,$$

where $\exp(x) = e^x$. In this case, Proposition 5.1 holds with $w_k = w_k(x) = \frac{\left(\frac{x-a}{b-a}\right)^k}{n^k k!}$. Therefore, we have

$$\sum_{k=0}^{\infty} \frac{\left(\frac{x-a}{b-a}\right)^k}{n^k k!} \mathcal{D}^k f_i(x, n; a, b) \mathcal{D}^k f_j(x, n; a, b) = \delta_{i,j} f_i(x, n; a, b).$$

Remark 13. If $a = 0$ and $b = 1$ in Example 2, then we arrive at the *Szasz-Mirakjan basis functions which are given in* [1, p. 300, Example 2].

Acknowledgement. The author would like to thank Professor Ronald Goldman (Rice University, Houston, USA) for his very valuable comments, criticisms and very useful suggestions on this paper.

The present investigation was supported by the *Scientific Research Project Administration of Akdeniz University.*

References

[1] Abel, U. and Li, Z., *A new proof of an identity of Jetter and Stöckler for multivariate Bernstein polynomials*, Comput. Aided Geom. Design. **23** (3), 297–301, 2006.

[2] Acikgoz, M. and Araci, S., *On generating function of the Bernstein polynomials*, Numerical Anal. Appl. Math., Amer. Inst. Phys. Conf. Proc. CP**1281**, 1141–1143, 2010.

[3] Bernstein, S. N., *Démonstration du théorème de Weierstrass fondée sur la calcul des probabilités*, Comm. Soc. Math. Charkow Sér. 2 t. **13**, 1-2**,** 1912–1913.

[4] Busé, L. and Goldman, R., *Division algorithms for Bernstein polynomials*, Comput. Aided Geom. Design. **25** (9), 850–865, 2008.

[5] Farouki, R. T. and Goodman, T. N. T., *On the optimal stability of the Bernstein basis*, Math. Comput. **65**,1553–1566, 1996.

[6] Büyükyazıcı, I. and Ibikli, E., *The approximation properties of generalized Bernstein polynomials of two variables*, Appl. Math. Comput. **156**, 367–380, 2004.

[7] Goldman, R., *An Integrated Introduction to Computer Graphics and Geometric Modeling*, (CRC Press, Taylor and Francis, New York, 2009).

[8] Goldman, R., *Pyramid Algorithms: A Dynamic Programming Approach to Curves and Surfaces for Geometric Modeling*, (Morgan Kaufmann Publishers, R. Academic Press, San Diego, 2002).

[9] Goldman, R., *Identities for the Univariate and Bivariate Bernstein Basis Functions*, Graphics Gems V, (edited by Alan Paeth, Academic Press, 1995) 149–162.

[10] Jetter, K. and Stöckler, J., *An identity for multivariate Bernstein poynomials*, Comput. Aided Geom. Design. **20**, 563–577, 2003.

[11] Kim, M. S., Kim, D. and Kim, T., *On the q-Euler numbers related to modified q-Bernstein polynomials*, Abstr. Appl. Anal. **2010**, Art. ID 952384, 15 pages.

[12] Lewanowicz, S. and Woźny, P., *Generalized Bernstein polynomials,* BIT Numer. Math. **44**, 63–78, 2004.

[13] Lorentz, G. G., *Bernstein Polynomials*, (Chelsea Pub. Comp. New York, N. Y. 1986).

[14] Phillips, G. M., *Interpolation and approximation by polynomials*, (CMS Books in Mathematics/ Ouvrages de Mathématiques de la SMC, 14. Springer–Verlag, New York, 2003).

[15] Phillips, G. M., *Bernstein polynomials based on the q-integers*, Ann. Numer. Math. **4**, 511–518, 1997.

[16] Oruc, H. and Phillips, G. M., *A generalization of the Bernstein polynomials*, Proc. Edinb. Math. Soc. **42**, 403–413, 1999.

[17] Simsek, Y., *Interpolation function of generalized q-Bernstein type polynomials and their application*, Lecture Notes in Computer Science **6920**, (Springer-Verlag, Berlin, 2011), 647–662.

[18] Simsek, Y., *Functional equations from generating functions: a novel approach to deriving identities for the Bernstein basis functions*, Fixed Point Theory and Applications, 2013, 2013:80.

[19] Simsek, Y. and Acikgoz, M., *A new generating function of (q-) Bernstein-type polynomials and their interpolation function*, Abstr. Appl. Anal. **2010**, Art. ID 769095, 12 pp.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 15 – 26

INTERNAL STATE VARIABLES IN DIPOLAR THERMOELASTIC BODIES

M. Marin [a] *, S. R. Mahmoud [b c], and G. Stan [a]

Received 27 : 01 : 2012 : Accepted 03 : 10 : 2012

Abstract

The aim of our study is prove that the presence of the internal state variables in a thermoelastic dipolar body do not influence the uniqueness of solution. After the mixed initial boundary value problem in this context is formulated, we use the Gronwall's inequality to prove the uniqueness of solution of this problem.

Keywords: thermoelastic, dipolar, internal state variables, uniqueness, Gronwall's inequality

2000 AMS Classification: 35A25, 35G46, 74A60, 74H25, 80A20

1. Introduction

Interest to consider the internal state variables as a means to estimate mechanical properties has grown rapidly in recent years.

The theories of internal state variables in different kind of materials represent a material length scale and are quite sufficient for a large number of the solid mechanics applications.

The internal state variables are the smallest possible subset of system variables that can represent the entire state of the system at any given time. The minimum number of state variables required to represent a given system, n, is usually equal to the order of the differential equations system's defining. If the system is represented in the transfer function form, the minimum number of state variables is equal to the order of the transfer function's denominator after it has been reduced to a proper fraction. It is important to understand that converting a state space realization to a transfer function form may lose some internal information about the system, and may provide a description of a system which is stable, when the state-space realization is unstable at certain points. For instance, in the electric circuits, the number of state variables is often, though not

[a]Dept. of Mathematics and Computer Science, Transilvania University of Brasov, Romania.
*Corresponding author
[b]Department of Mathematics, King Abdulaziz University, Jeddah, Saudi Arabia.
[c]Department of Mathematics, Science Faculty, Sohag University, Egypt

always, the same as the number of energy storage elements in the circuit such as capacitors and inductors.

The theory of bodies with internal state variables has been first formulated for the thermo-viscoelastic materials (see, for instance Chirita [3]). Then the internal state variables has been considered for different kind of materials.

The study [9] of Nachlinger and Nunziato is dedicated to the internal state variables approach of finite deformations without heat conduction in the one-dimensional case.

In the paper [12] the authors describe how the so-called Bammann internal state variable constitutive approach, which has proven highly successful in modelling deformation processes in metals, can be applied with great benefit to silicate rocks and other geological materials in modelling their deformation dynamics. In its essence, the internal state variables theory provides a constitutive framework to account for changing history states that arise from inelastic dissipative microstructural evolution of a polycrystalline solid.

A thermodynamically consistent framework is proposed for modeling the hysteresis of capillarity in partially saturated porous media in the paper [14]. Capillary hysteresis is viewed as an intrinsic dissipation mechanism, which can be characterized by a set of internal state variables. The volume fractions of pore fluids are assumed to be additively decomposed into a reversible part and an irreversible part. The irreversible part of the volumetric moisture content is introduced as one of the internal variables. It is shown that the pumping effect occurring in a porous medium experiencing a wetting/drying cycle is thermodynamically admissible.

The paper [2] presents the formulation of a constitutive model for amorphous thermoplastics using a thermodynamic approach with physically motivated internal state variables. The formulation follows current internal state variable methodologies used for metals and departs from the spring-dashpot representation generally used to characterize the mechanical behavior of polymers.

Anand and Gurtin develop in the paper [1] a continuum theory for the elastic-viscoplastic deformation of amorphous solids such as polymeric and metallic glasses. Introducing an internal-state variable that represents the local free-volume associated with certain metastable states, the authors are able to capture the highly non-linear stress-strain behavior that precedes the yield-peak and gives rise to post-yield strain softening.

In the study [13], is presented a formulation of state variable based gradient theory to model damage evolution and alleviate numerical instability associated within the post-bifurcation regime. This proposed theory is developed using basic microforce balance laws and appropriate state variables within a consistent thermodynamic framework. The proposed theory provides a strong coupling and consistent framework to prescribe energy storage and dissipation associated with internal damage. For other paper in this topic, see [10], [11].

Other results on some generalizations of thermoelastic bodies can be found in the papers [4]-[8].

2. Basic equations

Let us consider B be an open region of three-dimensional Euclidean space R^3 occupied, at time $t = 0$, by the reference configuration of a thermoelastic dipolar body with internal state variables.

We assume that the boundary of the domain B, denoted by ∂B, is a closed, bounded and piece-wise smooth surface which allows us the application of the divergence theorem. A fixed system of rectangular Cartesian axes is used and we adopt the Cartesian tensor notations. The points in B are denoted by (x_i) or (x). The variable t is the time and $t \in [0, t_0)$. We shall employ the usual summation over repeated subscripts while subsripts preceded by a comma denote the partial differentiation with respect to the spatial argument. Also, we use a superposed dot to denote the partial differentiation with respect to t. The Latin indices are understood to range over the integers (1, 2, 3), while the Greek subsripts have the range $1,\ 2, \ldots, n$.

In the following we designate by n_i the components of the outward unit normal to the surface ∂B. The closure of the domain B, denoted by $\bar{B}$, means $\bar{B} = B \cup \partial B$.

Also, the spatial argument and the time argument of a function will be ommited when there is no likelihood of confusion.

The behaviour of a thermoelastic dipolar body is characterized by the following kinematic variables:

$$u_i = u_i(x,t),\ \varphi_{jk} = \varphi_{jk}(x,t),\ (x,t) \in B \times [0, t_0)$$

where u_i are the components of the displacement field and φ_{jk} - the components of the dipolar displacement field.

The fundamental system of field equations, in the theory of dipolar thermoelastic bodies with internal state variables, consists of:

- the equations of motion:

$$\begin{aligned} &(\tau_{ij} + \eta_{ij})_{,j} + \varrho F_i = \varrho \ddot{u}_i, \\ &\mu_{ijk,i} + \sigma_{jk} + \varrho G_{jk} = I_{kr} \ddot{\varphi}_{jr}; \end{aligned} \tag{2.1}$$

- the energy equation:

$$T_0 \dot{\eta} = q_{i,i} + \varrho r; \tag{2.2}$$

- the constitutive equations:

$$\begin{aligned} \tau_{ij} &= C_{ijmn}\,\varepsilon_{mn} + G_{mnij}\,\gamma_{mn} + F_{mnrij}\,\kappa_{mnr} - B_{ij}\,\theta + B_{ij\alpha}\,\omega_\alpha, \\ \sigma_{ij} &= G_{ijmn}\,\varepsilon_{mn} + B_{ijmn}\,\gamma_{mn} + D_{ijmnr}\,\kappa_{mnr} - D_{ij}\,\theta + D_{ij\alpha}\,\omega_\alpha, \\ \mu_{ijk} &= F_{ijkmn}\,\varepsilon_{mn} + D_{mnijk}\,\gamma_{mn} + A_{mnrijk}\,\kappa_{mnr} - F_{ijk}\,\theta + F_{ijk\alpha}\,\omega_\alpha, \\ \eta &= B_{ij}\,\varepsilon_{ij} + D_{ij}\,\gamma_{ij} + F_{ijs}\,\kappa_{ijs} - a\,\theta - G_\alpha\,\omega_\alpha, \\ q_i &= a_{ijk}\,\varepsilon_{jk} + b_{ijk}\,\gamma_{jk} + c_{ijsm}\,\kappa_{jsm} + d_i\,\theta + f_{i\alpha}\,\omega_\alpha + K_{ij}\,\theta_{,\,j}; \end{aligned} \tag{2.3}$$

- the geometric equations:

$$\begin{aligned} &\varepsilon_{ij} = \frac{1}{2}\left(u_{j,i} + u_{i,j}\right),\quad \gamma_{ij} = u_{j,i} - \varphi_{ij}, \\ &\kappa_{ijk} = \varphi_{jk,i}. \end{aligned} \tag{2.4}$$

Usually, the internal state variables are denoted by $\xi_\alpha,\ \alpha = 1,\ 2,\ \ldots,\ n$. In the linear theory, we denote by ω_α the internal state variables measured from the internal state variables ξ^0_α of the initial state. Also, the temperature θ represents the difference between the absolute temperature T and the temperature T_0, $T_0 > 0$, of the initial state. Thus we have:

$$\xi_\alpha = \xi^0_\alpha + \omega_\alpha,\ T = T_0 + \theta. \tag{2.5}$$

Within the linear approximation, from the entropy production inequality, it follows (see, for instance, [1]):

$$\dot{\omega}_\alpha = f_\alpha, \tag{2.6}$$

where

$$f_\alpha = g_{ij\alpha}\varepsilon_{ij} + h_{ij\alpha}\gamma_{ij} + l_{ijk\alpha}\kappa_{ijk} + p_\alpha\theta + q_{\alpha\beta}\omega_\beta + r_{i\alpha}\theta_{,\,i}. \tag{2.7}$$

The other notations used in the above equations have the following meanings:

- ϱ - the constant mass density;
- $\tau_{ij},\ \sigma_{ij},\ \mu_{ijk}$ - the components of the stress tensors;

- I_{ij} - the coefficients of inertia;
- F_i - the components of body force per unit mass;
- G_{jk} - the components of dipolar body force per unit mass;
- r - the heat supply per unit mass and unit time;
- η - the entropy per unit mass;
- q_i - the components of the heat flux;
- $\varepsilon_{ij}, \gamma_{ij}, \kappa_{ijk}$ - the kinematic characteristics of the strain tensors.

The above coefficients $C_{ijmn},\ B_{ijmn}, \ldots,\ D_{ijm},\ E_{ijm}, \ldots, a_{ijk},\ \ldots,\ g_{ij\alpha}, \ldots,\ r_{i\alpha}$ are functions of x and characterize the thermoelastic properties of the material with internal state variable (the constitutive coefficients). For a homogeneous medium these quantities are constants. The constitutive coefficients obey to the following symmetry relations

$$\begin{aligned} &C_{ijmn} = C_{mnij} = C_{ijnm},\ B_{ijmn} = B_{mnij}, \\ &G_{ijmn} = G_{ijnm}, F_{ijkmn} = F_{ijknm},\ A_{ijkmnr} = A_{mnrijk}, \\ &B_{ij} = B_{ji},\ a_{ijk} = a_{ikj},\ K_{ij} = K_{ji},\ g_{ij\alpha} = g_{ji\alpha}. \end{aligned} \tag{2.8}$$

We supplement the above equations with the following initial conditions

$$\begin{aligned} &u_i\left(x_s, 0\right) = u_{0i}\left(x_s\right),\ \dot{u}_i\left(x_s, 0\right) = u_{1i}\left(x_s\right), \\ &\varphi_{ij}\left(x_s, 0\right) = \varphi_{0ij}\left(x_s\right),\ \dot{\varphi}_{ij}\left(x_s, 0\right) = \varphi_{1ij}\left(x_s\right), \\ &\theta\left(x_s, 0\right) = \theta_0\left(x_s\right),\ \omega_\alpha\left(x_s, 0\right) = \omega_{0\alpha}\left(x_s\right),\ \left(x_s\right) \in B \end{aligned} \tag{2.9}$$

and the prescribed boundary conditions

$$\begin{aligned} &u_i = \tilde{u_i},\ \text{on } \overline{\partial B_1} \times [0, t_0]\,,\ t_i \equiv \left(\tau_{ij} + \sigma_{ij}\right) n_j = \tilde{t_i},\ \text{on } \partial B_2 \times [0, t_0]\,, \\ &\varphi_{ij} = \tilde{\varphi}_{ij},\ \text{on } \overline{\partial B_3} \times [0, t_0]\,,\ \mu_{jk} \equiv \mu_{ijk} n_i = \tilde{\mu}_{jk},\ \text{on } \partial B_4 \times [0, t_0]\,, \\ &\theta = \tilde{\theta},\ \text{on } \overline{\partial B_5} \times [0, t_0]\,,\ q \equiv q_i n_i = \tilde{q},\ \text{on } \partial B_6 \times [0, t_0]\,. \end{aligned} \tag{2.10}$$

Here $\overline{\partial B_1}$, $\overline{\partial B_3}$, $\overline{\partial B_5}$ and ∂B_2, ∂B_4, ∂B_6 are subsets of the boundary ∂B which satisfay the relations

$$\begin{aligned} &\overline{\partial B_1} \cup \partial B_2 = \overline{\partial B_3} \cup \partial B_4 = \overline{\partial B_5} \cup \partial B_6 = \partial B \\ &\partial B_1 \cap \partial B_2 = \partial B_3 \cap \partial B_4 = \partial B_5 \cap \partial B_6 = \emptyset \end{aligned}$$

In the above conditions 2.9 and 2.10, the functions u_{0i}, u_{1i}, φ_{0ij}, φ_{1ij}, θ_0 $\omega_{0\alpha}$, $\tilde{u_i}$, $\tilde{t_i}$, $\tilde{\varphi_{ij}}$, $\tilde{\mu_{jk}}$, $\tilde{\theta}$ and $\tilde{q}$ are prescribed in their domain of definition.

In conclusion, the mixed initial boundary value problem of the thermoelasticity of dipolar bodies with internal variables consists of the equations (2.1), (2.2) and (2.6), the initial conditions (2.9) and the boundary conditions (2.10).

By a solution of this problem we mean a state of deformation $(u_i,\ \varphi_{ij},\ \theta,\ \omega_\alpha)$ satisfying the Eqns. (2.1), (2.2) and (2.6) and the conditions (2.9) and (2.10).

3. Main results

In the main section of our paper we will deduce some estimations and then, as a consequence, we obtain in simple manner the uniqueness theorem of the solution of the above problem.

In order to prove these results, we shall need the following assumptions

- (i) the mass density ϱ is strictly positive, i.e.

$$\varrho\left(x_s\right) \geq \varrho_0 > 0,\ \text{on } B;$$

- (ii) there exists a positive constant λ_1 such that

$$I_{ij}\xi_i\xi_j \geq \lambda_1\xi_i\xi_i,\ \forall\ \xi_i;$$

- (iii) the specific heat a from $(3)_4$ is strictly positive, i.e.

$$a(x_s) \geq a_0 > 0, \text{ on } B;$$

- (iv) the constitutive tensors C_{ijmn}, B_{ijmn} and A_{ijkmnr} are positive definite:

$$\int_B C_{ijmn}\, \xi_{ij}\, \xi_{mn}\, dv \geq \lambda_2 \int_B \xi_{ij}\, \xi_{ij}\, dv, \;\forall\, \xi_{ij}$$
$$\int_B B_{ijmn}\, \xi_{ij}\, \xi_{mn}\, dv \geq \lambda_3 \int_B \xi_{ij}\, \xi_{ij}\, dv, \;\forall\, \xi_{ij}$$
$$\int_B A_{ijkmnr}\, \xi_{ijk}\, \xi_{mnr}\, dv \geq \lambda_4 \int_B \xi_{ijk}\, \xi_{ijk}\, dv, \;\forall\, \xi_{ijk}$$

where λ_2, λ_3 and λ_4 are positive constants;

- (v) the symmetric part $\tilde{K}_{ij}$ of the thermal conductivity tensor K_{ij} is positive definite, in the sense that there exists a positive constant μ such that

$$\int_B \tilde{K}_{ij}\, \xi_i\, \xi_j\, dv \geq \mu \int_B \xi_i\, \xi_i\, dv, \text{ for all vectors } \xi_i.$$

Let us consider

$$\left(u_i^{(\nu)},\, \varphi_{ij}^{(\nu)},\, \theta^{(\nu)},\, \omega_\alpha^{(\nu)}\right),\ \nu = 1,\ 2$$

two solutions of our initial boundary value problem.

Because of the linearity of the problem, their difference is also solution of the problem. We denote by $(v_i,\ \psi_{ij}, \kappa, w_\alpha)$ the differences,

$$v_i = u_i^{(2)} - u_i^{(1)},\ \psi_i = \varphi_{ij}^{(2)} - \varphi_{ij}^{(1)},\ \kappa = \theta^{(2)} - \theta^{(1)},\ w_\alpha = \omega_\alpha^{(2)} - \omega_\alpha^{(1)}$$

In order to prove the desired uniquness theorem, it suffice to prove that the above considered problem, consists of the equations (2.1), (2.2) and (2.6) and the conditions (2.9) and (2.10), in which

$$F_i = G_{jk} = r = 0$$
$$u_{0i} = u_{1i} = \varphi_{0ij} = \varphi_{1ij} = \theta_0 = \omega_{0\alpha} = 0$$

and

$$\tilde{u}_i = \tilde{t}_i \tilde{\varphi}_{ij} = \tilde{\mu}_{ij} = \tilde{\theta} = \tilde{q} = 0$$

imply that

$$u_i = \varphi_{ij} = \theta = \omega_\alpha = 0,$$

in $B \times [0, t_0]$, provided that the hypotheses (i) - (v) hold.
Therefore, we consider the new problem P_0 defined by the following equations

$$(\tau_{ij} + \sigma_{ij})_{,j} = \varrho \ddot{u}_i,$$
$$\mu_{ijk,i} + \sigma_{jk} = I_{kr} \ddot{\varphi}_{jr} \tag{3.1}$$

$$T_0 \dot{\eta} = q_{i,i} \tag{3.2}$$

$$\dot{\omega}_\alpha = f_\alpha, \tag{3.3}$$

with the initial conditions

$$u_i(x_s, 0) = 0,\ \dot{u}_i(x_s, 0) = 0,\ \varphi_{ij}(x_s, 0) = 0,$$
$$\dot{\varphi}_{ij}(x_s, 0) = 0,\ \theta(x_s, 0) = 0,\ \omega_\alpha(x_s, 0) = 0,\ (x_s) \in B \tag{3.4}$$

and the boundary conditions

$$
\begin{aligned}
&u_i = 0, \text{ on } \overline{\partial B_1} \times [0, t_0]\,, \; t_i \equiv (\tau_{ij} + \sigma_{ij})\, n_j = 0, \text{ on } \partial B_2 \times [0, t_0]\,, \\
&\varphi_{ij} = 0, \text{ on } \overline{\partial B_3} \times [0, t_0]\,, \; \mu_{jk} \equiv \mu_{ijk} n_i = 0, \text{ on } \partial B_4 \times [0, t_0]\,, \\
&\theta = 0, \text{ on } \overline{\partial B_5} \times [0, t_0]\,, \; q \equiv q_i n_i = 0, \text{ on } \partial B_6 \times [0, t_0]\,.
\end{aligned} \tag{3.5}
$$

To these equations and conditions we adjoin the constitutive relations (2.3) and (2.7). In order to prove that the problem P_0 admits the null solution, we will show that the function $y(t)$ defined by

$$
y(t) = \int_B \left(\dot{u}_i \dot{u}_i + \dot{\varphi}_{ij} \dot{\varphi}_{ij} + \varepsilon_{ij} \varepsilon_{ij} + \gamma_{ij} \gamma_{ij} + \kappa_{ijr} \kappa_{ijr} + \theta^2 + \omega_\alpha \omega_\alpha \right) dV
$$

vanishes on $[0, t_0]$.
To this aim, we first prove some useful estimations.

3.1. Theorem. *If the ordered array* $(u_i,\ \varphi_{ij},\ \theta,\ \omega_\alpha)$ *is a solution of the problem* P_0*, then the following relation hold*

$$
\begin{aligned}
&\frac{1}{2} \int_B \left(C_{ijmn} \varepsilon_{ij} \varepsilon_{mn} + 2G_{ijmn} \varepsilon_{ij} \gamma_{mn} + 2F_{mnrij} \varepsilon_{ij} \kappa_{mnr} + \right. \\
&+ B_{ijmn} \gamma_{ij} \gamma_{mn} + A_{ijsmnr} \kappa_{ijs} \kappa_{mnr} + 2D_{ijmnr} \gamma_{ij} \kappa_{mnr} + 2B_{ij\alpha} \varepsilon_{ij} \omega_\alpha + \\
&\left. + 2D_{ij\alpha} \gamma_{ij} \omega_\alpha + 2F_{ijr\alpha} \kappa_{ijr} \omega_\alpha + a\theta^2 + \varrho \dot{u}_i \dot{u}_i + I_{kr} \dot{\varphi}_{jr} \dot{\varphi}_{jk} \right) dV = \\
&\int_0^t \int_B \left[\left(B_{ij\alpha} \varepsilon_{ij} + D_{ij\alpha} \gamma_{ij} + F_{ijr\alpha} \kappa_{ijr} \right) \dot{\omega}_\alpha - \frac{1}{T_0} q_i \theta_{,i} \right] dV ds.
\end{aligned} \tag{3.6}
$$

Proof. By using the constitutive equations (2.3) and the symmetry relations (2.8), we obtain

$$
\begin{aligned}
&\tau_{ij} \dot{u}_{j,i} + \sigma_{ij} \dot{\gamma}_{ij} + \mu_{ijs} \dot{\kappa}_{ijs} = \\
&\frac{1}{2} \frac{\partial}{\partial t} \left(C_{ijmn} \varepsilon_{ij} \varepsilon_{mn} + 2G_{mnij} \varepsilon_{ij} \gamma_{mn} + 2F_{mnrij} \varepsilon_{ij} \kappa_{mnr} + \right. \\
&+ B_{ijmn} \gamma_{ij} \gamma_{mn} + A_{ijsmnr} \kappa_{ijs} \kappa_{mnr} + 2D_{ijmnr} \gamma_{ij} \kappa_{mnr} + \\
&\left. + 2B_{ij\alpha} \varepsilon_{ij} \omega_\alpha + 2D_{ij\alpha} \gamma_{ij} \omega_\alpha + 2F_{ijs\alpha} \kappa_{ijs} \omega_\alpha + a\theta^2 \right) - \\
&- B_{ij\alpha} \varepsilon_{ij} \dot{\omega}_\alpha - D_{ij\alpha} \gamma_{ij} \dot{\omega}_\alpha - F_{ijs\alpha} \kappa_{ijs} \dot{\omega}_\alpha - G_\alpha \theta \dot{\omega}_\alpha.
\end{aligned} \tag{3.7}
$$

On the other hand, in view of (3.1) and (3.2) we deduce:

$$
\begin{aligned}
&\tau_{ij} \dot{u}_{j,i} + \sigma_{ij} \dot{\gamma}_{ij} + \mu_{ijs} \dot{\kappa}_{ijs} = \\
&= \left[(\tau_{ij} + \sigma_{ij})\, \dot{u}_j + \mu_{ijs} \dot{\varphi}_{js} + \frac{1}{T_0} q_i \theta \right]_{,i} - \\
&- \frac{1}{2} \frac{\partial}{\partial t} \left(\varrho \dot{u}_i \dot{u}_i + I_{kr} \dot{\varphi}_{jr} \dot{\varphi}_{jk} \right) - \frac{1}{T_0} q_i \theta_{,i}
\end{aligned} \tag{3.8}
$$

From the equalities (3.7) and (3.8) we have

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\left(C_{ijmn}\varepsilon_{ij}\varepsilon_{mn}+2G_{mnij}\varepsilon_{ij}\gamma_{mn}+2F_{mnrij}\varepsilon_{ij}\kappa_{mnr}+\right.\\
&\quad+B_{ijmn}\gamma_{ij}\gamma_{mn}+A_{ijsmnr}\kappa_{ijs}\kappa_{mnr}+2D_{ijmnr}\gamma_{ij}\kappa_{mnr}+\\
&\quad+2B_{ij\alpha}\varepsilon_{ij}\omega_\alpha+2D_{ij\alpha}\gamma_{ij}\omega_\alpha+2F_{ijs\alpha}\kappa_{ijs}\omega_\alpha+\\
&\quad\left.+a\theta^2+\varrho\dot{u}_i\dot{u}_i+I_{kr}\dot{\varphi}_{jr}\dot{\varphi}_{jk}\right)=\\
&=\left[(\tau_{ij}+\sigma_{ij})\,\dot{u}_j+\mu_{ijs}\dot{\varphi}_{js}+\frac{1}{T_0}q_i\theta\right]_{,\,i}-\frac{1}{T_0}q_i\theta_{,\,i}+\\
&\quad+\left(B_{ij\alpha}\varepsilon_{ij}+D_{ij\alpha}\gamma_{ij}+F_{ijs\alpha}\kappa_{ijs}+G_\alpha\theta\right)\dot{\omega}_\alpha
\end{aligned} \tag{3.9}$$

Now, we integrate relation (3.9) over the domain B. By using the divergence theorem and the boundary conditions (3.5), we conclude that

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\int_B\left(C_{ijmn}\varepsilon_{ij}\varepsilon_{mn}+2G_{mnij}\varepsilon_{ij}\gamma_{mn}+2F_{mnrij}\varepsilon_{ij}\kappa_{mnr}+\right.\\
&B_{ijmn}\gamma_{ij}\gamma_{mn}+A_{ijsmnr}\kappa_{ijs}\kappa_{mnr}+2D_{ijmnr}\gamma_{ij}\kappa_{mnr}+\\
&2B_{ij\alpha}\varepsilon_{ij}\omega_\alpha+2D_{ij\alpha}\gamma_{ij}\omega_\alpha+2F_{ijs\alpha}\kappa_{ijs}\omega_\alpha+\\
&\left.+a\theta^2+\varrho\dot{u}_i\dot{u}_i+I_{kr}\dot{\varphi}_{jr}\dot{\varphi}_{jk}\right)dV=\\
&\int_B\left[\left(B_{ij\alpha}\varepsilon_{ij}+D_{ij\alpha}\gamma_{ij}+F_{ijs\alpha}\kappa_{ijs}+G_\alpha\theta\right)\dot{\omega}_\alpha-\frac{1}{T_0}q_i\theta_{,\,i}\right]dV.
\end{aligned} \tag{3.10}$$

Finally, we integrate the equality (20) from 0 to t and, by using the initial condition (3.4), we arrive at the desired result (3.6). □

3.2. Theorem. *Let $(u_i,\ \varphi_{ij},\ \theta,\ \omega_\alpha)$ be a solution of the problem P_0. Then there exists the positive constants m_1 and m_2 such that the following relation hold*

$$\begin{aligned}
&\int_B\left[\left(B_{ij\alpha}\varepsilon_{ij}+D_{ij\alpha}\gamma_{ij}+F_{ijs\alpha}\kappa_{ijs}+G_\alpha\theta\right)\dot{\omega}_\alpha-\frac{1}{T_0}q_i\theta_{,\,i}\right]dV\le\\
&\le -m_1\int_B\theta_{,\,i}\theta_{,\,j}dV+m_2\int_B\left(\varepsilon_{ij}\varepsilon_{ij}+\gamma_{ij}\gamma_{ij}+\kappa_{ijs}\kappa_{ijs}+\theta^2+\omega_\alpha\omega_\alpha\right)dV.
\end{aligned} \tag{3.11}$$

Proof. Taking into account the relations (2.6), (2.7) and $(2.3)_5$, we can write:

$$\begin{aligned}
&\int_B\left[\left(B_{ij\alpha}\varepsilon_{ij}+D_{ij\alpha}\gamma_{ij}+F_{ijs\alpha}\kappa_{ijs}+G_\alpha\theta\right)\dot{\omega}_\alpha-\frac{1}{T_0}q_i\theta_{,\,i}\right]dV=\\
&\int_B\left[\left(B_{ij\alpha}\varepsilon_{ij}+D_{ij\alpha}\gamma_{ij}+F_{ijs\alpha}\kappa_{ijs}+G_\alpha\theta\right)\left(g_{ij\alpha}\varepsilon_{ij}+h_{ij\alpha}\gamma_{ij}+\right.\right.\\
&\left.+l_{ijs\alpha}\kappa_{ijs}+p_\alpha\theta+q_{\alpha\beta}\omega_\beta+r_{i\alpha}\theta_{,\,i}\right)-\\
&\left.-\frac{1}{T_0}\left(a_{ijk}\varepsilon_{jk}+b_{ij}\gamma_{jk}+c_{ijsm}\kappa_{jsm}+d_i\theta+f_{i\alpha}\omega_\alpha+K_{ij}\theta_{,\,j}\right)\theta_{,\,i}\right]dV=\\
&-\int_B\frac{1}{T_0}K_{ij}\theta_{,\,i}\theta_{,\,j}dV+\int_B\left(\mathcal{B}_{ij}\varepsilon_{ij}\theta+\mathcal{D}_{ij}\gamma_{ij}\theta+\mathcal{F}_{ijs}\kappa_{ijs}\theta+\right.\\
&\mathcal{M}\theta^2+\mathcal{L}_\alpha\omega_\alpha\theta+\mathcal{D}_i\theta\theta_{,\,i}+\mathcal{C}_{ijmn}\varepsilon_{ij}\varepsilon_{mn}+\mathcal{D}_{ijmn}\varepsilon_{ij}\gamma_{mn}+\\
&\mathcal{F}_{ijmnr}\varepsilon_{ij}\kappa_{mnr}+\mathcal{B}_{ij\alpha}\varepsilon_{ij}\omega_\alpha+\mathcal{B}_{ijk}\varepsilon_{ij}\theta_{,\,k}+\mathcal{B}_{ijmn}\gamma_{ij}\gamma_{mn}+\\
&\mathcal{D}_{ijmnr}\gamma_{ij}\kappa_{mnr}+\mathcal{D}_{ij\alpha}\gamma_{ij}\omega_\alpha+\mathcal{D}_{ijk}\gamma_{ij}\theta_{,\,k}+\mathcal{A}_{ijsmnr}\kappa_{ijs}\kappa_{mnr}+\\
&\left.+\mathcal{F}_{ijs\alpha}\kappa_{ijs}\omega_\alpha+\mathcal{F}_{ijsm}\kappa_{ijs}\theta_{,\,m}+\mathcal{P}_{i\alpha}\omega_\alpha\theta_{,\,i}\right)dV,
\end{aligned} \tag{3.12}$$

where we have used the following notations

$$
\begin{aligned}
&\mathcal{A}_{ijsmnr} = \frac{1}{2}\left(F_{ijk\alpha}l_{mnr\alpha} + F_{mnr\alpha}l_{ijk\alpha}\right), \ \mathcal{C}_{ijmn} = \frac{1}{2}\left(B_{ij\alpha}g_{mn\alpha} + B_{mn\alpha}g_{ij\alpha}\right),\\
&\mathcal{B}_{ij} = B_{ij\alpha}p_\alpha + G_\alpha g_{ij\alpha}, \ \mathcal{B}_{ij\alpha} = B_{ij\beta}q_{\beta\alpha}, \ \mathcal{B}_{ijk} = B_{ij\alpha}\gamma_{k\alpha} - \frac{1}{T_0}a_{kji}\\
&\mathcal{D}_{ij} = D_{ij\alpha}p_\alpha + G_\alpha h_{ij\alpha}, \ \mathcal{D}_i = G_\alpha r_{i\alpha} - \frac{1}{T_0}d_i, \ \mathcal{D}_{ij\alpha} = D_{ij\beta}q_{\beta\alpha},\\
&\mathcal{D}_{ijk} = D_{ij\alpha}r_{k\alpha} - \frac{1}{T_0}b_{kij}, \ \mathcal{F}_{ijmnr} = D_{ij\alpha}l_{mnr\alpha} + F_{ijk\alpha}h_{mn\alpha},\\
&\mathcal{F}_{ijk} = G_\alpha l_{ijk\alpha} + F_{ijk\alpha}p_\alpha, \ \mathcal{F}_{ijk\alpha} = F_{ijk\beta}q_{\beta\alpha}, \ \mathcal{F}_{ijkm} = F_{ijk\alpha}r_{m\alpha} - \frac{1}{T_0}c_{mijk},\\
&\mathcal{D}_{ijmn} = B_{ij\alpha}h_{mn\alpha} + D_{mn\alpha}g_{ij\alpha}, \ \mathcal{L}_\alpha = G_\beta q_{\beta\alpha}, \ \mathcal{M} = G_\alpha p_\alpha, \ \mathcal{P}_{i\alpha} = -\frac{1}{T_0}f_{i\alpha}.
\end{aligned} \tag{3.13}
$$

By using the Schwarz's inequality and the arithmetic - geometric mean inequality

$$
ab \leq \frac{1}{2}\left(\frac{a^2}{\pi^2} + b^2\pi^2\right) \tag{3.14}
$$

to the last term in the relation (3.12), we are lead to

$$
\begin{aligned}
&\int_B \left[\left(B_{ij\alpha}\,\varepsilon_{ij} + D_{ij\alpha}\,\gamma_{ij} + F_{ijs\alpha}\,\kappa_{ijs} + G_\alpha\,\theta\right)\dot{\omega}_\alpha - \frac{1}{T_0}\,q_i\,\theta_{,\,i}\right]dV \leq\\
&\leq \left(-2\mu + \pi_1^2 + \pi_2^2 + \pi_3^2 + \pi_4^2 + \pi_5^2\right)\int_B \theta_{,\,i}\,\theta_{,\,i}dV +\\
&\left(\frac{M_2^2}{\pi_2^2} + M_6^2 + M_{11}^2 + M_{12}^2 + M_{13}^2 + M_{14}^2\right)\int_B \varepsilon_{ij}\,\varepsilon_{ij}dV +\\
&\left(\frac{M_3^2}{\pi_3^2} + M_7^2 + M_{15}^2 + M_{16}^2 + M_{17}^2 + 1\right)\int_B \gamma_{ij}\,\gamma_{ij}dV +\\
&\left(\frac{M_4^2}{\pi_4^2} + M_8^2 + M_{18}^2 + M_{19}^2 + 2\right)\int_B \kappa_{ijs}\,\kappa_{ijs}dV +\\
&\left(\frac{M_5^2}{\pi_5^2} + M_{10}^2 + 3\right)\int_B \omega_\alpha\,\omega_\alpha dV + \left(\frac{M_1^2}{\pi_1^2} + M_9^2 + 4\right)\int_B \theta^2 dV,
\end{aligned} \tag{3.15}
$$

where π_1, π_2, π_3, π_4 and π_5 are arbitrary positive constants. Also, in the inequality (3.15) we have used the notations

$$
\begin{aligned}
&M_1^2 = \max\left(\mathcal{D}_i\,\mathcal{D}_i\right)(x_s), \ M_2^2 = \max\left(\mathcal{B}_{ijk}\,\mathcal{B}_{ijk}\right)(x_s),\\
&M_3^2 = \max\left(\mathcal{D}_{ijk}\,\mathcal{D}_{ijk}\right)(x_s), \ M_4^2 = \max\left(\mathcal{F}_{ijkm}\,\mathcal{F}_{ijkm}\right)(x_s),\\
&M_5^2 = \max\left(\mathcal{P}_{i\alpha}\,\mathcal{P}_{i\alpha}\right)(x_s), \ M_6^2 = \max\left(\mathcal{B}_{ij}\,\mathcal{B}_{ij}\right)(x_s),\\
&M_7^2 = \max\left(\mathcal{D}_{ij}\,\mathcal{D}_{ij}\right)(x_s), \ M_8^2 = \max\left(\mathcal{F}_{ijk}\,\mathcal{F}_{ijk}\right)(x_s),\\
&M_9^2 = 2\ \max\left|\mathcal{M}\left(x_s\right)\right|, \ M_{10}^2 = \max\left(\mathcal{L}_\alpha\,\mathcal{L}_\alpha\right)(x_s),\\
&M_{11}^2 = 2\ \max\left[\left(\mathcal{C}_{ijmn}\mathcal{C}_{ijmn}\right)(x_s)\right]^{1/2}, \ M_{12}^2 = \max\left(\mathcal{D}_{ijmn}\,\mathcal{D}_{ijmn}\right)(x_s),\\
&M_{13}^2 = \max\left(\mathcal{D}_{ijmnr}\,\mathcal{D}_{ijmnr}\right)(x_s), \ M_{14}^2 = \max\left(\mathcal{B}_{ij\alpha}\,\mathcal{B}_{ij\alpha}\right)(x_s),\\
&M_{15}^2 = 2\ \max\left[\left(\mathcal{B}_{ijmn}\mathcal{B}_{ijmn}\right)(x_s)\right]^{1/2}, M_{16}^2 = \max\left(\mathcal{F}_{ijmnr}\,\mathcal{F}_{ijmnr}\right)(x_s),\\
&M_{17}^2 = \max\left(\mathcal{D}_{ij\alpha}\,\mathcal{D}_{ij\alpha}\right)(x_s), \ M_{18}^2 = 2\ \max\left[\left(\mathcal{A}_{ijkmnr}\mathcal{A}_{ijkmnr}\right)(x_s)\right]^{1/2},\\
&M_{19}^2 = \max\left(\mathcal{F}_{ijk\alpha}\,\mathcal{F}_{ijk\alpha}\right)(x_s).
\end{aligned} \tag{3.16}
$$

We choose the arbitrary constants π_1, π_2, π_3, π_4 and π_5 so that the quantity m_1 defined by

$$m_1 = \mu - \frac{1}{2}\left(\pi_1^2 + \pi_2^2 + \pi_3^2 + \pi_4^2 + \pi_5^2\right)$$

is strictly positive. Next, if we choose the constant m_2 as follows

$$\begin{aligned} m_2 = \frac{1}{2}\max\Big\{ & \frac{M_2^2}{\pi_2^2} + M_6^2 + M_{11}^2 + M_{12}^2 + M_{13}^2 + M_{14}^2, \\ & \frac{M_3^2}{\pi_3^2} + M_7^2 + M_{15}^2 + M_{16}^2 + M_{17}^2 + 1, \\ & \frac{M_4^2}{\pi_4^2} + M_8^2 + M_{18}^2 + M_{19}^2 + 2, \\ & \frac{M_5^2}{\pi_5^2} + M_{10}^2 + 3,\ \frac{M_1^2}{\pi_1^2} + M_9^2 + 4\Big\} \end{aligned}$$

then we arrive to the estimate (21) and this conclude the proof of Theorem 3.2. □

3.3. Theorem. *Let $(u_i,\ \varphi_{ij},\ \theta,\ \omega_\alpha)$ be a solution of the problem P_0 and suppose that the assumptions (i) - (v) are satisfied. Then there exists a positive constant m_3 such that we have the following inequality*

$$\begin{aligned} &\int_B \left(\dot{u}_i\ \dot{u}_i + \dot{\varphi}_{ij}\ \dot{\varphi}_{ij} + \varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijk}\kappa_{ijk} + \theta^2 + \omega_\alpha\omega_\alpha\right) dV \le \\ &m_3 \int_0^t \int_B \left(\dot{u}_i\ \dot{u}_i + \dot{\varphi}_{ij}\ \dot{\varphi}_{ij} + \varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijk}\kappa_{ijk} + \theta^2 + \omega_\alpha\omega_\alpha\right) dV ds \end{aligned} \tag{3.17}$$

for any $t \in [0, t_0]$.

Proof. First, taking into account the hypotheses (i) - (v), we have

$$\begin{aligned} &m_0 \int_B \left(\dot{u}_i\ \dot{u}_i + \dot{\varphi}_{ij}\ \dot{\varphi}_{ij} + \varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2\right) dV \le \\ &\int_B (C_{ijmn}\varepsilon_{ij}\varepsilon_{mn} + B_{ijmn}\gamma_{ij}\gamma_{mn} + A_{ijsmnr}\kappa_{ijs}\kappa_{mnr} + \\ &a\ \theta^2 + \varrho\ \dot{u}_i\ \dot{u}_i + I_{kr}\ \dot{\varphi}_{jr}\ \dot{\varphi}_{jk}\big) dV, \end{aligned} \tag{3.18}$$

where we have used the notation

$$m_0 = \min\{\varrho,\ a,\ \lambda_1,\ \lambda_2,\ \lambda_3,\ \lambda_4\}$$

Next, we use the Schwarz's inequality and the arithmetic - geometric mean inequality (3.14) to the left side of the relation (3.18). So, we are lead to the inequality

$$\begin{aligned} m_0 \int_B &\left(\dot{u}_i\ \dot{u}_i + \dot{\varphi}_{ij}\ \dot{\varphi}_{ij} + \varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2\right) dV \le \\ &\le \left(\pi_6^2 + N_4^2 + N_5^2\right)\int_B \varepsilon_{ij}\varepsilon_{ij} dV + \left(\pi_7^2 + N_6^2 + 2\right)\int_B \gamma_{ij}\gamma_{ij} dV + \\ &+ \left(\pi_8^2 + 3\right)\int_B \kappa_{ijs}\kappa_{ijs} dV + \left(\frac{N_1^2}{\pi_6^2} + \frac{N_2^2}{\pi_7^2} + \frac{N_3^2}{\pi_8^2}\right)\int_B \omega_\alpha\omega_\alpha dV - \\ &+ m_2 \int_0^t \int_B \left(\varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2 + \omega_\alpha\omega_\alpha\right) dV ds \\ &- m_1 \int_0^t \int_B \theta,_i\theta,_i dV ds \end{aligned} \tag{3.19}$$

where $t \in [0, t_0]$.
In this inequality we have used the notations

$$\begin{aligned}
&N_1^2 = \max\left(B_{ij\alpha}\ B_{ij\alpha}\right)(x_s)\,,\ N_2^2 = \max\left(D_{ij\alpha}\ D_{ij\alpha}\right)(x_s)\,,\\
&N_3^2 = \max\left(F_{ijk\alpha}\ F_{ijk\alpha}\right)(x_s)\,,\ N_4^2 = \max\left(G_{mnij}\ G_{mnij}\right)(x_s)\,,\\
&N_5^2 = \max\left(F_{mnrij}\ F_{mnrij}\right)(x_s)\,,\ N_6^2 = \max\left(D_{ijmnr}\ D_{ijmnr}\right)(x_s)\,,
\end{aligned} \tag{3.20}$$

where $(x_s) \in \overline{B}$.

On the other hand, by using the initial conditions (3.4) and the consitutive relation (2.7), we arrive to the conclusion that:

$$\begin{aligned}
\int_B \omega_\alpha \omega_\alpha dV &= \int_0^t \frac{d}{ds}\left(\int_B \omega_\alpha \omega_\alpha dV\right) ds = 2\int_0^t \left(\int_B \omega_\alpha \dot{\omega}_\alpha dV\right) ds = \\
&= 2\int_0^t \int_B (g_{ij\alpha}\varepsilon_{ij}\omega_\alpha + h_{ij\alpha}\gamma_{ij}\omega_\alpha + l_{ijs\alpha}\kappa_{ijs}\omega_\alpha + \\
&\qquad + p_\alpha \theta \omega_\alpha + q_{\alpha\beta}\omega_\alpha\omega_\beta + r_i \omega_\alpha \theta_{,\,i})\, dV ds
\end{aligned} \tag{3.21}$$

Now, by using, again, the Schwarz's inequality and the arithmetic - geometric mean inequality (3.14) to the right side of the relation (3.21). So, we deduce that for an arbitrary positive constant π_9 the following inequality hold:

$$\begin{aligned}
\int_B \omega_\alpha\ \omega_\alpha\ dV &\le \pi_9^2 \int_0^t \int_B \theta_{,\,i}\ \theta_{,\,i}\ dV ds + \\
&+ \left(\frac{Q_1^2}{\pi_9^2} + Q_5^2 + Q_6^2 + 3\right) \int_0^t \int_B \omega_\alpha\ \omega_\alpha\ dV ds + \\
&+ Q_2^2 \int_0^t \int_B \varepsilon_{ij}\ \varepsilon_{ij}\ dV ds + Q_3^2 \int_0^t \int_B \gamma_{ij}\ \gamma_{ij}\ dV ds + \\
&+ Q_4^2 \int_0^t \int_B \kappa_{ijs}\ \kappa_{ijs}\ dV ds + \int_0^t \int_B \theta^2 dV ds
\end{aligned} \tag{3.22}$$

where $t \in [0, t_0]$.
In this inequality we have used the notations

$$\begin{aligned}
&Q_1^2 = \max\left(r_{i\alpha}\ r_{i\alpha}\right)(x_s)\,,\ Q_2^2 = \max\left(g_{ij\alpha}\ g_{ij\alpha}\right)(x_s)\,,\\
&Q_3^2 = \max\left(h_{ij\alpha}\ h_{ij\alpha}\right)(x_s)\,,\ Q_4^2 = \max\left(l_{ijk\alpha}\ l_{ijk\alpha}\right)(x_s)\,,\\
&Q_5^2 = \max\left(p_\alpha\ p_\alpha\right)(x_s)\,,\ Q_6^2 = \max\left[\left(q_{i\alpha}\ q_{i\alpha}\right)(x_s)\right]^{1/2},
\end{aligned} \tag{3.23}$$

where $(x_s) \in \overline{B}$.
If we denote by m_4 the quantity

$$m_4 = \max\left\{\frac{Q_1^2}{\pi_9^2} + Q_5^2 + Q_6^2 + 3,\ Q_2^2,\ Q_3^2,\ Q_4^2,\ 1\right\},$$

then, from (3.21) we obtain the following inequality

$$\begin{aligned}
\int_B \omega_\alpha\ \omega_\alpha\ dV &\le \pi_9^2 \pi_{10}^2 \int_0^t \int_B \theta_{,\,i}\ \theta_{,\,i}\ dV ds + \\
&+ m_4 \pi_{10}^2 \int_0^t \int_B \left(\varepsilon_{ij}\varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2 + \omega_\alpha\ \omega_\alpha\right) dV ds
\end{aligned} \tag{3.24}$$

which is satisfied for an arbitrary positive constant π_{10}.
From (3.19) and (3.24) we obtain

$$\begin{aligned} & m_0 \int_B \left(\dot{u}_i \dot{u}_i + \dot{\varphi}_{ij} \dot{\varphi}_{ij} + \theta^2\right) dV + \left[m_0 - \left(\pi_6^2 + N_4^2 + N_5^2\right)\right] \int_B \varepsilon_{ij}\, \varepsilon_{ij}\, dV + \\ & + \left(m_0 - \pi_7^2 - N_6^2 - 2\right) \int_B \gamma_{ij}\, \gamma_{ij}\, dV + \left(m_0 - \pi_8^2 - 3\right) \int_B \kappa_{ijs}\, \kappa_{ijs}\, dV + \\ & + \left(\pi_{10}^2 - \frac{N_1^2}{\pi_6^2} - \frac{N_2^2}{\pi_7^2} - \frac{N_3^2}{\pi_8^2}\right) \int_B \omega_\alpha \omega_\alpha dV \leq \left(m_1 - \pi_9^2 - \pi_{10}^2\right) \int_0^t \int_B \theta,_i \theta,_i dV ds + \\ & + \left(m_2 + m_4 \pi_{10}^2\right) \int_0^t \int_B \left(\varepsilon_{ij}\varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2 + \omega_\alpha\, \omega_\alpha\right) dV ds \end{aligned} \tag{3.25}$$

We choose the arbitrary constants π_6, π_7, π_8, π_9 and π_{10} so that

$$m_5 \equiv m_0 - \pi_6^2 - N_4^2 - N_5^2 > 0,\ m_6 \equiv m_0 - \pi_7^2 - N_6^2 - 2 > 0,$$
$$m_7 \equiv m_0 - \pi_8^2 - 3 > 0,\ m_8 \equiv \pi_{10}^2 - \frac{N_1^2}{\pi_6^2} - \frac{N_2^2}{\pi_7^2} - \frac{N_3^2}{\pi_8^2} > 0,$$
$$m_9 \equiv m_1 - \pi_9^2\ \pi_{10}^2 > 0,$$

and thus we are lead to

$$\begin{aligned} & \left(m_2 + m_4 \pi_{10}^2\right) \int_0^t \int_B \left(\varepsilon_{ij}\varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2 + \omega_\alpha\, \omega_\alpha\right) dV ds \geq \\ & \geq m_0 \int_B \left(\dot{u}_i \dot{u}_i + \dot{\varphi}_{ij} \dot{\varphi}_{ij} + \theta^2\right) dV + m_5 \int_B \varepsilon_{ij}\varepsilon_{ij} dV + m_6 \int_B \gamma_{ij}\gamma_{ij} dV + \\ & + m_7 \int_B \kappa_{ijs}\kappa_{ijs} dV + m_8 \int_B \omega_\alpha\, \omega_\alpha dV + m_9 \int_B \theta,_i \theta,_i dV dV \geq \\ & \geq m_{10} \int_B \left(\dot{u}_i \dot{u}_i + \dot{\varphi}_{ij} \dot{\varphi}_{ij} + \varepsilon_{ij}\varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijs}\kappa_{ijs} + \theta^2 + \omega_\alpha\, \omega_\alpha\right) dV, \end{aligned} \tag{3.26}$$

where the signification of the constant m_{10} is

$$m_{10} = \min\left\{m_0,\ m_5,\ m_6,\ m_7,\ m_8\right\}.$$

It is easy to observe that

$$\begin{aligned} & \int_0^t \int_B \left(\dot{u}_i \dot{u}_i + \dot{\varphi}_{ij} \dot{\varphi}_{ij} + \varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\ \gamma_{ij} + \kappa_{ijs}\ \kappa_{ijs} + \theta^2 + \omega_\alpha\, \omega_\alpha\right) dV ds \geq \\ & \geq \int_0^t \int_B \left(\varepsilon_{ij}\ \varepsilon_{ij} + \gamma_{ij}\ \gamma_{ij} + \kappa_{ijs}\ \kappa_{ijs} + \theta^2 + \omega_\alpha\, \omega_\alpha\right) dV ds \end{aligned} \tag{3.27}$$

Finally, if we choose

$$m_3 = \frac{\left(m_2 + m_4 \pi_{10}^2\right)}{m_{10}}$$

then from (3.26) and (3.27) we arrive at the desired result (3.17) and Theorem 3.3 is proved. □

Theorem 3.1, Theorem 3.2 and Theorem 3.3 form the basis of the main result of this study: the uniqueness of mixed initial-boundary value problem for thermoelastic dipolar body with internal state variables.

3.4. Theorem. *Assume that the hypotheses (i) - (v) hold. Then there exists at most one solution of the problem defined by the equations (2.1), (2.2) and (2.6) with the initial conditions (2.9) and the boundary conditions (2.10).*

Proof. Suppose that the mixed problem has two solutions. Then the difference of these solutions is solution for the above mentioned problem P_0. For our aim it is suffice to show that the function $y(t)$ defined by

$$y(t) = \int_B \left(\dot{u}_i\dot{u}_i + \dot{\varphi}_{ij}\dot{\varphi}_{ij} + \varepsilon_{ij}\varepsilon_{ij} + \gamma_{ij}\gamma_{ij} + \kappa_{ijr}\kappa_{ijr} + \theta^2 + \omega_\alpha\omega_\alpha\right) dV$$

vanishes on $[0, t_0]$.

If we assume the contrary, i.e. $y(t) \neq 0$, this is absurdum because the inequality (3.17) and Gronwall's inequality imply that $y(t) \equiv 0$ on $[0, t_0]$ and Theorem 3.4 is concluded. □

Conclusion. The existence of internal state variables do not affect the uniqueness of solution of the mixed problem for dipolar thermoelastic materials.

Acknowledgement

We express our gratitude to the referees for their criticism of the manuscript and for helpful suggestions.

References

[1] Anand, L. and Gurtin, M. E. *A theory of amorphous solids undergoing large deformations*, Int. J. Solids Struct. **40**, 1465–1487, 2003

[2] Bouvard, J. L., Ward, D. K., Hossain, D., Marin, E. B., Bammann, D. J. and Horstemeyer M. F., *A general inelastic internal state variable model for amorphous glassy polymers*, Acta Mechanica, **213** 1–2, 71-96, 2010

[3] Chirita, S. *On the linear theory of thermo-viscoelastic materials with internal state variables*, Arch. Mech., **33**, 455–464, 1982

[4] Marin, M. *An evolutionary equation in thermoelasticity of dipolar bodies*, Journal of Mathematical Physics, **40** 3, 1391–1399, 1999

[5] Marin, M. *A partition of energy in thermoelsticity of microstretch bodies*, Nonlinear Analysis: RWA, **11** 4, 2436–2447, 2010

[6] Marin, M. *Some estimates on vibrations in thermoelasticity of dipolar bodies*, Journal of Vibration and Control, **16** 1, 33–47, 2010

[7] Marin, M. *Lagrange identity method for microstretch thermoelastic materials* J. Mathematical Analysis and Applications, **363** 1, 275–286, 2010

[8] Marin, M., Agarwal, R. P. and Mahmoud, S. R. *Modeling a microstretch thermoelastic body with two temperature*, Abstract and Applied Analysis, **2013**, 1–7, 2013

[9] Nachlinger, R. R. and Nunziato, J. W. *Wave propagation and uniqueness theorem for elastic materials with ISV*, Int. J. Engng. Sci., **14**, 31-38, 1976

[10] Pop, N., *An algorithm for solving nonsmooth variational inequalities arising in frictional quasistatic contact problems*, Carpathian Journal of Mathematics, **24** 1, 110–119, 2008

[11] Pop, N., Cioban, H. and Horvat-Marc, A., *Finite element method used in contact problems with dry friction*, Computational Materials Science, **50** 4, 1283–1285, 2011

[12] Sherburn, J. A., Horstemeyer, M. F., Bammann, D. J. and Baumgardner, R. R. *Application of the Bammann inelasticity internal state variable constitutive model to geological materials*, Geophysical J. Int., **184** 3, 1023–1036, 2011

[13] Solanki, K. N. and Bammann, D. J. *A thermodynamic framework for a gradient theory of continuum damage*, (American Acad.Mech.Conf., New Orleans, 2008).

[14] Wei, C. and Dewoolkar, M. M. *Formulation of capillary hysteresis with internal state variables*, Water Resources Research, **42**, 16 pp., 2006

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 27 – 41

ON A FUNCTIONAL EQUATION ORIGINATING FROM A MIXED ADDITIVE AND CUBIC EQUATION AND ITS STABILITY

Mohammad Janfada [a] [*], Tayebe Laal Shateri [b] [†], and Rahele Shourvarzi [c] [‡]

Received 07 : 08 : 2011 : Accepted 25 : 12 : 2012

Abstract

In this paper, we study solutions of the 2-variable mixed additive and cubic functional equation

$$\begin{aligned} f(2x+y, 2z+t) + f(2x-y, 2z-t) &= 2f(x+y, z+t) \\ &\quad + 2f(x-y, z-t) + 2f(2x, 2z) - 4f(x, z), \end{aligned}$$

which has the cubic form $f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$ as a solution. Also the Hyers–Ulam–Rassias stability of this equation in the non-Archimedean Banach spaces is investigated.

Keywords: Hyers–Ulam–Rassias stability, Cubic functional equation, Non–Archimedean normed space, Derivation.

2000 AMS Classification: 39B22, 39B82, 46S10

1. Introduction and preliminaries

The study of stability problems for functional equations is related to a question of Ulam [26] concerning the stability of group homomorphisms, affirmatively answered for Banach spaces by Hyers [8]. Subsequently, the result of Hyers was generalized by Aoki [1], Bourgin [5] and Rassias [24].
During the last decades several stability problems for various functional equations have been investigated by numerous mathematicians. We refer the reader to the survey articles [7, 9, 23] and monographs [6, 10, 13, 22] and references therein.

---

[a]Department of Mathematics, Ferdowsi University of Mashhad, Mashhad, P. O. Box 1159-91775 Iran.
[*]Email: mjanfada@gmail.com
[b]Facaulty of Mathematics and computer sciences, Hakim Sabzevari University, Sabzevar, Iran.
[†]Email: t.shateri@gmail.com
[c]Department of Mathematics, Hakim Sabzevari University, Sabzevar, Iran.
[‡]Email: rshurvarzy@gmail.com

Let X and Y be real vector spaces. For a mapping $f : X \times X \to Y$, consider the following 2-variable mixed additive and cubic functional equation:

$$\begin{aligned} f(2x+y,2z+t)+f(2x-y,2z-t) &= 2f(x+y,z+t)+2f(x-y,z-t) \\ &\quad + 2f(2x,2z)-4f(x,z) \end{aligned} \tag{1.1}$$

One can see that the cubic form $f(x,y) = ax^3 + bx^2y + cxy^2 + dy^3$ is a solution of (1.1), when $X = Y = \mathbb{R}$.

The one variable cubic equation

$$f(2x+y)+f(2x-y) = 2f(x+y)+2f(x-y)+2f(2x)-4f(x) \tag{1.2}$$

is considered in [17] and the general solutions of this equation and its Hyers–Ulam–Rassias stability in quasi-Banach spaces is studied.

Several-variable functional equations and their stability have been studied in many papers (see, for example, [3, 4], [12, 11], [14], [18, 19], [20, 21], [25]).

In this paper first we study solutions of (1.1) and its relations with (1.2) and then the Hyers–Ulam–Rassias stability of (1.1) in non-Archimedean Banach spaces is investigated.

By a *non-Archimedean field* we mean a field $\mathbb{K}$ equipped with a function (valuation) $|\cdot|$ from $\mathbb{K}$ into $[0,\infty)$ such that $|r| = 0$ if and only if $r = 0$, $|rs| = |r|\,|s|$, and $|r+s| \leq \max\{|r|,|s|\}$ for all $r, s \in \mathbb{K}$. Clearly $|1| = |-1| = 1$ and $|n| \leq 1$ for all $n \in \mathbb{N}$. By the trivial valuation we mean the mapping $|\cdot|$ taking everything but 0 into 1 and $|0| = 0$. Let X be a vector space over a field $\mathbb{K}$ with a non-Archimedean non-trivial valuation $|\cdot|$. A function $\|\cdot\| : X \to [0,\infty)$ is called a *non-Archimedean norm* if it satisfies the following conditions:

(i) $\|x\| = 0$ if and only if $x = 0$;

(ii) for any $r \in \mathbb{K}, x \in X$, $\|rx\| = |r|\|x\|$;

(iii) the strong triangle inequality (ultrametric); namely,

$$\|x+y\| \leq \max\{\|x\|,\|y\|\} \quad (x, y \in X).$$

Then $(X, \|\cdot\|)$ is called a non-Archimedean normed space. Due to the fact that

$$\|x_n - x_m\| \leq \max\{\|x_{j+1} - x_j\| : m \leq j \leq n-1\} \qquad (n > m),$$

a sequence $\{x_n\}$ is Cauchy if and only if $\{x_{n+1} - x_n\}$ converges to zero in a non-Archimedean normed space. By a complete non-Archimedean normed space we mean one in which every Cauchy sequence is convergent.

The stability problem in non-Archimedean normed spaces has been studied by many authors. In [2], the stability of approximate additive mappings $f : \mathbb{Q}_p \to \mathbb{R}$ is investigated. In [15, 16], the authors investigated the stability of Cauchy, quadratic and cubic functional equations, in the context of non-Archimedean normed spaces.

We need the following lemmas from [17] for our stability results.

1.1. Lemma. *If a mapping $f : X \to Y$ with $f(0) = 0$ satisfies (1.2), then the mapping $g : X \to Y$ defined by $g(x) = f(2x) - 8f(x)$ is additive.*

1.2. Lemma. *If a mapping $f : X \to Y$ with $f(0) = 0$ satisfies (1.2), then the mapping $h : X \to Y$ defined by $h(x) = f(2x) - 2f(x)$ is cubic.*

2. Relations between (1.2) and (1.1)

In this section we show that equations (1.2) and (1.1) are closely related and so by knowing the solutions of (1.2), we may find solutions of (1.1). Next some useful examples are considered.

2.1. Theorem. *Suppose $f : X \times X \to Y$ is a mapping satisfying (1.1), then $g : X \to Y$ defined by $g(x) := f(x, x)$ satisfies (1.2).*

Proof. From (1.1) and definition of g,

$$\begin{aligned} g(2x+y)+g(2x-y) &= f(2x+y,2x+y)+f(2x-y,2x-y) \\ &= 2f(x+y,x+y)+2f(x-y,x-y)+2f(2x,2x)-4f(x,x) \\ &= 2g(x+y)+2g(x-y)+2g(2x)-4g(x). \end{aligned}$$

□

2.2. Theorem. *Let $a, b, c, d \in \mathbb{R}$ and $g : X \to Y$ be a mapping satisfying (1.2). If $f : X \times X \to Y$ is defined by*

$$f(x,y) = (a-\frac{c}{3})g(x) + (\frac{c+b}{6})g(x+y) + (\frac{c-b}{6})g(x-y) + (d-\frac{b}{3})g(y), \tag{2.1}$$

then f satisfies (1.1). Furthermore if $f(0,0) = 0$, $a + d = 1$ and $c = -b$, then $g(x) = f(x, x)$.

Proof. We have

$$\begin{aligned} f(2x+y,2z+t)+f(2x-y,2z-t) &= (a-\frac{c}{3})g(2x+y)+(\frac{c+b}{6})g(2x+y+2z+t) \\ &+ (\frac{c-b}{6})g(2x+y-(2z+t))+(d-\frac{b}{3})g(2z+t) \\ &+ (a-\frac{c}{3})g(2x-y)+(\frac{c+b}{6})g(2x-y+2z-t) \\ &+ (\frac{c-b}{6})g(2x-y-(2z-t))+(d-\frac{b}{3})g(2z-t) \\ &= (a-\frac{c}{3})\big[g(2x+y)+g(2x-y)\big] \\ &+ (\frac{c+b}{6})\big[g(2x+2zy+t)+g(2x+2z-(y+t))\big] \\ &+ (\frac{c-b}{6})\big[g(2x-2z+y-t)+g(2x-2z-(y-t))\big] \\ &+ (d-\frac{b}{3})\big[g(2z+t)+g(2z-t)\big]. \end{aligned} \tag{2.2}$$

On the other hand

$$\begin{aligned} &\quad 2f(x+y,z+t)+2f(x-y,z-t)+2f(2x,2z)-4f(x,z) \\ &= (a-\frac{c}{3})\big[2g(x+y)+2g(x-y)+2g(2x)-4g(x)\big] \\ &+ (\frac{c+b}{6})\big[2g(x+y+z+t)+2g(x+z-(y+t))+2g(2x+2z)-4g(x+z)\big] \\ &+ (\frac{c-b}{6})\big[2g(x-z+y-t)+2g(x-z-(y-t))+2g(2x-2z)-4g(x-z)\big] \\ &+ (d-\frac{b}{3})\big[2g(z+t)+2g(z-t)+2g(2z)-4g(z)\big] \\ &= (a-\frac{c}{3})\big[g(2x+y)+g(2x-y)\big] \\ &+ (\frac{c+b}{6})\big[g(2x+2zy+t)+g(2x+2z-(y+t))\big] \\ &+ (\frac{c-b}{6})\big[g(2x-2z+y-t)+g(2x-2z-(y-t))\big] \\ &+ (d-\frac{b}{3})\big[g(2z+t)+g(2z-t)\big]. \end{aligned} \tag{2.3}$$

Thus (2.2) and (2.3) imply that f satisfies (1.1). □

For the following example, we recall that a mapping D from an algebra X into itself is called *derivation* if, for any $x, y \in X$, $D(xy) = D(x)y + xD(y)$.

2.3. Example. Let X be a real algebra and let D_1 be a derivation on X. Suppose $D_2 : X \to X$ satisfies

$$D_2(xy) = D_2(x)y + D_1(x)D_1(y) + xD_2(y).$$

Now define $f : X \times X \to X$ by $f(x, y) = D_2(xy)$, then f satisfies (1.1). Also $g : X \to X$ defined by $g(x) = D_2(x^2)$ satisfies (1.2).

2.4. Example. Let M_n be the algebra of $n \times n$-real matrices. Define the mapping $g : M_n \to M_n$ by $g(A) = A^3$, $A \in M_n$, then one can easily see that g satisfies (1.2). For $a, b, c, d \in \mathbb{R}$, set

$$f(A, B) = aA^3 + \frac{2b}{3}A^2oB + \frac{2c}{3}AoB^2 + \frac{b}{3}ABA + \frac{b}{3}BAB + dBAB,$$

where AoB is the Jordan product $\frac{1}{2}(AB + BA)$ of A and B, for any $A, B \in M_n$. Then f satisfies (2.1). So by Theorem 2.2, f satisfies (1.1).

3. Stability of Eq. (1.2)

Throughout this section, assume that X is a vector space and that Y is a non-Archimedean Banach space. In this section, we study some stability results from [17] in non-Archimedean Banach spaces. Indeed, we consider the stability of functional equation (1.1), and the fact the $X \times X$ with the point-wise operations is also a vector space implies a similar stability result for (1.2). For convenience, we use the following abbreviation for a given mapping $f : X \to Y$,

$$Df(x, y) := f(2x + y) + f(2x - y) - 2f(x + y) - 2f(x - y) - 2f(2x) + 4f(x)$$

for all $x, y \in X$.

3.1. Theorem. *Let $\varphi_a : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty} \frac{1}{|2|^n}\varphi_a(2^n x, 2^n y) = 0 \tag{3.1}$$

$$M_a(x, y) := \lim_{n\to\infty} \max\{\frac{1}{|2|^i}\varphi_a(2^i x, 2^i y) :\ 0 \le i < n\} < \infty \tag{3.2}$$

$$\lim_{t\to\infty} \lim_{n\to\infty} \max\{\frac{1}{|2|^i}\varphi_a(2^i x, 2^i y) : t \le i < t + n\} = 0 \tag{3.3}$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x, y)\|_Y \le \varphi_a(x, y) \tag{3.4}$$

for all $x, y \in X$. Then the limit

$$A(x) = \lim_{n\to\infty} \frac{1}{2^n}[f(2^{n+1}x) - 8f(2^n x)]$$

exists, for all $x \in X$, and the mapping $A : X \to Y$ is the unique additive mapping satisfying

$$\|f(2x) - 8f(x) - A(x)\|_Y \le \frac{1}{|2|}\tilde{\varphi}_a(x), \tag{3.5}$$

for all $x \in X$, where

$$\tilde{\varphi}_a(x) := \max\{|2|M_a(x, x), |2|M_a(0, x), M_a(x, 2x)\}$$

Proof. Letting $x = 0$ in (3.4), we get

$$\|f(y) + f(-y)\|_Y \le \varphi_a(0, y) \tag{3.6}$$

for all $y \in X$. Replacing y by x and $2x$ in (3.4), respectively, we get the following inequalities

$$\|f(3x) - 4f(2x) + 5f(x)\|_Y \le \varphi_a(x, x) \tag{3.7}$$

$$\|f(4x) - 2f(3x) - 2f(2x) - 2f(-x) + 4f(x)\|_Y \le \varphi_a(x, 2x) \tag{3.8}$$

for all $x \in X$. It follows from (3.6)-(3.8) that for any $x \in X$,

$$\|f(4x) - 10f(2x) + 16f(x)\|_Y \le \max\{|2|\varphi_a(x, x), |2|\varphi_a(0, x), \varphi_a(x, 2x)\}. \tag{3.9}$$

Let $g : X \to Y$ be a mapping defined by $g(x) := f(2x) - 8f(x)$ and let

$$\psi_a(x) := \max\{|2|\varphi_a(x, x), |2|\varphi_a(0, x), \varphi_a(x, 2x)\},$$

for all $x \in X$. Therefore (3.9) means

$$\|g(2x) - 2g(x)\|_Y \le \psi_a(x), \tag{3.10}$$

for all $x \in X$. By relations (3.1)-(3.3) we infer that for all $x \in X$,

$$\max\{\frac{\psi_a(2^i x)}{|2|^i} : \ 0 \le i < n\} < \infty, \qquad \lim_{n\to\infty} \frac{1}{|2|^n}\psi_a(2^n x) = 0. \tag{3.11}$$

Replacing x by $2^n x$ in (3.10) and dividing both sides (3.10) by $|2|^{n+1}$ we get

$$\|\frac{1}{2^{n+1}}g(2^{n+1}x) - \frac{1}{2^n}g(2^n x)\|_Y \le \frac{1}{|2|^{n+1}}\psi_a(2^n x) \tag{3.12}$$

for all $x \in X$ and all non-negative integer n, and so for any $x \in X$ and every non-negative integers n and m with $n \ge m$,

$$\|\frac{1}{2^n}g(2^n x) - \frac{1}{2^m}g(2^m x)\|_Y \le \frac{1}{|2|}\max\{\frac{\psi_a(2^i x)}{|2|^i} : m \le i < n\}. \tag{3.13}$$

Therefore we conclude from (3.11) and (3.12) that the sequence $\{\frac{1}{2^n}g(2^n x)\}$ is a Cauchy sequence in Y, for all $x \in X$. The sequence $\{\frac{1}{2^n}g(2^n x)\}$ converges in Y for any $x \in X$, since Y is complete. So one can define the mapping $A : X \to Y$ by

$$A(x) := \lim_{n\to\infty} \frac{1}{2^n}g(2^n x) \tag{3.14}$$

for all $x \in X$. Letting $m = 0$ and passing to the limit when $n \to \infty$ in (3.13), we get (3.5). Now we show that A is an additive mapping. It follows from (3.11), (3.12) and (3.14) that

$$\begin{aligned}\|A(2x) - 2A(x)\|_Y &= \lim_{n\to\infty} \|\frac{1}{2^n}g(2^{n+1}x) - \frac{1}{2^{n-1}}g(2^n x)\|_Y \\ &= |2| \lim_{n\to\infty} \|\frac{1}{2^{n+1}}g(2^{n+1}x) - \frac{1}{2^n}g(2^n x)\|_Y \\ &\le \lim_{n\to\infty} \frac{1}{|2|^n}\psi_a(2^n x),\end{aligned}$$

for all $x \in X$. So

$$A(2x) = 2A(x) \tag{3.15}$$

for all $x \in X$. On the other hand it follows from (3.1), (3.4) and (3.14) that

$$\begin{aligned}\|DA(x,y)\|_Y &= \lim_{n\to\infty} \frac{1}{|2|^n}\|Dg(2^n x, 2^n y)\|_Y \\ &= \lim_{n\to\infty} \frac{1}{|2|^n}\{\|Df(2^{n+1}x, 2^{n+1}y) - 8Df(2^n x, 2^n y)\|_Y\} \\ &\le \lim_{n\to\infty} \frac{1}{|2|^n}\max\{\|Df(2^{n+1}x, 2^{n+1}y)\|_Y, |8|\|Df(2^n x, 2^n y)\|_Y\} \\ &\le \lim_{n\to\infty} \frac{1}{|2|^n}\max\{\varphi_a(2^{n+1}x, 2^{n+1}y), |8|\varphi_a(2^n x, 2^n y)\} = 0,\end{aligned}$$

for all $x, y \in X$. Hence the mapping A satisfies (1.2). So by Lemma 1.1, the mapping $x \mapsto A(2x) - 8A(x)$ is additive. Therefore (3.15) implies that A is additive. To prove the uniqueness of A, let $T : X \to Y$ be another additive mapping satisfying (3.5). So it follows from (3.5), (3.14) and (3.3) that

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{|2|^t}\tilde{\varphi}_a(2^{tx}) &= \lim_{t\to\infty}\max\{|2|\frac{M_a(2^t x, 2^t x)}{|2|^t}, |2|\frac{M_a(0, 2^t x)}{|2|^t}, \frac{M_a(2^t x, 2^t 2x)}{|2|^t}\} \\ &= \lim_{t\to\infty}\max\{\lim_{n\to\infty}\max\{|2|\frac{\varphi_a(2^{i+t}x, 2^{i+t}x)}{|2|^{t+i)}} :\ 0 \le i < n\}, \\ &\quad \lim_{n\to\infty}\max\{|2|\frac{\varphi_a(0, 2^{i+t}x)}{|2|^{t+i}} 0 \le i < n\}, \\ &\quad \lim_{n\to\infty}\max\{\frac{\varphi_a(2^{i+t}x, 2^{i+t+1}x)}{|2|^{t+i)}} : 0 \le i < n\}\} \\ &= \lim_{t\to\infty}\lim_{n\to\infty}\{|2|\frac{\varphi_a(2^i x, 2^i x)}{|2|^i}, |2|\frac{\varphi_a(0, 2^i x)}{|2|^i}, \frac{\varphi_a(2^i x, 2^{i+1}x)}{|2|^i} :\ t \le i < t+n\} \\ &= 0.\end{aligned}$$

Hence it follows

$$\begin{aligned}\|A(x) - T(x)\|_Y &= \lim_{t\to\infty}\frac{1}{|2|^t}\|g(2^t x) - T(2^t x)\|_Y \\ &\le \frac{1}{|2|}\lim_{t\to\infty}\frac{1}{|2|^t}\tilde{\varphi}_a(2^{tx}) = 0,\end{aligned}$$

for all $x \in X$. So $A = T$ □

3.2. Theorem. *Let $\varphi_a : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty}|2|^n\varphi_a(\frac{x}{2^n}, \frac{y}{2^n}) = 0,$$

$$M_a(x,y) = \lim_{n\to\infty}\max\{|2|^i\varphi_a(\frac{x}{2^i}, \frac{y}{2^i}) :\quad 1 \le i < n\} < \infty, \tag{3.16}$$

$$\lim_{t\to\infty}\lim_{n\to\infty}\max\{|2|^i\varphi_a(\frac{x}{2^i}, \frac{y}{2^i}) :\quad t+1 \le i < t+n\} = 0, \tag{3.17}$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x,y)\|_Y \le \varphi_a(x,y),$$

for all $x, y \in X$. Then the limit

$$A(x) = \lim_{n\to\infty}2^n[f(\frac{x}{2^{n-1}}) - 8f(\frac{x}{2^n})]$$

exists for all $x \in X$ and the mapping $A : X \to Y$ is the unique additive mapping satisfying

$$\|f(2x) - 8f(x) - A(x)\|_Y \le \frac{1}{|2|}\tilde{\varphi}_a(x), \tag{3.18}$$

for all $x \in X$*, where*

$$\tilde{\varphi}_a(x) := \max\{|2|M_a(x,x), |2|M_a(0,x), M_a(x,2x)\}.$$

Proof. Let $g : X \to Y$ be a mapping defined by $g(x) := f(2x) - 8f(x)$ and let

$$\psi_a(x) := \max\{|2|\varphi_a(x,x), |2|\varphi_a(0,x), \varphi_a(x,2x)\},$$

for all $x \in X$. Similar to the proof of Theorem 3.1, we have

$$\|g(2x) - 2g(x)\|_Y \leq \psi_a(x), \tag{3.19}$$

for all $x \in X$. From our assumptions, we infer that

$$\lim_{n\to\infty} |2|^n \psi_a(\frac{x}{2^n}) = 0, \quad \lim_{n\to\infty} \max\{|2|^i \psi_a(\frac{x}{2^i}) : \quad 1 \leq i < n\} < \infty, \tag{3.20}$$

for all $x \in X$. Replacing x by $\frac{x}{2^{n+1}}$ in (3.19) and multiplying both sides of (3.19) by $|2|^n$, we get

$$\|2^{n+1} g(\frac{x}{2^{n+1}}) - 2^n g(\frac{x}{2^n})\|_y \leq |2|^n \psi_a(\frac{x}{2^{n+1}}), \tag{3.21}$$

for all $x \in X$ and all non-negative integer n. So we have

$$\|2^n g(\frac{x}{2^n} - 2^m g(\frac{x}{2^m}\|_Y \leq \frac{1}{|2|} \max\{|2|^{(i+1)} \psi_a(\frac{x}{2^{i+1}}) : \quad m \leq i < n\}, \tag{3.22}$$

for all $x \in X$ and all non-negative integers n and m with $n \geq m$. Therefore we conclude from (3.20) and (3.21) that the sequence $\{2^n g(\frac{x}{2^n}\}$ is a Cauchy sequence in Y, for all $x \in X$ and so converges in Y, for all $x \in X$, since Y is complete. Thus one can define the mapping $A : X \to Y$ by

$$A(x) := \lim_{n\to\infty} 2^n g(\frac{x}{2^n}),$$

for all $x \in X$. Letting $m = 0$ and passing to the limit when $n \to \infty$ in (3.22) we get (3.18). The rest of the proof is similar to the proof of Theorem 3.1. □

3.3. Corollary. *Let* θ, r, s *be non-negative real numbers such that* $r, s > 1$ *or* $0 \leq r, s < 1$ *and* $|2| < 1$*. Suppose that a function* $f : X \to Y$ *with* $f(0) = 0$ *satisfies the inequality*

$$\|Df(x,y)\|_Y \leq \varphi_a(x,y) := \begin{cases} \theta, & r = s = 0 \\ \theta\|x\|_X^r, & r > 0, s = 0 \\ \theta\|y\|_X^s, & r = 0, s > 0 \\ \theta(\|x\|_X^r + \|y\|_X^s), & r, s > 0 \end{cases}$$

for all $x, y \in X$*. Then there exists a unique additive mapping* $A : X \to Y$ *satisfying*

$$\|f(2x) - 8f(x) - A(x)\|_Y \leq$$
$$\frac{\theta}{|2|} \begin{cases} 1, & r = s = 0 \\ \|x\|_X^r, & r > 0, s = 0 \\ |2|\|x\|_X^s, & r = 0, s > 0 \\ \max\{|2|(\|x\|_X^r + \|x\|_X^s), (\|x\|_X^r + \|2x\|_X^s)\}, & r, s > 0 \end{cases}$$

for all $x \in X$ *where* $r, s > 1$ *and satisfying*

$$\|f(2x) - 8f(x) - A(x)\|_Y \leq \frac{\theta}{|2|} \begin{cases} |2|, & r = s = 0 \\ \frac{|2|\|x\|_X^r}{|2|^r}, & r > 0, s = 0 \\ \frac{|4|\|x\|_X^s}{|2|^s}, & r = 0, s > 0 \\ |2|(\frac{\|x\|_X^r}{|2|^r} + \|x\|_X^s), & r, s > 0 \end{cases}$$

for all $x \in X$ *where* $r, s < 1$.

Proof. The result follows by Theorem 3.1 when $0 < r, s < 1$, and by Theorem 3.2 when $r, s > 1$. $\square$

The following corollary also can be deduced from Theorems 3.1 and 3.2.

3.4. Corollary. *Let $\theta \geq 0$ and $r, s > 0$ be non-negative real numbers such that $\lambda := r + s \neq 1$. Suppose that the function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality*

$$\|Df(x, y)\|_Y \leq \varphi_a(x, y) := \theta\|x\|_X^r\|y\|_Y^s \tag{3.23}$$

for all $x, y \in X$. Then there exists a unique additive mapping $A : X \to Y$ satisfying the inequality

$$\|f(2x) - 8f(x) - A(x)\|_Y \leq \frac{W\theta}{|2|}\|x\|_X^r\|y\|_Y^s$$

for all $x, y \in X$ when $\lambda > 1$ with $W = \max\{|2|, |2|^s\}$, and satisfying

$$\|f(2x) - 8f(x) - A(x)\|_Y \leq \frac{|2|\theta}{|2|^\lambda}\|x\|_X^r\|y\|_Y^s$$

for all $x, y \in X$ when $\lambda < 1$.

3.5. Theorem. *Let $\varphi_c : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty} \frac{1}{|8|^n}\varphi_c(2^n x, 2^n y) = 0 \tag{3.24}$$

$$M_c(x, y) = \lim_{n\to\infty} \max\{\frac{1}{|8|^i}\varphi_c(2^i x, 2^i y) : \quad 0 \leq i < n\} < \infty \tag{3.25}$$

$$\lim_{t\to\infty} \lim_{n\to\infty} \max\{\frac{1}{|8|^i}\varphi_c(2^i x, 2^i y) : \quad t \leq i < t + n\} = 0 \tag{3.26}$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x, y)\|_Y \leq \varphi_c(x, y) \tag{3.27}$$

for all $x, y \in X$. Then the limit

$$C(x) := \lim_{n\to\infty} \frac{1}{8^n}[f(2^{n+1}x) - 2f(2^n x)]$$

exists, for all $x \in X$, and the mapping $C : X \to Y$ is the unique cubic mapping satisfying

$$\|f(2x) - 2f(x) - C(x)\|_Y \leq \frac{1}{|8|}\tilde{\varphi}_c(x) \tag{3.28}$$

for all $x \in X$, where

$$\tilde{\varphi}_c(x) := \max\{|2|M_c(x, x), |2|M_c(0, x), M_c(x, 2x)\}.$$

Proof. Similar to the proof of Theorem 3.1 we have

$$\|f(4x) - 10f(2x) + 16f(x)\|_Y \leq \psi_c(x), \tag{3.29}$$

for all $x \in X$, where $\psi_c(x) := \max\{|2|\varphi_a(x, x), |2|\varphi_a(0, x), \varphi_a(x, 2x)\}$. Let $h : X \to Y$ be a mapping defined by $h(x) := f(2x) - 2f(x)$ for all $x \in X$. Therefore (3.29) means that

$$\|h(2x) - 8h(x)\|_Y \leq \psi_c(x) \tag{3.30}$$

for all $x \in X$. By the relations (3.24) and (3.25), we infer that

$$\lim_{n\to\infty} \max\{\frac{\psi_c(2^i x)}{|8|^i} : \quad 0 \leq i < n\} < \infty, \quad \lim_{n\to\infty} \frac{1}{|8|^n}\psi_c(2^n x) = 0 \tag{3.31}$$

for all $x \in X$. Replacing x by $2^n x$ in (3.30) and dividing both sides of (3.30) by $|8|^{n+1}$ we get

$$\|\frac{1}{8^{n+1}}h(2^{n+1}x) - \frac{1}{8^n}h(2^n x)\|_Y \leq \frac{1}{|8|^{n+1}}\psi_c(2^n x), \tag{3.32}$$

for all $x \in X$ and all non-negative integer n. So we have

$$\|\frac{1}{8^n}h(2^n x) - \frac{1}{8^m}h(2^m x)\|_Y \leq \frac{1}{|8|}\max\{\frac{\psi_c(2^i x)}{|8|^i} m \leq i < n\}, \tag{3.33}$$

for all $x \in X$ and all non-negative integers n and m with $n \geq m$. Therefore we conclude from (3.31) and (3.32) that the sequence $\{\frac{1}{8^n}h(2^n x)\}$ is a Cauchy sequence in Y for all $x \in X$. The sequence $\{\frac{1}{8^n}h(2^n x)\}$ converges in Y, for all $x \in X$, since Y is complete. So one can define the mapping $C : X \to Y$ by

$$C(x) := \lim_{n\to\infty} \frac{1}{8^n}h(2^n x), \tag{3.34}$$

for all $x \in X$. Letting $m = 0$ and passing to the limit when $n \to \infty$ in (3.33), we get (3.28). Now we show that C is a cubic mapping. It follows from (3.31), (3.32) and (3.34) that

$$\begin{aligned}\|C(2x) - 8C(x)\|_Y &= \lim_{n\to\infty} \|\frac{1}{8^n}h(2^{n+1}x) - \frac{1}{8^{n-1}}h(2^n x)\|_Y \\ &\leq \lim_{n\to\infty} \frac{1}{|8|^n}\psi_c(2^n x) = 0,\end{aligned}$$

for all $x \in X$. So

$$C(2x) = 8C(x), \tag{3.35}$$

for all $x \in X$. On the other hand it follows from (3.24), (3.27) and (3.34) that

$$\begin{aligned}\|DC(x,y)\|_Y &= \lim_{n\to\infty} \frac{1}{|8|^n}\|Dh(2^n x, 2^n y)\|_Y \\ &= \lim_{n\to\infty} \frac{1}{|8|^n}\{\|Df(2^{n+1}x, 2^{n+1}y) - 2Df(2^n x, 2^n y)\|_Y\} \\ &\leq \lim_{n\to\infty} \frac{1}{|8|^n}\max\{\|Df(2^{n+1}x, 2^{n+1}y)\|_Y, |2|\|Df(2^n x, 2^n y)\|_Y\} \\ &\leq \lim_{n\to\infty} \frac{1}{|8|^n}\max\{\varphi_c(2^{n+1}x, 2^{n+1}y), |2|\varphi_c(2^n x, 2^n y)\} = 0\end{aligned}$$

for all $x, y \in X$. Hence the mapping C satisfies (1.2). So by Lemma (1.2), the mapping $x \mapsto C(2x) - 2C(x)$ is cubic. Therefore (3.35) implies that C is cubic.

To prove the uniqueness of C, let $T : X \to Y$ be another cubic mapping satisfying (3.28). So it follows from (3.26) that $\lim_{t\to\infty} \frac{\tilde{\varphi}_c(2^t x)}{|8|^t} = 0$, for all $x, y \in X$ and $x \in \{0, y, \frac{y}{2}\}$. So by relations (3.28) and (3.24)

$$\|C(x) - T(x)\|_Y = \lim_{n\to\infty} \frac{1}{|8|^n}\|h(2^n x) - T(2^n x)\|_Y \leq \frac{1}{|8|}\lim_{n\to\infty} \frac{1}{|8|^n}\tilde{\varphi}_c(2^{nx}) = 0,$$

for all $x \in X$. So $C = T$. □

3.6. Theorem. *Let $\varphi_c : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty} |8|^n \varphi_c(\frac{x}{2^n}, \frac{y}{2^n}) = 0 \tag{3.36}$$

$$M_c(x,y) = \lim_{n\to\infty} \max\{|8|^i \varphi_c(\frac{x}{2^i}, \frac{y}{2^i}) : \quad 1 \leq i < n\} < \infty \tag{3.37}$$

$$\lim_{t\to\infty} \lim_{n\to\infty} \max\{|8|^i \varphi_c(\frac{x}{2^i}, \frac{y}{2^i}) : t+1 \leq i < t+n\} = 0 \tag{3.38}$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x,y)\|_Y \leq \varphi_c(x,y)$$

for all $x, y \in X$. Then the limit

$$C(x) = \lim_{n\to\infty} 8^n[f(\frac{x}{2^{n-1}}) - 2f(\frac{x}{2^n})]$$

exists, for all $x \in X$, and the mapping $C : X \to Y$ is the unique cubic mapping satisfying

$$(3.39) \qquad \|f(2x) - 2f(x) - C(x)\|_Y \leq \frac{1}{|8|}\tilde{\varphi}_c(x),$$

for all $x \in X$, where

$$\tilde{\varphi}_c(x) := \max\{|2|M_c(x,x), |2|M_c(0,x), M_c(x,2x)\}.$$

Proof. Let $h : X \to Y$ be a mapping defined by $h(x) := f(2x) - 2f(x)$ and let

$$\psi_c(x) := \max\{|2|\varphi_a(x,x), |2|\varphi_c(0,x), \varphi_c(x,2x)\}$$

for any $x \in X$. Similar to the proof of Theorem 3.5, for every $x \in X$, we have

$$(3.40) \qquad \|h(2x) - 8h(x)\|_Y \leq \psi_c(x).$$

From (3.36) and (3.37) we infer that

$$(3.41) \qquad \lim_{n\to\infty} |8|^n\psi_c(\frac{x}{2^n}) = 0, \text{ and } \lim_{n\to\infty} \max\{|8|^i\psi_c(\frac{x}{2^i}) : 1 \leq i < n\} < \infty,$$

for all $x \in X$. Replacing x by $\frac{x}{2^{n+1}}$ in (3.40) and multiplying both sides of (3.40) by $|8|^n$, we get

$$(3.42) \qquad \|8^{n+1}h(\frac{x}{2^{n+1}}) - 8^n h(\frac{x}{2^n})\|_y \leq |8|^n\psi_c(\frac{x}{2^{n+1}})$$

for any $x \in X$ and all non-negative integer n. Thus we have

$$(3.43) \qquad \|8^n h(\frac{x}{2^n}) - 8^m h(\frac{x}{2^m})\|_Y \leq \frac{1}{|8|}\max\{|8|^{i+1}\psi_c(\frac{x}{2^{i+1}}) :\ m \leq i < n\},$$

for all $x \in X$ and all non-negative integers n and m with $n \geq m$. Therefore we conclude from (3.41) and (3.42) that the sequence $\{8^n h(\frac{x}{2^n}\}$ is a Cauchy sequence in Y for all $x \in X$. Hence the sequence $\{8^n h(\frac{x}{2^n}\}$ converges in Y, for all $x \in X$, since Y is complete. So one can define the mapping $C : X \to Y$ by

$$C(x) := \lim_{n\to\infty} 8^n h(\frac{x}{2^n})$$

for all $x \in X$. Letting $m = 0$ and passing to the limit when $n \to \infty$ in (3.43) we get (3.39). The rest of the proof is similar to the proof of Theorem 3.5. □

The following two corollaries follow from Theorems 3.5 and 3.6.

3.7. Corollary. *Let θ, r, s be non-negative real numbers such that $r, s > 3$ or $0 \leq r, s < 3$ and $|2| < 1$. Suppose that a function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality of Corollary 3.3, for all $x, y \in X$. Then there exists a unique cubic mapping $C : X \to Y$ satisfying*

$$\|f(2x) - 2f(x) - C(x)\|_Y$$
$$\leq \frac{\theta}{|8|}\begin{cases} 1, & r = s = 0 \\ \|x\|_X^r, & r > 0, s = 0 \\ |2|\|x\|_X^s, & r = 0, s > 0 \\ \max\{|2|(\|x\|_X^r + \|x\|_X^s), (\|x\|_X^r + \|2x\|_X^s), & r, s > 0, \end{cases}$$

for all $x \in X$ when $r, s > 3$ and satisfying

$$\|f(2x) - 2f(x) - C(x)\|_Y$$
$$\leq \theta \begin{cases} 1, & r = s = 0 \\ \frac{\|x\|_X^r}{|2|^r}, & r > 0, s = 0 \\ \frac{\|x\|_X^s}{|2|^s} \max\{|2|, |2|^s\}, & r = 0, s > 0 \\ \max\{|2|(\frac{\|x\|_X^r}{|2|^r} + \frac{\|x\|_X^s}{|2|^s}), (\frac{\|x\|_X^r}{|2|^r} + \|x\|_X^s), & r, s > 0, \end{cases}$$

for all $x \in X$ when $r, s < 3$.

3.8. Corollary. *Let $\theta \geq 0$ and $r, s > 0$ be non-negative real numbers such that $\lambda := r + s \neq 3$ and $|2| < 1$. Suppose that the function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality (3.23), for all $x, y \in X$. Put $W = \max\{|2|, |2|^s\}$. Then there exists a unique cubic mapping $C : X \to Y$ satisfying the inequality*

$$\|f(2x) - 2f(x) - C(x)\|_Y \leq \frac{1}{|8|} W\theta \|x\|_X^r \|x\|_Y^s,$$

for all $x \in X$, when $\lambda > 3$ and satisfying

$$\|f(2x) - 2f(x) - C(x)\|_Y \leq \frac{1}{|2|^\lambda} W\theta \|x\|_X^r \|x\|_Y^s$$

for all $x \in X$, when $\lambda < 3$.

3.9. Theorem. *Let $\varphi : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty} \frac{1}{|8|^n} \varphi(2^n x, 2^n y) = 0$$
$$M_c(x, y) = \lim_{n\to\infty} \max\{\frac{1}{|8|^i} \varphi(2^i x, 2^i y) : \quad 0 \leq i < n\} < \infty$$
$$\lim_{t\to\infty} \lim_{n\to\infty} \max\{\frac{1}{|8|^i} \varphi(2^i x, 2^i y) : \quad t \leq i < t + n\} = 0$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x, y)\|_Y \leq \varphi(x, y), \quad x, y \in X.$$

Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that for every $x \in X$,

$$\|f(x) - A(x) - C(x)\|_Y \leq \frac{1}{|48|} \max\{|4|\tilde{\varphi}_a(x), \tilde{\varphi}_c(x)\} \tag{3.44}$$

where

$$M_a(x, y) := \lim_{n\to\infty} \max\{\frac{1}{|2|^i} \varphi(2^i x, 2^i y) 0 \leq i < n\}$$
$$\tilde{\varphi}_a(x) := \max\{|2|M_a(x, x), |2|M_a(0, x), M_a(x, 2x)\}$$
$$\tilde{\varphi}_c(x) := \max\{|2|M_c(x, x), |2|M_c(0, x).M_c(x, 2x)\}$$

Proof. By Theorems 3.1 and 3.5, there exist an additive mapping $A_0 : X \to Y$ and a cubic mapping $C_0 : X \to Y$ such that

$$\|f(2x) - 8f(x) - A_0(x)\|_Y \leq \frac{1}{|2|} \tilde{\varphi}_a(x), \quad \|f(2x) - 2f(x) - C_0(x)\|_Y \leq \frac{1}{|8|} \tilde{\varphi}_c(x)$$

for all $x \in X$. This implies that for any $x \in X$,

$$\|f(x) + \frac{1}{6} A_0(x) - \frac{1}{6} C_0(x)\|_Y \leq \frac{1}{|48|} \max\{|4|\tilde{\varphi}_a(x), \tilde{\varphi}_c(x)\}.$$

So we obtain (3.44), by letting $A(x) = \frac{1}{6}A_0(x)$ and $C(x) = \frac{1}{6}C_0(x)$, for all $x \in X$.

To prove the uniqueness of A and C, let $A_1, C_1 : X \to Y$ be other additive and cubic mappings satisfying (3.44).

Put $A' = A - A_1$ and $C' = C - C_1$. So

$$\begin{aligned}\|A'(x) + C'(x)\|_Y &\le \max\{\|f(x) - A(x) - C(x)\|_Y, \|f(x) - A_1(x) - C_1(x)\|_Y\} \\ &\le \frac{1}{|48|}\max\{|4|\tilde{\varphi}_a(x), \tilde{\varphi}_c(x)\} \qquad (3.45)\end{aligned}$$

for all $x \in X$. The fact that for every $x \in X$,

$$\lim_{n\to\infty} \frac{1}{|8|^n}\tilde{\varphi}_c(2^n x) = \lim_{n\to\infty} \frac{1}{|2|^n}\tilde{\varphi}_a(2^n x) = 0,$$

and (3.45) imply that

$$\lim_{n\to\infty} \frac{1}{|2|^n}\|A'(2^n x) + C'(2^n x)\|_Y = 0,$$

for all $x \in X$. Therefore $A' = 0$. So it follows from (3.45) that

$$\|C'(x)\|_Y = \lim_{n\to\infty} \|\frac{C'(2^n x)}{8^n}\|_Y \le \lim_{n\to\infty} \frac{1}{|4|}\max\{|4|\frac{\tilde{\varphi}_a(x)}{|8|^n},\ \frac{\tilde{\varphi}_c(x)}{|8|^n}\},$$

for all $x \in X$. Therefore $C' = 0$. $\square$

The next theorem is an alternative result of Theorem 3.9.

3.10. Theorem. *Let $\varphi : X \times X \to [0, \infty)$ be a mapping such that*

$$\lim_{n\to\infty} |2|^n \varphi(\frac{x}{2^n}, \frac{y}{2^n}) = 0$$

$$M_a(x, y) := \lim_{n\to\infty} \max\{|2|^i \varphi(\frac{x}{2^i}, \frac{y}{2^i}) : \quad 1 \le i < n\} < \infty$$

$$\lim_{t\to\infty} \lim_{n\to\infty} \max\{|2|^i \varphi(\frac{x}{2^i}, \frac{y}{2^i}) : \quad t + 1 \le i < t + n\} = 0,$$

for all $y \in X$ and all $x \in \{0, y, \frac{y}{2}\}$. Suppose that a mapping $f : X \to Y$ with $f(0) = 0$ satisfies the inequality

$$\|Df(x, y)\|_Y \le \varphi(x, y),$$

for all $x, y \in X$. Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that

$$\|f(x) - A(x) - C(x)\|_Y \le \frac{1}{|48|}\max\{|4|\tilde{\varphi}_a(x), \tilde{\varphi}_c(x)\}, \quad x \in X,$$

where

$$M_c(x, y) := \lim_{n\to\infty} \max\{|8|^i \varphi(\frac{x}{2^i}, \frac{y}{2^i}) : \quad 1 \le i < n\},$$

$$\tilde{\varphi}_a(x) := \max\{|2|M_a(x, x), |2|M_a(0, x), M_a(x, 2x)\}$$

$$\tilde{\varphi}_c(x) := \max\{|2|M_c(x, x), |2|M_c(0, x), M_c(x, 2x)\}$$

for all $x \in X$.

3.11. Corollary. *Let r, s, θ be non-negative real numbers such that $r, s > 3$ or $0 \le r, s < 1$ and $|2| < 1$. Suppose that a function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality Corollary 3.3, for all $x, y \in X$. Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that*

$$\|f(x) - A(x) - C(x)\|_Y \le \frac{1}{|48|}\gamma_{a1} \text{ for all } x \in X \text{ when } 0 \le r, s < 1$$

and

$$\|f(x)-A(x)-C(x)\|_Y \le \frac{1}{|48|}\max\{\gamma_{a2},\gamma_c\} \text{ for all } x\in X \text{ when } r,s>3,$$

where

$$\gamma_{a1}=\begin{cases} |8|\theta, & r=s=0\\ \frac{|8|\theta\|x\|_X^r}{|2|^r}, & r>0, s=0\\ \frac{|8||2|^s\theta\|x\|_X^s}{|2|^s}, & r=0, s>0\\ |8|\theta(\frac{\|x\|_X^r}{|2|^r}+\|x\|_X^s), & r,s>0,\end{cases}$$

$$\gamma_{a2}=\begin{cases} |4|\theta, & r=s=0\\ |4|\theta\|x\|_X^r, & r>0, s=0\\ |8|\theta\|x\|_X^s, & r=0, s>0\\ |4|\theta\max\{|2|(\|x\|_X^r+\|x\|_X^s),(\|x\|_X^r+\|2x\|_X^s), & r,s>0,\end{cases}$$

$$\gamma_{c}=\begin{cases} |8|\theta, & r=s=0\\ \frac{|8|\theta\|x\|_X^r}{|2|^r}, & r>0, s=0\\ \frac{|8|\theta\|x\|_X^s}{|2|^s}\max\{|2|^s,|2|\}, & r=0, s>0\\ |8|\theta\max\{|2|(\frac{\|x\|_X^r}{|2|^r}+\frac{\|x\|_X^s}{|2|^s}),(\frac{\|x\|_X^r}{|2|^r}+\|x\|_X^s)\}, & r,s>0.\end{cases}$$

3.12. Corollary. *Let $\theta\ge 0$ and $r,s>0$ be real numbers such that $\lambda := r+s\in (0,1)\bigcup(3,+\infty)$ and $|2|<1$. Suppose that a function $f:X\to Y$ with $f(0)=0$ satisfies the inequality (3.23) for all $x,y\in X$. Then there exist a unique additive mapping $A: X\to Y$ and a unique cubic mapping $C:X\to Y$ such that*

$$\|f(x)-A(x)-C(x)\|_Y \le \frac{1}{|48|}W\theta\|x\|_X^r\|x\|_Y^s \text{ for all } x\in X \text{ and } \lambda>3,$$

where $W=\max\{|2|,|2|^s\}$

$$\|f(x)-A(x)-C(x)\|_Y \le \frac{|2|^s}{|6||2|^\lambda}\theta\|x\|_X^r\|x\|_Y^s \text{ for all } x\in X \text{ and } 0<\lambda<1.$$

3.13. Theorem. *Let $\varphi: X\times X\to[0,\infty)$ be a mapping such that*

$$\lim_{n\to\infty}\frac{1}{|2|^n}\varphi(2^nx,2^ny)=0,$$

$$M_a(x,y):=\lim_{n\to\infty}\max\{\frac{1}{|2|^i}\varphi(2^ix,2^iy):\quad 0\le i<n\}<\infty,$$

$$\lim_{t\to\infty}\lim_{n\to\infty}\max\{\frac{1}{|2|^i}\varphi(2^ix,2^iy):\quad t\le i<t+n\}=0,$$

for all $y\in X$ and all $x\in\{0,y,\frac{y}{2}\}$. Also suppose

$$\lim_{n\to\infty}|8|^n\varphi(\frac{x}{2^n},\frac{y}{2^n})=0,$$

$$M_c(x,y):=\lim_{n\to\infty}\max\{|8|^i\varphi(\frac{x}{2^i},\frac{y}{2^i}):\quad 1\le i<n\}<\infty,$$

$$\lim_{t\to\infty}\lim_{n\to\infty}\max\{|8|^i\varphi(\frac{x}{2^i},\frac{y}{2^i}):\quad t+1\le i<t+n\}=0,$$

for all $x,y\in X$ and all $x\in\{0,y,\frac{y}{2}\}$. Suppose that a mapping $f:X\to Y$ with $f(0)=0$ satisfies the inequality

$$\|Df(x,y)\|_Y\le\varphi(x,y)$$

for all $x, y \in X$. Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that

$$\|f(x) - A(x) - C(x)\|_Y \leq \frac{1}{|48|} \max\{|4|\tilde{\varphi}_a(x), \tilde{\varphi}_c(x)\}$$

where

$$\tilde{\varphi}_a(x) := \max\{|2|M_a(x,x), |2|M_a(0,x), M_a(x,2x)\}$$

$$\tilde{\varphi}_c(x) := \max\{|2|M_c(x,x), |2|M_c(0,x), M_c(x,2x)\}$$

for all $x \in X$.

Proof. The proof is similar to the proof of Theorem 3.9. □

3.14. Corollary. *Let θ, r, s be non-negative real numbers such that $1 < r, s < 3$ and $|2| < 1$. Suppose that a function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality $\|Df(x,y)\|_Y \leq \theta(\|x\|_X^r + \|y\|_X^s)$ for all $x, y \in X$. Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that*

$$\|f(x) - A(x) - C(x)\|_Y \leq \frac{1}{|48|} \max\{|4|\gamma_a(x), \gamma_c(x)\} \quad \textit{for all } x \in X,$$

where

$$\gamma_a(x) = \max\{|2|\theta(\|x\|_X^r + \|x\|_X^s), \theta(\|x\|_X^r + \|2x\|_X^s)\},$$

$$\gamma_c(x) = |8|\theta \max\{|2|(\frac{\|x\|_X^r}{|2|^r} + \frac{\|x\|_X^s}{|2|^s}), (\frac{\|x\|_X^r}{|2|^r} + \|x\|_X^s).$$

3.15. Corollary. *Let θ, r, s be non-negative real numbers such that $1 < \lambda := r + s < 3$. Suppose that a function $f : X \to Y$ with $f(0) = 0$ satisfies the inequality*

$$\|Df(x,y)\|_Y \leq \begin{cases} \theta\|x\|_X^r, & r > 0, s = 0 \\ \theta\|y\|_X^s, & r = 0, s > 0 \\ \theta\|x\|_X^r\|y\|_X^s, & r, s > 0 \end{cases}$$

for all $x, y \in X$. Then there exist a unique additive mapping $A : X \to Y$ and a unique cubic mapping $C : X \to Y$ such that

$$\|f(x) - A(x) - C(x)\|_Y \leq \frac{1}{|48|} \max\{|4|\gamma_a(x), \gamma_c(x)\} \quad \textit{for all } x \in X$$

where,

$$\gamma_a(x) = \begin{cases} \theta\|x\|_X^r, & r > 0, s = 0 \\ |2|\theta\|x\|_X^s, & r = 0, s > 0 \\ \theta\|x\|_X^\lambda \max\{|2|, |2|^s\}, & r, s > 0, \end{cases}$$

$$\gamma_c(x) = \begin{cases} \frac{|8|\theta\|x\|_X^r}{|2|^r}, & r > 0, s = 0 \\ \frac{|8|\theta\|x\|_X^s}{|2|^s} \max\{|2|, |2|^s\}, & r = 0, s > 0 \\ \frac{|8|\theta\|x\|_X^\lambda}{|2|^\lambda} \max\{|2|, |2|^s\}, & r, s > 0. \end{cases}$$

3.16. Remark. The hypothesis $f(0) = 0$ is not essential in the statement of the theorems, since it is possible to deal with the auxiliary function $g(x) := f(x) - f(0)$ for which we have $Dg(x,y) = Df(x,y)$.

References

[1] T. Aoki, *On the stability of the linear transformation in Banach spaces*, J. Math. Soc. Japan **2**, 64–66, 1950.

[2] L. M. Arriola and W. A. Beyer, *Stability of the Cauchy functional equation over p-adic fields*, Real Anal. Exchange **31**, 125–132, 2055/2006.

[3] C. Baak, S. -K. Hong, M. -J. Kim, *Generalized quadratic mappings of r-type in several variables*, J. Math. Anal. Appl. **310**, 116–127, 2005.

[4] J. -H. Bae and W. -G. Park, *A functional equation originating from quadratic forms*, J. Math. Anal. Appl. **326**, 1142–1148, 2007.

[5] D. G. Bourgin, *Classes of transformations and bordering transformations*, Bull. Amer. Math. Soc. **57**, 223–237, 1951.

[6] S. Czerwik, *Stability of Functional Equations of Ulam–Hyers–Rassias Type*, (Hadronic Press, Palm Harbor, Florida, 2003).

[7] G. L. Forti, *Hyers-Ulam stability of functional equations in several variables* Aequationes Math. **50** (1-2), 143–190, 1995.

[8] D. H. Hyers, *On the stability of the linear functional equation*, Proc. Nat. Acad. Sci. U. S. A. **27**, 222–224, 1941.

[9] D. H. Hyers and Th. M. Rassias, *Approximate homomorphisms*, Aequationes Math. **44**, 125–153, 1992.

[10] D. H. Hyers, G. Isac and Th. M. Rassias, *Stability of Functional Equations in Several Variables*, (Birkhäuser, Basel, 1998).

[11] S.-M. Jung, *On the Hyers-Ulam Stability of the Functional Equations That Have the Quadratic Property*, Journal of Mathematical Analysis and Applications **222**, 126–137, 1998.

[12] S.-M. Jung, *On the Hyers-Ulam-Rassias Stability of a Quadratic Functional Equation*, Journal of Mathematical Analysis and Applications **232**, 384–393, 1999.

[13] S. -M. Jung, *Hyers–Ulam–Rassias Stability of Functional Equations in Mathematical Analysis*, (Hadronic Press lnc. , Palm Harbor, Florida, 2001).

[14] G. H. Kim, *On the Hyers–Ulam–Rassias stability of functional equations in n-variables*, J. Math. Anal. Appl. **299**, 375–391, 2004.

[15] M. S. Moslehian and Th. M. Rassias, *Stability of functional equations in non-Archimedian spaces*, Appl. Anal. Disc. Math. **1**, 325–334, 2007.

[16] M. S. Moslehian, and Gh. Sadeghi, *Stability of two type of cubic functional equations in non-Archimedian spaces*, Real. Anal. Exchange, **33** (2), 375–383, 2008.

[17] A. Najati, and G. Z. Eskandani, *Stability of a mixed additive and cubic functional equation in quasi-Banach spaces*, J. Math. Anal. Appl. **342**, 1318–1331, 2008.

[18] W. -G. Park, J. -H. Bae, *On a bi-quadratic functional equation and its stability*, Nonlinear Analysis **6** No.2, 643–654, 2005.

[19] C. G. Park, *Generalized quadratic mappings in several variables*, Nonlinear Analysis **57**, 713–722, 2004.

[20] D. Popa and I. Rasa, *The Frechet functional equation with application to the stability of certain operators*, J. Approx. Theory **164**, 138–144, 2012.

[21] D. Popa and I. Rasa, *On the Hyers–Ulam stability of the linear differential equation*, J. Math. Anal. Appl. **381**, 530–537, 2011.

[22] Th. M. Rassias, *Functional Equations, Inequalities and Applications*, (Kluwer Academic Publishers, Dordrecht, Boston and London, 2003).

[23] Th. M. Rassias, *On the stability of functional equations and a problem of Ulam*, Acta Appl. Math. **62** (1), 23–130, 2000.

[24] Th. M. Rassias, *On the stability of the linear mapping in Banach spaces*, Proc. Amer. Math. Soc. **72**, 297–300, 1978.

[25] T. Xu, J.M. Rassias ,W. Xu, A *fixed point approach to the stability of a general mixed additive-cubic equation on Banach modules*, Acta Mathematica Scientia *32*, 866-892, 2012.

[26] S. M. Ulam, *Problems in Modern Mathematics*, (Chapter VI, Science Editions, Wiley, New York, 1960).

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 43 – 50

RESULTS ON THE COMPOSITION AND NEUTRIX COMPOSITION OF THE DELTA FUNCTION

Brian Fisher [a] * and Biljana Jolevska–Tuneska [b] †

Received 09 : 02 : 2012 : Accepted 14 : 11 : 2012

Abstract

The neutrix composition $F(f(x)))$ of a distribution $F(x)$ and a locally summable function $f(x)$ is said to exist and be equal to the distribution $h(x)$ if the neutrix limit of the sequence $\{F_n(f(x))\}$ is equal to $h(x)$, where $F_n(x) = F(x) * \delta_n(x)$ and $\{\delta_n(x)\}$ is a certain sequence of infinitely differentiable functions converging to the Dirac delta-function $\delta(x)$. It is proved that the neutrix composition $\delta^{(s)}\{[\exp_+(x) - 1]^r\}$ exists and

$$\delta^{(s)}\{[\exp_+(x) - 1]^r\} = \sum_{k=0}^{rs+r-1} \frac{(-1)^{s+k} s! c_{rs+r-1,k}}{2rk!} \delta^{(k)}(x),$$

for $r = 1, 2, \ldots$ and $s = 0, 1, 2, \ldots$. Further results are also proved.

Keywords: distribution, dirac-delta function, composition of distributions, neutrix, neutrix limit.

2000 AMS Classification: 46F10.

1. Introduction

Certain operations on smooth functions (such as addition, and multiplication by scalars) can be extended without difficulty to arbitrary distributions. Others (such as multiplication, convolution, and change of variables) can be defined only for particular distributions. Note that it is a difficult task to give a meaning to the expression $F(f(x))$, if F and f are singular distributions.

The technique of neglecting appropriately defined infinite quantities was devised by Hadamard and the resulting finite value extracted from the divergent integral is usually referred to as the Hadamard finite part. In fact, Hadamard's method can be regarded

[a]Department of Mathematics, University of Leicester, Leicester, LE1 7RH, England.

*Email: fbr@le.ac.uk

[b]Ss. Cyril and Methodius University in Skopje Faculty of Electrical Engineering, Karpos II bb, Skopje, Republic of Macedonia.

†Email: biljanaj@feit.ukim.edu.mk

as a particular application of the neutrix calculus developed by van der Corput, see [1]. This is a very general principle for the discarding of unwanted infinite quantities from asymptotic expansions and has been widely exploited in the context of distributions, by Fisher in connection with the problem of compositions of distributions, see [2] or [3].

In the following, we let $\mathcal{D}$ be the space of infinitely differentiable functions φ with compact support and let $\mathcal{D}[a,b]$ be the space of infinitely differentiable functions with support contained in the interval $[a,b]$. We let $\mathcal{D}'$ be the space of distributions defined on $\mathcal{D}$ and let $\mathcal{D}'[a,b]$ be the space of distributions defined on $\mathcal{D}[a,b]$.

Now let $\rho(x)$ be a function in $\mathcal{D}[-1,1]$ having the following properties:
(i) $\rho(x) = 0$ for $|x| \geq 1$,
(ii) $\rho(x) \geq 0$,
(iii) $\rho(x) = \rho(-x)$,
(iv) $\displaystyle\int_{-1}^{1} \rho(x)\,dx = 1$.

Putting $\delta_n(x) = n\rho(nx)$ for $n = 1, 2, \dots$, it follows that $\{\delta_n(x)\}$ is a regular sequence of infinitely differentiable functions converging to the Dirac delta-function $\delta(x)$. Further, if F is a distribution in $\mathcal{D}'$ and $F_n(x) = \langle F(x-t), \delta_n(x)\rangle$, then $\{F_n(x)\}$ is a regular sequence of infinitely differentiable functions converging to $F(x)$.

There have been several attempts recently to define distributions of the form $F(f(x))$ in $\mathcal{D}'$, where F and f are distributions in $\mathcal{D}'$, see [6] and [4]. At the beginning, we look at the following definition which is a generalization of Gel'fand and Shilov's definition of the composition involving the delta function, see [10]. This definition was given in [2] by Fisher, it involves neutrix limit and was originally called the neutrix composition of distributions.

1.1. Definition. Let F be a distribution in $\mathcal{D}'$ and let f be a locally summable function. We say that the neutrix composition $F(f(x))$ exists and is equal to h on the open interval (a,b) if

$$\mathop{\mathrm{N-lim}}_{n\to\infty} \int_{-\infty}^{\infty} F_n(f(x))\varphi(x)dx = \langle h(x), \varphi(x)\rangle$$

for all φ in $\mathcal{D}[a,b]$, where $F_n(x) = F(x) * \delta_n(x)$ for $n = 1, 2, \dots$ and N is the neutrix, see [1], having domain N' the positive integers and range N'' the real numbers, with negligible functions which are finite linear sums of the functions

$$n^\lambda \ln^{r-1} n,\ \ln^r n : \qquad \lambda > 0,\ r = 1, 2, \dots$$

and all functions which converge to zero in the usual sense as n tends to infinity.

If f, g are two distributions then in the ordinary sense the composition $f(g)$ does not necessarily exist, but the neutrix composition can exist. Thus the definition of the neutrix composition is an extension of the regular definition of compositions of distributions. Some neutrix composition of distributions are considered in [9], [11] and [12].

Recently, Jack Ng and van Dam applied the neutrix calculus, in conjuction with the Hadamard integral, developed by van der Corput, to quantum field theories, in particular, to obtain finite results for the cofficients in the perturbation series. They also applied neutrix calculus to quantum field theory, obtaining finite renormalization in the loop calculations, see [13] and [14].

Now let $f(x)$ be an infinitely differentiable function having a single simple root at the point $x = x_0$. Gel'fand and Shilov defined the distribution $\delta^{(r)}(f(x))$ by the equation

$$\delta^{(r)}(f(x)) = \frac{1}{|f'(x_0)|}\left[\frac{1}{|f'(x)|}\frac{d}{dx}\right]^r \delta(x - x_0),$$

for $r = 0, 1, 2, \dots$, see [10].

The following theorems were proved in [5], [6], [8] and [7] respectively.

1.2. Theorem. *The neutrix composition* $\delta^{(s)}(\operatorname{sgn} x|x|^\lambda)$ *exists and*

$$\delta^{(s)}(\operatorname{sgn} x|x|^\lambda) = 0$$

for $s = 0, 1, 2, \dots$ *and* $(s+1)\lambda = 1, 3, \dots$ *and*

$$\delta^{(s)}(\operatorname{sgn} x|x|^\lambda) = \frac{(-1)^{(s+1)(\lambda+1)} s!}{\lambda[(s+1)\lambda - 1]!}\delta^{((s+1)\lambda-1)}(x)$$

for $s = 0, 1, 2, \dots$ *and* $(s+1)\lambda = 2, 4, \dots$.

1.3. Theorem. *The compositions* $\delta^{(2s-1)}(\operatorname{sgn} x|x|^{1/s})$ *and* $\delta^{(s-1)}(|x|^{1/s})$ *exist and*

$$\delta^{(2s-1)}(\operatorname{sgn} x|x|^{1/s}) = \frac{(2s)!}{2}\delta'(x),$$
$$\delta^{(s-1)}(|x|^{1/s}) = (-1^s \delta(x)$$

for $s = 1, 2, \dots$.

1.4. Theorem. *The neutrix composition* $\delta^{(s)}[\ln^r(1+|x|)]$ *exists and*

$$\delta^{(s)}[\ln^r(1+|x|)] = \sum_{k=0}^{sr+r-1}\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^{s-i}[1+(-1)^k]s!(i+1)^{rs+r-1}}{2r(rs+r-1)!k!}\delta^{(k)}(x)$$

for $s = 0, 1, 2, \dots$ *and* $r = 1, 2, \dots$.

In particular, the composition $\delta[\ln(1+|x|)]$ *exists and*

$$\delta[\ln(1+|x|)] = \delta(x).$$

1.5. Theorem. *The neutrix composition* $\delta^{(s)}(\sinh^{-1} x_+)$ *exists and*

$$\delta^{(s)}(\sinh^{-1} x_+) = \sum_{k=0}^{s}\sum_{i=0}^{k}\binom{k}{i}(-1)^{s+i+k}\frac{(k-2i+1)^s + (k-2i-1)^s}{2^k k!}\delta^{(k)}(x)$$

for $s = 0, 1, 2, \dots$.

2. Main Results

In the following, the functions $\exp_+(x)$ and $\exp_-(x)$ are defined by

$$\exp_+(x) = \begin{cases} \exp(x), & x \ge 0, \\ 0, & x < 0 \end{cases} \quad \text{and} \quad \exp_-(x) = \begin{cases} \exp(x), & x \le 0, \\ 0, & x > 0. \end{cases}$$

The constants $c_{i,k}$ are defined by the expansion

$$\frac{\ln^k(1+x)}{1+x} = \sum_{i=1}^{\infty} c_{i,k}x^i \tag{2.1}$$

for $i, k = 1, 2, \dots$ and by the expansion

$$(1+x)^{-1} = \sum_{i=0}^{\infty} c_{i,0}x^i = \sum_{i=0}^{\infty}(-1)^i x^i \tag{2.2}$$

for $i = 0, 1, 2, \dots$ and $k = 0$.

We also need the following lemma, which can be easily proved by induction:

2.1. Lemma.

$$\int_{-1}^{1} t^i \rho^{(s)}(t)\,dt = \begin{cases} 0, & 0 \le i < s, \\ (-1)^s s!, & i = s \end{cases}$$

and

$$\int_0^1 t^s \rho^{(s)}(t)\,dt = \tfrac{1}{2}(-1)^s s!$$

for $s = 0, 1, 2, \dots$.

We now prove the following theorem.

2.2. Theorem. *The neutrix composition* $\delta^{(s)}\{[\exp_+(x) - 1]^r\}$ *exists and*

$$\delta^{(s)}\{[\exp_+(x) - 1]^r\} = \sum_{k=0}^{rs+r-1} \frac{(-1)^{s+k} s! c_{rs+r-1,k}}{2rk!} \delta^{(k)}(x), \tag{2.3}$$

for $r = 1, 2, \dots$ *and* $s = 0, 1, 2, \dots$, *where the constants* $c_{rs+r-1,k}$ *are defined with relations (2.1) and (2.2).*

In particular

$$\delta[\exp_+(x) - 1] = \frac{1}{2}\delta(x), \tag{2.4}$$

$$\delta\{[\exp_+(x) - 1]^2\} = -\frac{1}{4}\delta(x) + \frac{1}{4}\delta'(x), \tag{2.5}$$

$$\delta'\{[\exp_+(x) - 1]^2\} = \frac{1}{2}\delta(x) - \frac{1}{2}\delta'(x). \tag{2.6}$$

Proof. We will first of all prove equation (2.3) on the interval $[-1, 1]$. To do this, we need to evaluate

$$\begin{aligned} \int_{-1}^{1} x^k \delta_n^{(s)}\{[\exp_+(x) - 1]^r\}\,dx &= \\ &= \int_0^1 x^k \delta_n^{(s)}\{[\exp(x) - 1]^r\}\,dx + \int_{-1}^{0} x^k \delta_n^{(s)}[(-1)^r]\,dx \\ &= n^{s+1}\int_0^1 x^k \rho^{(s)}\{n[\exp(x) - 1]^r\}\,dx + 0 \\ &= I. \end{aligned} \tag{2.7}$$

Making the substitution $n[\exp(x) - 1]^r = t$ or

$$x = \ln[1 + (t/n)^{1/r}],$$

we have

$$dx = \frac{t^{1/r-1}\,dt}{rn^{1/r}[1 + (t/n)^{1/r}]}.$$

Then for for $n > 1$, we have

$$\begin{aligned} I &= \frac{n^{s+1}}{rn^{1/r}} \int_0^1 \frac{\ln^k[1 + (t/n)^{1/r}] t^{1/r-1}}{1 + (t/n)^{1/r}} \rho^{(s)}(t)\,dt \\ &= \sum_{i=0}^{\infty} \frac{c_{i,k}}{r} \int_0^1 \frac{t^{(i+1)/r-1}}{n^{(i+1)/r-s-1}} \rho^{(s)}(t)\,dt. \end{aligned}$$

It follows that

$$
\begin{aligned}
\operatorname*{N-lim}_{n\to\infty} I &= \operatorname*{N-lim}_{n\to\infty} \int_0^1 x^k \delta_n^{(s)}\{[\exp(x)-1]^r\}\,dx \\
&= \frac{(-1)^s s! c_{rs+r-1,k}}{2r},
\end{aligned} \tag{2.8}
$$

on using the lemma 2.1, for $k = 0,1,2,\dots,rs+r-1$, $r = 1,2,\dots$ and $s = 0,1,2,\dots$.

Next, when $k = rs+r$, we have

$$
\begin{aligned}
\int_0^1 \left|x^{rs+r}\delta_n^{(s)}\{[\exp(x)-1]^r\}\right|dx &\le \frac{n^{s+1}}{rn^{1/r}} \int_0^1 \left|\frac{\ln^{rs+r}[1+(t/n)^{1/r}]t^{1/r-1}}{1+(t/n)^{1/r}}\rho^{(s)}(t)\right| dt \\
&= O(n^{-1/r}),
\end{aligned}
$$

since $|\ln^{rs+r}[1+(t/n)^{1/r}]| = O(n^{-s-1})$. Hence, if $\psi(x)$ is an arbitrary continuous function, then

$$
\lim_{n\to\infty} \int_0^1 x^{rs+r}\delta_n^{(s)}\{[\exp(x)-1]^r\}\psi(x)\,dx = 0, \tag{2.9}
$$

for $r = 1,2,\dots$ and $s = 0,1,2,\dots$.

Further,

$$
\begin{aligned}
\operatorname*{N-lim}_{n\to\infty} \int_{-1}^0 x^{rs+r}\delta_n^{(s)}(0)\psi(x)\,dx &= \operatorname*{N-lim}_{n\to\infty} n^{s+1}\int_{-1}^0 x^{rs+r)}\rho^{(s)}(0)\psi(x)\,dx \\
&= 0,
\end{aligned} \tag{2.10}
$$

for $r = 1,2,\dots$ and $s = 0,1,2,\dots$.

Now let φ be an arbitrary function in $\mathcal{D}[-1,1]$. By Taylor's Theorem we have

$$
\varphi(x) = \sum_{k=0}^{rs+r-1} \frac{x^k\varphi^{(k)}(0)}{k!} + \frac{x^{rs+r}\varphi^{(rs+r)}(\xi x)}{s!},
$$

where $0 < \xi < 1$. Then

$$\begin{aligned}
\operatorname*{N-lim}_{n\to\infty}\langle\delta_n^{(s)}\{[\exp(x)-1]^r\},\varphi(x)\rangle &= \\
&= \operatorname*{N-lim}_{n\to\infty}\sum_{k=0}^{rs+r-1}\frac{\varphi^{(k)}(0)}{k!}\int_{-1}^{1}x^k\delta_n^{(s)}\{[\exp(x)-1]^r\}\,dx \\
&\quad + \operatorname*{N-lim}_{n\to\infty}\int_{-1}^{1}\frac{x^{rs+r}}{(rs+r)!}\delta_n^{(s)}\{[\exp(x)-1]^r\}\varphi^{(rs+s)}(\xi x)\,dx \\
&= \operatorname*{N-lim}_{n\to\infty}\sum_{k=0}^{rs+r-1}\frac{\varphi^{(k)}(0)}{k!}\int_{0}^{1}x^k\delta_n^{(s)}\{[\exp(x)-1]^r\}\,dx \\
&\quad + \operatorname*{N-lim}_{n\to\infty}\sum_{k=0}^{rs+r-1}\frac{\varphi^{(k)}(0)}{k!}\int_{-1}^{0}x^k\delta_n^{(s)}(0)\,dx \\
&\quad + \operatorname*{N-lim}_{n\to\infty}\int_{0}^{1}\frac{x^{rs+r}}{(rs+r)!}\delta_n^{(s)}\{[\exp(x)-1]^r\}\varphi^{(s)}(\xi x)\,dx \\
&\quad + \operatorname*{N-lim}_{n\to\infty}\int_{-1}^{0}\frac{x^{rs+r-1}}{(rs+r-1)!}\delta_n^{(s)}(0)\varphi^{(s)}(\xi x)\,dx \\
&= \sum_{k=0}^{rs+r-1}\frac{(-1)^s s!c_{rs+r-1,k}}{2rk!}\varphi^{(k)}(0) \\
&= \sum_{k=0}^{rs+r-1}\frac{(-1)^{s+k} s!c_{rs+r-1,k}}{2rk!}\langle\delta^{(k)}(x),\varphi(x)\rangle,
\end{aligned}$$

on using equations (2.7), (2.8), (2.9) and (2.10), for $r = 2,3,\dots$ and $s = 1,2,\dots$.

This proves that the neutrix composition $\delta^{(s)}\{[\exp_+(x)-1]^r\}$ exists and

$$\delta^{(s)}\{[\exp_+(x)-1]^r\} = \sum_{k=0}^{rs+r-1}\frac{(-1)^{s+k}s!c_{rs+r-1,k}}{2rk!}\delta^{(k)}(x),$$

on the interval $[-1,1]$ for $r = 1,2,\dots$ and $s = 0,1,2,\dots$.

It is obvious that $\delta^{(s)}\{[\exp_+(x)-1]^r\} = 0$, if $x \neq 0$ and so the neutrix composition $\delta^{(s)}\{[\exp_+(x)-1]^r\}$ exists on the real line.

Equations (2.4), (2.5) and (2.6) follow on noting that $c_{0,0} = 1$, $c_{1,0} = -1$ and $c_{1,1} = -1$.

Finally note that when $r = 1$ and $s = 0$, the normal limits exist and so the composition $\delta[\exp_+(x)-1]$ exists. This completes the proof of the theorem 2.2.

2.3. Corollary. *The neutrix composition $\delta^{(s)}\{[1-\exp_-(x)]^r\}$ exists and*

$$\delta^{(s)}\{[1-\exp_-(x)]^r\} = \sum_{k=0}^{rs+r-1}\frac{(-1)^{rs+r+s+k-1}s!c_{rs+r-1,k}}{2rk!}\delta^{(k)}(x), \tag{2.11}$$

for $r = 1,2,\dots$ and $s = 0,1,2,\dots$.

In particular

$$\delta\{[1-\exp_-(x)]\} = \frac{1}{2}\delta(x), \tag{2.12}$$

$$\delta\{[1-\exp_-(x)]^2\} = \frac{1}{4}\delta(x) - \frac{1}{4}\delta'(x), \tag{2.13}$$

$$\delta'\{[1-\exp_-(x)]^2\} = \frac{1}{2}\delta(x) - \frac{1}{2}\delta'(x). \tag{2.14}$$

Proof. To prove equation (2.11) on the interval $[-1, 1]$, we need to evaluate

$$\begin{aligned}
\int_{-1}^{1} x^k \delta_n^{(s)}\{[1-\exp_-(x)]^r\}\,dx &= \\
&= \int_{-1}^{0} x^k \delta_n^{(s)}\{[1-\exp(x)]^r\}\,dx + \int_0^1 x^k \delta_n^{(s)}(1)\,dx \\
&= n^{s+1}\int_{-1}^{0} x^k \rho^{(s)}\{n[1-\exp(x)]^r\}\,dx + 0 \\
&= I. \qquad (2.15)
\end{aligned}$$

Making the substitution $n[1-\exp(x)]^r = t$ or

$$x = \ln[1-(t/n)^{1/r}],$$

we have

$$dx = -\frac{t^{1/r-1}\,dt}{rn^{1/r}[1-(t/n)^{1/r}]}.$$

Then for for $n > 1$, we have

$$\begin{aligned}
I &= \frac{n^{s+1}}{rn^{1/r}}\int_0^1 \frac{\ln^k[1-(t/n)^{1/r}]t^{1/r-1}}{1-(t/n)^{1/r}}\rho^{(s)}(t)\,dt \\
&= \sum_{i=0}^{\infty}\frac{(-1)^i c_{i,k}}{r}\int_0^1 \frac{t^{(i+1)/r-1}}{n^{(i+1)/r-s-1}}\rho^{(s)}(t)\,dt.
\end{aligned}$$

It follows that

$$\begin{aligned}
\operatorname*{N-lim}_{n\to\infty} I &= \operatorname*{N-lim}_{n\to\infty}\int_0^1 x^k \delta_n^{(s)}\{[1-\exp(x)]^r\}\,dx \\
&= \frac{(-1)^{rs+r+s-1}s!c_{rs+r-1,k}}{2r}, \qquad (2.16)
\end{aligned}$$

on using the lemma 2.1, for $k = 0, 1, 2, \ldots, rs+r-1$, $r = 1, 2, \ldots$ and $s = 0, 1, 2, \ldots$.

Next, when $k = rs + r$, we have

$$\begin{aligned}
\int_{-1}^{0} \left|x^{rs+r}\delta_n^{(s)}\{[1-\exp(x)]^r\}\right|dx &\le \\
&\le \frac{n^{s+1}}{rn^{1/r}}\int_{-1}^{0}\left|\frac{\ln^{rs+r}[1-(t/n)^{1/r}]t^{1/r-1}}{1-(t/n)^{1/r}}\rho^{(s)}(t)\right|dt \\
&= O(n^{-1/r}).
\end{aligned}$$

Hence, if $\psi(x)$ is an arbitrary continuous function, then

$$\lim_{n\to\infty}\int_0^1 x^{rs+r}\delta_n^{(s)}\{[1-\exp(x)]^r\}\psi(x)\,dx = 0, \qquad (2.17)$$

for $r = 1, 2, \ldots$ and $s = 0, 1, 2, \ldots$.

The proof of the corollary now follows as in the proof of Theorem 2.2, using (2.15), (2.16) and (2.17). Equations (2.12), (2.13) and (2.14) follows immediately.

2.4. Corollary. *The neutrix composition $\delta^{(s)}[|\exp(x)-1|^r]$ exists and*

$$\delta^{(s)}[|\exp(x)-1|^r] = \begin{cases} \displaystyle\sum_{k=0}^{rs+r-1}\frac{(-1)^k s!c_{rs+r-1,k}}{rk!}\delta^{(k)}(x), & r \text{ odd } s \text{ even}, \\ 0, & r \text{ even}, \\ 0, & r, s \text{ odd} \end{cases} \qquad (2.18)$$

for $r = 1, 2, \ldots$ and $s = 0, 1, 2, \ldots$.

Proof. Equation (2.18) follows on noting that we have

$$\delta^{(s)}[|\exp(x)-1|^r] = \delta^{(s)}[|\exp_+(x)-1|^r] + \delta^{(s)}[|\exp_-(x)-1|^r]$$

and

$$\delta^{(s)}[|\exp_-(x)-1|^r] = \begin{cases} \delta^{(s)}\{[\exp(x)-1]^r\} & r \text{ odd}, \ s \text{ even}, \\ \delta^{(s)}\{[1-\exp(x)]^r\}, & r \text{ even}, \\ -\delta^{(s)}\{[1-\exp(x)]^r\}, & r, s \text{ odd}. \end{cases}$$

References

[1] van der Corput, J. G. *Introduction to the neutrix calculus,* J. Analyse Math., **7**, 291–398, 1959.

[2] Fisher, B. *On defining the change of variable in distributions,* Rostock. Math. Kolloq., **28**, 75–86, 1985.

[3] Fisher, B. *On defining the distribution* $\delta^{(r)}(f(x))$, Rostock. Math. Kolloq., **23**, 73–80, 1993.

[4] Fisher, B. *On defining the distribution* $(x_+^r)_-^{-s}$, Univ. u Novom Sadu Zb. Rad. Prirod.-Mat. Fak. Ser. Mat., **15**, 119–129, 1985.

[5] Fisher, B. *The delta function and the composition of distributions*, Dem. Math. **35**(1), 117–123, 2002.

[6] Fisher, B. *The composition and neutrix composition of distributions*, Proc. Math. Methods Engineering, Springer, 59–69, 2007.

[7] Fisher B. and Kılıçman, A. *On the composition and neutrix composition of the delta function and powers of the inverse hyperbolic sine function*, Integral Transforms Spec. Funct., **21**(12), 935–944, 2010.

[8] Fisher, B., Kraiweeradechachai, T., and Özçağ, E. *Results on the neutrix composition of the delta function*, Hacet. J. Math. Stat., **36**(2), 147–156, 2007.

[9] Fisher, B., Jolevska-Tuneska, B., and Özçağ, E. *Further results on the composition of distributions*, Integral Transforms and special functions, **13** (2), 109–116, 2002

[10] Gel'fand, I. M. and Shilov, G. E. *Generalized Functions*, Vol. I, (Academic Press, 1964).

[11] Jolevska-Tuneska, B. and Ozcag, E. *On the composition of distributions* $x^{-s}\ln|x|$ and $|x|^\mu$, International Journal of Mathematics and Mathematical Sciences, **2007**, 9 pages., doi:10.1155/2007/60129

[12] Kraiweeradechachai, T., Orankitjaroen, S., Fisher. B., and Özçağ, E. *Further results on the neutrix composition of the delta function*, East-West J. Math., **11**(2), 151–164, 2009.

[13] Ng, Jack Y. and van Dam, H. *Neutrix calculus and finite quantum field theory,* J. Phys. A: Math. Gen. **38**, 317–323, 2005.

[14] Ng, Jack Y. and van Dam, H. *An application of neutrix calculus to quantum field theory,* International Journal of Modern Physics A, **21** (2), 297–312, 2006.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 51 – 64

ON STRONGLY AND SEPARABLY ω_1-$p^{\omega+n}$-PROJECTIVE ABELIAN p-GROUPS

Peter Danchev*

Received 02 : 04 : 2012 : Accepted 17 : 12 : 2012

Abstract

Let $n \geq 0$ be an arbitrary integer. We prove some results for strongly n-simply presented abelian p-groups with C-decomposable property, extending classical achievements due to Keef in Commun. Algebra (1990). As applications we define the classes of strongly ω_1-$p^{\omega+n}$-projective and separably ω_1-$p^{\omega+n}$-projective abelian p-groups which are also properly contained in all ω_1-$p^{\omega+n}$-projectives, recently defined by Keef in J. Alg. Numb. Th. Acad. (2010). Moreover, some principal descriptions concerning these new objects are obtained as well.

Keywords: C-decomposable groups, $p^{\omega+n}$-projective groups, strongly n-simply presented groups, ω_1-$p^{\omega+n}$-projective groups, strongly ω_1-$p^{\omega+n}$-projective groups, bounded subgroups, countable subgroups, nice subgroups, Ulm subgroups, Ulm factors.

2000 AMS Classification: 20K10.

1. Introduction and Terminology

Let all groups into consideration throughout the paper be abelian p-torsion groups where p is a fixed prime integer. As usual, for some ordinal $\alpha \geq 0$ and a group G, we state the *α-th Ulm subgroup* $p^\alpha G$, consisting of all elements of G with height $\geq \alpha$, inductively as follows: $p^0 G = G$, $pG = \{pg \mid g \in G\}$, $p^\alpha G = p(p^{\alpha-1}G)$ if $\alpha - 1$ exists (so α is non-limit) and $p^\alpha G = \cap_{\beta<\alpha} p^\beta G$ if $\alpha - 1$ does not exist (so α is limit). The group G is named *p^α-bounded* if $p^\alpha G = \{0\}$; note that these groups have to be reduced. We shall say that G is *Σ-cyclic* if it is a direct sum of cyclic groups, and *separable* if it is p^ω-bounded - notice that Σ-cyclic groups are separable. Most of the important unexplained here notations and notions will follow mainly those from [9].

The class of $p^{\omega+n}$-projective groups, defined originally as in [14], plays an important if not facilitating role in the theory of abelian groups whenever $n \geq 0$ is an integer. There are two similar characterizations of the $p^{\omega+n}$-projectives given in [14] and [1], respectively.

*Department of Mathematics, Plovdiv State University P. Hilendarski, Plovdiv 4000, Bulgaria. Emails: `pvdanchev@yahoo.com, peter.danchev@yahoo.com`

1.1. Theorem. *The group G is $p^{\omega+n}$-projective if and only if precisely one of the following conditions holds:*

(a) there exists a p^n-bounded subgroup P of G such that G/P is Σ-cyclic.

(b) there exists a Σ-cyclic group S with a p^n-bounded subgroup B such that $G \cong S/B$.

Observe that when $n = 0$ we obtain the classical Σ-cyclic groups, i.e., the p^ω-projective groups. Moreover, note that P is of necessity nice in G because G/P is separable.

On the other hand, a few years ago, Keef established in ([12], Proposition 1.4 and Theorem 1.2 (a1)) the following intriguing generalization of $p^{\omega+n}$-projective groups:

1.2. Theorem. *The group G is ω_1-$p^{\omega+n}$-projective if and only if exactly one of the following conditions is valid:*

(i) there is a countable subgroup C of G such that $C \subseteq p^\omega G$ and G/C is $p^{\omega+n}$-projective.

(ii) there is a p^n-bounded subgroup H of G such that G/H is the direct sum of a countable group and a Σ-cyclic group.

Notice that the subgroup C of point (i) of the last theorem is necessarily nice in G satisfying the inequalities $p^{\omega+n}G \subseteq C \subseteq p^\omega G$. So, it is interesting to know whether or not the subgroup H in point (ii) of the same theorem can be chosen to be nice in G. Unfortunately or not, the answer is "no" as it will be demonstrated in the sequel.

Thus adding the niceness will be a non-trivial procedure, and thereby we come to the main concept which motivates the writing of this article.

Definition 1.1. A group G is called *strongly ω_1-$p^{\omega+n}$-projective* if it contains a p^n-bounded nice subgroup A such that G/A is a direct sum of a countable group and a Σ-cyclic group.

Each $p^{\omega+n}$-projective group is necessarily strongly ω_1-$p^{\omega+n}$-projective, while the converse is untrue provided that the group has length strictly greater $\omega + n$. However, $p^{\omega+n}$-bounded strongly ω_1-$p^{\omega+n}$-projective groups must be $p^{\omega+n}$-projective, instead of ω_1-$p^{\omega+n}$-projectives (cf. [12]) which are not.

A weaker version of the last group class is the following:

Definition 1.2. A group G is said to be *separably ω_1-$p^{\omega+n}$-projective* if it contains a p^n-bounded nice subgroup M such that $M \cap p^\omega G = \{0\}$ and G/M is a direct sum of a countable group and a Σ-cyclic group.

It is worthwhile noticing that such a subgroup M, for which $G/(M \oplus p^\omega G)$ is Σ-cyclic, must be nice in G as it will be demonstrated below. Also, Σ-cyclic groups are separably ω_1-$p^{\omega+n}$-projective and, for $n = 1$, p^ω-bounded $p^{\omega+1}$-projective groups are necessarily separably ω_1-$p^{\omega+1}$-projective, whereas in both cases the converse is not true provided that the group has length greater than ω. Even more, p^ω-bounded separably ω_1-$p^{\omega+n}$-projective groups need not be Σ-cyclic; in fact they are $p^{\omega+n}$-projective.

On the other hand, in [4] we enlarged the Keef's concept to the so-termed *weakly ω_1-$p^{\omega+n}$-projective groups* that are groups G containing countable nice subgroups $N \subseteq p^\omega G$ such that $G/N/p^{\omega+n}(G/N) \cong G/(p^{\omega+n}G + N)$ is $p^{\omega+n}$-projective. Likewise, some other improvements of ω_1-$p^{\omega+n}$-projectivity were established in [2] and [5], respectively.

On another vein, in [8] the present author along with Keef defined the class of *(strongly) n-simply presented groups G* which are groups containing a (nice) p^n-bounded subgroup P such that G/P is simply presented. Clearly, (strongly) ω_1-$p^{\omega+n}$-projective groups are (strongly) n-simply presented.

Besides, in [10], it was introduced and investigated the class of *separably n-simply presented groups* that are strongly n-simply presented groups G for which $P\cap p^\omega G = \{0\}$. Evidently, all separably ω_1-$p^{\omega+n}$-projective groups are themselves separably n-simply presented.

In some of the next sections we shall study the above stated concepts more carefully.

2. A Survey of Known Results

In this brief section, we shall list a few more useful results, needed for applicable purposes in the next sections. These results are stated here only for the sake of completeness and for the readers' convenience, and will be utilized below without some more special and concrete referring.

2.1. Proposition. *([9])*

(j) (Nunke's property) A group G is simply presented if and only if $p^\alpha G$ and $G/p^\alpha G$ are both simply presented for some ordinal α.

(jj) (Direct summand property) Direct summands of simply presented groups are again simply presented.

2.2. Proposition. *([14]) Subgroups of $p^{\omega+n}$-projective groups are again $p^{\omega+n}$-projective.*

2.3. Proposition. *([12]) Subgroups of ω_1-$p^{\omega+n}$-projective groups are again ω_1-$p^{\omega+n}$-projective.*

2.4. Proposition. *([8])*

(j) If G is a strongly n-simply presented group with $p^{\omega+n}G = \{0\}$, then G is $p^{\omega+n}$-projective.

(jj) If G is a strongly n-simply presented group, then $G/p^\alpha G$ is strongly n-simply presented for some ordinal α. In particular, $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective.

Moreover, G is strongly n-simply presented if and only if $p^{\alpha+n}G$ and $G/p^{\alpha+n}G$ are both strongly n-simply presented.

3. C-Decomposable Strongly n-Simply Presented p-Groups

As mentioned in the first section, a strongly n-simply presented group is such a group G for which there is a p^n-bounded nice subgroup N with G/N being simply presented.

The next assertion strengthens ([11], Theorem 3).

3.1. Theorem. *Suppose G is a strongly n-simply presented group with $p^\omega G$ simply presented and $G \cong H \oplus K$ where K is a Σ-cyclic group whose final rank is at least $r(p^{\omega+n}G)$. Then G is a direct sum of a simply presented group and a $p^{\omega+n}$-projective group.*

Proof. Since G is strongly n-simply presented, in virtue of [8] the quotient $G/p^{\omega+n}G$ should be $p^{\omega+n}$-projective. But $H/p^{\omega+n}H$ is obviously isomorphic to a summand of $G/p^{\omega+n}G$, and hence it is $p^{\omega+n}$-projective as well. Moreover, $p^\omega H \cong p^\omega G$ is simply presented and hence so is $p^{\omega+n}H$ applying [9]. Therefore, H is strongly n-simply presented again by the utilization of [8]. It follows from Theorem 1.1 (a) that there exists a subgroup $Q \subseteq (H/p^{\omega+n}H)[p^n]$ such that $(H/p^{\omega+n}H)/Q$ is Σ-cyclic. Let P be the subgroup of H containing $p^{\omega+n}H \cong p^{\omega+n}G$ and defined by the equation $P/p^{\omega+n}H = Q$; thus $p^nP \subseteq p^{\omega+n}H$, and $H/p^{\omega+n}H/P/p^{\omega+n}H \cong H/P$ is Σ-cyclic with $p^\omega H \subseteq P$.

Using the idea behind a "standard $\omega+n$-decomposition", there is clearly a subgroup $P_1 \subseteq p^\omega H \subseteq P$ such that if L is a $p^{\omega+n}$-high subgroup of H, and thus it is $p^{\omega+n}$-bounded, then there exists a decomposition $p^\omega G \cong p^\omega H = P_1 \oplus p^\omega L$; so, in particular, $p^n P_1 = p^{\omega+n} H$ since $p^{\omega+n} L = \{0\}$. Indeed, we first claim that $p^\omega L$ is a maximal p^n-bounded summand of $p^\omega H$, so that it is pure and bounded in $p^\omega H$, whence its direct summand. To prove this, we foremost see that $p^n(p^\omega L) = p^{\omega+n} L = \{0\}$, hence $p^\omega L$ is bounded by p^n. Furthermore, because L is isotype in H and hence obviously $p^\omega L$ is pure in $p^\omega H$, we write $p^\omega H = P_1 \oplus p^\omega L$ (see, e.g., [9]). To show the maximality, also write $p^\omega H = X \oplus T$ for some $X \leq p^\omega H$ and $T \leq p^\omega H$ such that $p^n T = \{0\}$. It is apparently seen that $p^{\omega+n} H = p^n X$ and thus immediately $T \cap p^{\omega+n} H = \{0\}$. But $L \cap p^{\omega+n} H = \{0\}$ is maximal with this property, so that $T \subseteq L \cap p^\omega H = p^\omega L$ because as mentioned above L is isotype in H, as required. This gives the claim.

If now $P_2 = P \cap L$, we even have a valuated direct decomposition $P = P_1 \oplus P_2$. In fact, it is elementary to verify that $(p^{\omega+n} H)[p] = P_1[p] = (p^n P_1)[p]$. This insures at once that $P_1 \cap L = \{0\}$ and hence $P_1 \cap P_2 = \{0\}$. Next, since $H[p] = (p^{\omega+n} H)[p] \oplus L[p] = (p^n P_1)[p] \oplus L[p] = P_1[p] \oplus L[p]$ and since L is pure in H (see, cf. [9]), it easily follows that $H[p^n] = P_1[p^n] \oplus L[p^n]$. Therefore, intersecting the last equality with $P \leq H$, the modular law yields that $P[p^n] = P_1[p^n] \oplus (L \cap P)[p^n] = P_1[p^n] \oplus P_2[p^n]$. By what we have just shown above, $p^n P \subseteq p^{\omega+n} H = p^n P_1$ which, because of $P_1 \subseteq P$, is tantamount to $p^n P = p^n P_1$. The last equality directly implies that $P = P_1 + P[p^n]$, that is equivalent to $P = P_1 \oplus P_2$, as asserted. That this decomposition is valuated follows routinely, which technical details we leave to the reader. It is also worth noticing that the equality $P = P_1 \oplus P_2$ is an extension of the equality $p^\omega H = P_1 \oplus p^\omega L$; in fact the modular law ensures for $P_1 \leq p^\omega H \leq P$ that $p^\omega H = P_1 \oplus (P_2 \cap p^\omega H)$. But the latter summand is equal to $L \cap p^\omega H = p^\omega L$ because L is pure in H (e.g., [9]), and consequently we conclude that $p^\omega H = P_1 \oplus p^\omega L$ which was our initial pivotal relation.

We further observe that $L/P_2 \cong (L+P)/P \subseteq H/P$ is Σ-cyclic, and that $p^n P_2 \subseteq p^n P \cap p^n L \subseteq p^{\omega+n} H \cap L = \{0\}$, whence L is $p^{\omega+n}$-projective owing to Theorem 1.1 (a) as well.

Let us now T be a simply presented group with the following Ulm-Kaplansky function: $f_T(\alpha) = f_K(\alpha)$, when $\alpha < \omega$; $f_T(\alpha) = 0$, when $\omega \leq \alpha < \omega + n - 1$, and $f_T(\alpha) = f_G(\alpha)$, when $\omega + n - 1 \leq \alpha$. Note that the existence of such a group T is guaranteed by the fact that K has final rank no less than $r(p^{\omega+n} G)$ - see, for example, ([9], Theorem 83.6).

Next, consider the direct sum $A = T \oplus L$. If $B \subseteq A$ is the subgroup $p^\omega T \oplus P_2$, then apparently $A/B \cong (T/p^\omega T) \oplus (L/P_2)$ is Σ-cyclic. Moreover, $p^n P_1 = p^{\omega+n} H \cong p^{\omega+n} G$ is simply presented, hence in virtue of [9] so is P_1. But $p^\omega T$ is also simply presented (cf. [9]) and, in accordance with the preceding paragraph, it is readily checked that both $p^\omega T$ and P_1 have same Ulm-Kaplansky invariants. Thus [9] allows us to conclude that $p^\omega T \cong P_1$, and so there is an isometry $\phi : B = p^\omega T \oplus P_2 \to P_1 \oplus P_2 = P$. It is easy to check that $f_{G,P}(\alpha) = f_{A,B}(\alpha) = f_{L,P_2}(\alpha) + f_K(\alpha)$, when $\alpha < \omega$, or $f_{G,P}(\alpha) = f_{A,B}(\alpha) = 0$, when $\alpha \geq \omega$. This, however, implies in view of ([9], Theorem 83.4) that ϕ extends to an isomorphism $\Phi : A = T \oplus L \to G$, thus proving the result. $\square$

Remark. It is worth noting that the first part of the above proof actually demonstrates that any $p^{\omega+n}$-high subgroup of a strongly n-simply presented group is $p^{\omega+n}$-projective.

As a direct consequence, we derive a generalization of Corollary 4 from [11].

3.2. Corollary. *The group G is a summand of the direct sum of a simply presented group and a $p^{\omega+n}$-projective group if and only if G is a strongly n-simply presented group such that $p^\omega G$ is simply presented.*

Proof. "$\Rightarrow$". Write $T \oplus P = G \oplus H$ where T is simply presented and P is $p^{\omega+n}$-projective. Evidently, $p^\omega G$ is a summand of $p^\omega T \oplus p^\omega P$ which is simply presented. Therefore, $p^\omega G$ is simply presented referring to [9].

On the other hand, one may observe that $(T/p^{\omega+n}T) \oplus (P/p^{\omega+n}P) \cong (G/p^{\omega+n}G) \oplus (H/p^{\omega+n}H)$. Since $T/p^{\omega+n}T$ is a direct sum of countable groups of length $\omega+n$, hence it is $p^{\omega+n}$-projective, and $P/p^{\omega+n}P$ is $p^{\omega+n}$-projective, the left hand-side is $p^{\omega+n}$-projective too, whence so is $G/p^{\omega+n}G$. Finally, [8] applies to show that G is strongly n-simply presented, as desired.

"$\Leftarrow$". Let G be strongly n-simply presented with $p^\omega G$ simply presented. Also, let C be a Σ-cyclic group whose final rank exceeds the rank of $p^{\omega+n}G$. Then $G \oplus C$ is, by Theorem 3.1, a direct sum of a simply presented group and a $p^{\omega+n}$-projective group, as required. □

Recall that a group G is C-decomposable if $G \cong H \oplus C$ where C is a Σ-cyclic group with the same final rank as that of G.

An other (second) valuable consequence of the chief result of this section is the following generalization of Corollary 5 in [11].

3.3. Corollary. *If G is a C-decomposable strongly n-simply presented group such that $p^\omega G$ is simply presented, then G is the direct sum of a simply presented group and a $p^{\omega+n}$-projective group.*

Proof. It is clear that the final rank of G must be at least as large as the rank of $p^{\omega+n}G$. Furthermore, we apply Theorem 3.1 to get the claim. □

4. Strongly ω_1-$p^{\omega+n}$-Projective p-Groups

As stated in the introductory Section 1, a group G is strongly ω_1-$p^{\omega+n}$-projective if it has a nice subgroup $N \leq G[p^n]$ such that G/N is ω_1-p^ω-projective (= ω-totally Σ-cyclic in terms of [7]), that is, the direct sum of a countable group and a Σ-cyclic group. Respectively, a group G is separably ω_1-$p^{\omega+n}$-projective if it possesses a subgroup $L \leq G[p^n]$ with $L \cap p^\omega G = \{0\}$ such that G/L is ω_1-p^ω-projective (= ω-totally Σ-cyclic), i.e., the direct sum of a countable group and a Σ-cyclic group. It is pretty easy to check that $p^{\omega+n}$-projective groups are strongly ω_1-$p^{\omega+n}$-projective (in fact, the countable group in the direct decomposition of G/N must be exactly $\{0\}$) as well as separable $p^{\omega+n}$-projective groups are separably ω_1-$p^{\omega+n}$-projective (indeed, $p^\omega G = \{0\}$ and again the countable summand from the direct decomposition of G/L has to be precisely $\{0\}$).

In [13] the following useful technicality due to B. Charles was stated explicitly:

Lemma (Charles). *Suppose A is a group with a countable subgroup B such that A/B is Σ-cyclic. Then A is the direct sum of a countable group and a Σ-cyclic group.*

In the case when A/B is $p^{\omega+n}$-projective for some $n \geq 1$, the group A is defined in [12] to be ω_1-$p^{\omega+n}$-projective (compare also with Theorem 1.2 (i) stated above in Section 1) and it is not necessarily a direct sum of a countable group and a $p^{\omega+n}$-projective group; indeed there exists a $p^{\omega+n}$-bounded ω_1-$p^{\omega+n}$-projective group which is not $p^{\omega+n}$-projective (see the comments on pp. 56 and 57 of [12]).

However, it is rather natural to ask whether the following strengthening is true: For some group A let A/B be Σ-cyclic and let B be the direct sum of a countable group and a p^n-bounded group (i.e., p^nB is countable) for some $n \geq 1$. Does it follow that A is the direct sum of a countable group and a $p^{\omega+n}$-projective group? Unfortunately or not, it is untrue, and A is in general a proper subgroup of such a direct sum being an

ω_1-$p^{\omega+n}$-projective group (see, for instance, Theorem 1.2 (b1) and Theorem 1.5 (b) of [12]); that is why an equality may not be fulfilled.

Reciprocally, if A is a group with a Σ-cyclic subgroup C such that A/C is countable, then A is again a direct sum of a countable group and a Σ-cyclic group - see, e.g., [6], or Theorem 1.5 (b) from [12] when $n = 0$.

On the other hand, Megibben proved in [13] the following statement (for some non-trivial generalizations to that fact see also [7] and [3]).

Proposition (Megibben). *Suppose G is a group. Then the following are equivalent:*
(i) $G/p^\omega G$ is Σ-cyclic with $p^\omega G$ countable;
(ii) G is the direct sum of a countable group and a Σ-cyclic group.

Actually, the implication (i) $\Rightarrow$ (ii) in this assertion follows immediately from the above Lemma of Charles. Besides, a subgroup of the direct sum of a countable group and a Σ-cyclic group is again a direct sum of a countable group and a Σ-cyclic group; in fact, if $H \leq G$ where G is such a group, then $p^\omega G$ is countable and $G/p^\omega G$ is Σ-cyclic. But $H/(H \cap p^\omega G) \cong (H + p^\omega G)/p^\omega G \subseteq G/p^\omega G$ is Σ-cyclic as being a subgroup with countable intersection $H \cap p^\omega G$, so that the Lemma of Charles applies to get the assertion.

It is now quite usual to ask whether or not the following enlargement holds:

Question. Let G be a group and $n \geq 0$. Does it follow that the next two points are equivalent?
(a) $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective and $p^{\omega+n}G$ is countable;
(b) G is the direct sum of a countable group and a $p^{\omega+n}$-projective group.

This is true only when $n = 1$ – see Corollary 2.11 from [7]. However, when $n = 2$, the answer is negative – see Example on p. 533 from [7]. (See also [3] for more details when $n \geq 1$.)

Reciprocally, if A is a group with a $p^{\omega+n}$-projective subgroup S such that A/S is countable, then A need not be the direct sum of a countable group and a $p^{\omega+n}$-projective group whenever $n \geq 1$. Indeed, an appeal to Theorem 1.2 (c3) from [12] gives that A is ω_1-$p^{\omega+n}$-projective, whereas Theorem 1.5 (b) of [12] insures that A is only a (proper) subgroup of such a direct sum.

We will now provide the reader with some equivalent characterizations of strongly (respectively, separably) ω_1-$p^{\omega+n}$-projectives.

4.1. Lemma. *The group G is strongly ω_1-$p^{\omega+n}$-projective if and only if there exists a nice subgroup N of G such that $p^n N = \{0\}$, $G/(N + p^\omega G)$ is Σ-cyclic and $p^\omega(G/N) \cong p^\omega G/(p^\omega G \cap N)$ is countable.*

Proof. "$\Rightarrow$". Write $G/N = (A/N) \oplus (B/N)$ where A/N is countable and B/N is Σ-cyclic for some p^n-bounded nice subgroup N of G. Therefore $p^\omega(G/N) = p^\omega(A/N)$ is countable, i.e., same is true for $(p^\omega G + N)/N \cong p^\omega G/(N \cap p^\omega G)$. On the other hand, $G/N/p^\omega(G/N) = G/N/(p^\omega G + N)/N \cong G/(p^\omega G + N)$ should be Σ-cyclic, as stated.

"$\Leftarrow$". Since $G/(N + p^\omega G) \cong G/N/(N + p^\omega G)/N$ is Σ-cyclic and $(N + p^\omega G)/N \cong p^\omega G/(N \cap p^\omega G)$ is countable, the Lemma of Charles applies to deduce that G/N is the direct sum of a countable group and a Σ-cyclic group, as required. □

4.2. Lemma. *The group G is separably ω_1-$p^{\omega+n}$-projective if and only if there exists a nice subgroup P of G such that $p^n P = \{0\}$, $P \cap p^\omega G = \{0\}$ and $G/(P \oplus p^\omega G)$ is Σ-cyclic with countable $p^\omega G$.*

Proof. Follows in the same manner as the above Lemma 4.1, taking into account Lemma 1 from [10] which says that $P \oplus p^\omega G$ is nice in G if and only if P is nice in G (see [4] too). Also, $p^\omega G/(p^\omega G \cap P) \cong p^\omega G$ is now countable. □

4.3. Corollary. *If G is strongly (respectively, separably) ω_1-$p^{\omega+n}$-projective, then so is $p^\alpha G$ for any ordinal α.*

Proof. Let N be a nice p^n-bounded subgroup of G such that G/N is ω-totally Σ-cyclic (in addition, $p^\omega G \cap N = \{0\}$). Consequently, owing to ([7], Theorem 2.6) or to the comments after the Proposition of Megibben, one can see that $(p^\alpha G + N)/N \subseteq G/N$ is also ω-totally Σ-cyclic as being a subgroup, and thus $(p^\alpha G + N)/N \cong p^\alpha G/(p^\alpha G \cap N)$ is also ω-totally Σ-cyclic, where $p^\alpha G \cap N$ is p^n-bounded and nice in $p^\alpha G$. In addition, $p^\omega(p^\alpha G) \cap (p^\alpha G \cap N) = p^{\alpha+\omega} G \cap N \subseteq p^\omega G \cap N = \{0\}$, as needed. □

4.4. Corollary. *If G is strongly (respectively, separably) ω_1-$p^{\omega+n}$-projective, then so is $G/p^\alpha G$ for each ordinal α.*

Proof. Let N be a p^n-bounded nice subgroup of G such that G/N is the direct sum of a countable group and a Σ-cyclic group . Put $N' = (N + p^\alpha G)/p^\alpha G$, and it is easily seen that N' is p^n-bounded and nice in $G/p^\alpha G$. Likewise,

$$G/p^\alpha G/(N + p^\alpha G)/p^\alpha G \cong G/(N + p^\alpha G) \cong G/N/(N + p^\alpha G)/N = G/N/p^\alpha(G/N).$$

But $p^\alpha(G/N)$ is again countable whenever $\alpha \geq \omega$, hence $G/N/p^\alpha(G/N)$ remains a direct sum of a countable group and a Σ-cyclic group. Finally, $G/p^\alpha G$ is a strongly ω_1-$p^{\omega+n}$-projective group, as expected. In addition, the modular law from [9] ensures that $N' \cap p^\omega(G/p^\alpha G) = N' \cap (p^\omega G/p^\alpha G) = [(N + p^\alpha G) \cap p^\omega G]/p^\alpha G = (p^\alpha G + N \cap p^\omega G)/p^\alpha G = \{0\}$ provided $\alpha > \omega$ and $N \cap p^\omega G = \{0\}$. For $\alpha \leq \omega$, the intersection is again clearly equal to zero. □

The next two corollaries are also consequences of results from [8].

4.5. Corollary. *If G is a group such that $p^{\omega+n} G = \{0\}$, then G is strongly ω_1-$p^{\omega+n}$-projective if and only if G is $p^{\omega+n}$-projective.*

Proof. In accordance with Proposition 4.1, the quotient $G/(N + p^\omega G)$ is Σ-cyclic for some $N \leq G[p^n]$. Thus $p^n(N + p^\omega G) = \{0\}$ and Theorem 1.1 is manifestly applicable to obtain the claim. □

4.6. Corollary. *If G is strongly ω_1-$p^{\omega+n}$-projective, then $G/p^{\omega+n} G$ is $p^{\omega+n}$-projective.*

Proof. Follows directly from the combination of Corollaries 4.4 and 4.5. □

Remark. In ([12], Example 2.3) was constructed an example of an ω_1-$p^{\omega+n}$-projective group of length $\omega + n$ which is not $p^{\omega+n}$-projective; thereby in view of Corollary 4.5 it is not strongly ω_1-$p^{\omega+n}$-projective as well. Invoking [8], it is not even strongly n-simply presented.

Moreover, the following inclusions hold:

{separable $p^{\omega+n}$-projective groups} $\subseteq$ {$p^{\omega+n}$-projective groups} $\cap$ {separably ω_1-$p^{\omega+n}$-projective groups} $\subseteq$ {strongly ω_1-$p^{\omega+n}$-projective groups} $\subseteq$ {ω_1-$p^{\omega+n}$-projective groups} $\cap$ {strongly n-simply presented groups}.

Below we shall demonstrate that the last containment is actually tantamount to an equality - see Corollary 4.16.

On the other hand, Keef also showed in [12] that for any $n \geq 2$ there is a $p^{\omega+n}$-projective group G with the property that G is not separably ω_1-$p^{\omega+n}$-projective (see too the Example on p. 4382 of [11] where a $p^{\omega+n}$-projective group was exhibited which is not separably n-simply presented and thus not separably ω_1-$p^{\omega+n}$-projective; however every $p^{\omega+1}$-projective group is separably 1-simply presented). That is why there exists an example of a strongly ω_1-$p^{\omega+n}$-projective group that is not separably ω_1-$p^{\omega+n}$-projective (and even not separably n-simply presented) whenever $n > 1$. For $n = 1$ this will be illustrated below as well.

As a matter of fact, we begin with the following affirmation that restricts strong (separable) ω_1-$p^{\omega+1}$-projectivity to Ulm subgroups and Ulm factors.

4.7. Proposition. *The group G is strongly ω_1-$p^{\omega+1}$-projective if and only if*
(i) $p^{\omega+1}G$ is countable;
(ii) $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective.

Proof. The necessity being already established in the series of our previous assertions, we concentrate now on the sufficiency.

And so, using ([7], Corollary 2.11), the decomposition $G = K \oplus S$ holds, where K is countable and S is $p^{\omega+1}$-projective. Thus, by Theorem 1.1 (a), there is $T \leq S[p]$ with S/T being Σ-cyclic. Hence T is nice in S and so in G. Finally, $G/T \cong K \oplus (S/T)$ is the direct sum of a countable group and a Σ-cyclic group, as required in Definition 1.1. Besides, even $T \cap p^{\omega+1}G = T \cap p^{\omega+1}K \subseteq S \cap K = \{0\}$ is fulfilled. □

4.8. Proposition. *The group G is separably ω_1-$p^{\omega+1}$-projective if and only if*
(i) $p^{\omega}G$ is countable;
(ii) $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective.

Proof. The necessity being already obtained in the series of our preceding statements, we deal now with the sufficiency. And so, utilizing ([7], Corollary 2.11), one may decompose $G = L \oplus R$, where L is countable and R is separable $p^{\omega+1}$-projective. Thus, again an appeal to Theorem 1.1 (a), leads to the existence of $M \leq R[p]$ such that R/M is Σ-cyclic. Hence M is nice in R and so it is nice in G. Furthermore, $G/M \cong L \oplus (R/M)$ is the direct sum of a countable group and a Σ-cyclic group. Moreover, $M \cap p^{\omega}G = M \cap p^{\omega}L \subseteq R \cap L = \{0\}$, as required in Definition 1.2. □

As promised above, the wanted example of a strongly ω_1-$p^{\omega+1}$-projective non separably ω_1-$p^{\omega+1}$-projective group can be produced by choosing a group G whose subgroup $p^{\omega+1}G$ is countable but such that $p^{\omega}G$ is uncountable, and $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective. There exists an abundance of such groups; in fact, any $p^{\omega+1}$-projective group G with uncountable $p^{\omega}G$ may be applied in this situation. Nevertheless, each $p^{\omega+1}$-projective group G with countable $p^{\omega}G$ (in particular, each separable $p^{\omega+1}$-projective group) is separably ω_1-$p^{\omega+1}$-projective, as it will be seen below. This crucial property is due to the fact that $p^{\omega+1}$-projectives are C-decomposable (for more details see, for instance, [11] and [12]).

4.9. Proposition. *Suppose that G is a group whose $p^{\omega}G$ is countable. Then G is separably n-simply presented if and only if G is separably ω_1-$p^{\omega+n}$-projective.*

Proof. The sufficiency being trivial, we are now attack the necessity. Thus the application of [10] guarantees that $G/(M \oplus p^{\omega}G)$ is Σ-cyclic for some p^n-bounded nice subgroup M of G such that $M \cap p^{\omega}G = \{0\}$. But $G/(M \oplus p^{\omega}G) \cong G/M/(M \oplus p^{\omega}G)/M$ and since $(M \oplus p^{\omega}G)/M \cong p^{\omega}G$ is countable, the Lemma of Charles listed above applies to show that G/M is the direct sum of a countable group and a Σ-cyclic group. So, by Definition 1.2, the group G has to be separably ω_1-$p^{\omega+n}$-projective, as desired. $\square$

4.10. Proposition. *Let G be a group such that $p^\omega G$ is countable. Then G is both separably ω_1-$p^{\omega+1}$-projective and $p^{\omega+1}$-bounded if and only if G is $p^{\omega+1}$-projective.*

Proof. The necessity follows immediately from Corollary 4.5.

Concerning the sufficiency, it was proved in [11] that any $p^{\omega+1}$-projective groups belongs to the class of separably 1-simply presented groups. We now employ the preceding Proposition 4.9 to get the claim. $\square$

4.11. Corollary. *Suppose G is a group with countable $p^\omega G$. Then $G/p^{\omega+1}G$ is separably ω_1-$p^{\omega+1}$-projective if and only if $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective.*

Proof. Observe that $p^\omega(G/p^{\omega+1}G) = p^\omega G/p^{\omega+1}G$ is countable and we next apply Proposition 4.10. $\square$

4.12. Corollary. *If G is a separably ω_1-$p^{\omega+1}$-projective group and H is a subgroup such that $H \cap p^{\omega+1}G = p^{\omega+1}H$, then H is separably ω_1-$p^{\omega+1}$-projective.*

In particular, isotype subgroups of separably ω_1-$p^{\omega+1}$-projectives are separably ω_1-$p^{\omega+1}$-projective.

Proof. With the help of Proposition 4.8 write that $p^\omega G$ is countable and $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective. Hence $p^\omega H$ is countable, and $H/p^{\omega+1}H = H/(H \cap p^{\omega+1}G) \cong (H + p^{\omega+1}G)/p^{\omega+1}G \subseteq G/p^{\omega+1}G$ is $p^{\omega+1}$-projective. Consequently, again Proposition 4.8 works to get the assertion. The second half is immediate. $\square$

The above two reduction statements suggest the following stronger consideration. So we will now somewhat enlarge Propositions 4.7 and 4.8 to an arbitrary natural number $n \geq 1$ in an identical way, noticing also that Corollary 4.11 can be eventually derived from the next Theorem 4.13. In this aspect, Keef showed in [11] that a group G is separably n-simply presented if and only if $p^{\omega+n}G$ is simply presented and $G/p^{\omega+n}G$ is separably n-simply presented, while in [8] it was established that G is strongly n-simply presented if and only if $p^{\omega+n}G$ is strongly n-simply presented and $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective. Moreover, Keef proved in [12] that G is ω_1-$p^{\omega+n}$-projective if and only if $p^{\omega+n}G$ is countable and $G/p^{\omega+n}G$ is ω_1-$p^{\omega+n}$-projective.

So, keeping the similarity of the formulation, we are now able to formulate and prove our first central result of the present section.

4.13. Theorem. *(First Reduction Criterion). For every $n \geq 1$ the group G is strongly ω_1-$p^{\omega+n}$-projective if and only if*

(1) $p^{\omega+n}G$ is countable;

(2) $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective.

Proof. "$\Rightarrow$". According to Lemma 4.1, one may write that $p^\omega G/(p^\omega G \cap N)$ is countable for some p^n-bounded nice subgroup N of G. Thus $p^\omega G = p^\omega G \cap N + C$ where $C \leq p^\omega G$ is countable. Furthermore, $p^{\omega+n}G = p^nC$ is countable, so that clause (1) follows.

Next, point (2) follows directly from Corollary 4.6.

"$\Leftarrow$". Suppose that $P \leq G$ such that $p^{\omega+n}G \subseteq P$, $p^nP \subseteq p^{\omega+n}G$ (thereby $P/p^{\omega+n}G$ is p^n-bounded) and G/P is Σ-cyclic. Let Y be a maximal p^n-bounded summand of $p^\omega G$; so there is a decomposition $p^\omega G = X \oplus Y$ and thus the inclusions $X \subseteq p^\omega G \subseteq P$ hold. We may assume without loss of generality that X is countable; in fact, $p^{\omega+n}G = p^nX$ is countable and so we can decompose $X = K \oplus T$ where K is countable and T is p^n-bounded (whence T is a p^n-bounded summand of $p^\omega G$ and thereby $T \subseteq Y$; then even $T = T \cap Y \subseteq X \cap Y = \{0\}$ and $X = K$ - in any case $p^\omega G = K \oplus (T \oplus Y)$ where $T \oplus Y$ is

p^n-bounded). That is why $p^\omega G = K \oplus Y$ with a countable summand K, as desired. Another verification of this fact is like this: Note that $X[p] = (p^{\omega+n}G)[p] = (p^n X)[p]$, and hence $X[p]$ is countable. So X will be countable, provided that it is reduced.

Let us now H be a $p^{\omega+n}$-high subgroup of G containing Y (thus H is maximal with respect to $H \cap p^{\omega+n}G = \{0\}$). We next assert that $(G/p^{\omega+n}G)[p^n] = (X \oplus H[p^n])/p^{\omega+n}G$. To this aim, given $v \in G$ with $p^n v \in p^{\omega+n}G$, it suffices to prove that $v \in X \oplus H[p^n]$. If $x \in X$ is chosen such that $p^n x = p^n v$, then replacing v by $v - x$, we may assume that $p^n v = 0$. Since $G[p] = (p^{\omega+n}G)[p] \oplus H[p] = X[p] \oplus H[p]$ and H is pure in G, it easily follows that $G[p^n] = X[p^n] \oplus H[p^n]$. Therefore, $v = x' + h$ where $x' \in X[p^n]$ and $h \in H[p^n]$ as required. Moreover, $X \cap H = \{0\}$ because as noted above $X[p] = (p^{\omega+n}G)[p]$, which substantiates our assertion. Furthermore, by what we have just shown above, $P/p^{\omega+n}G \subseteq (G/p^{\omega+n}G)[p^n]$ implies that $P \subseteq X \oplus H[p^n]$. Note also the fact from above that $X \leq P$. Let $L = P \cap H[p^n] \subseteq H[p^n] \subseteq G[p^n]$; so $p^n L = \{0\}$. Clearly, the inclusion $L \subseteq H$ forces that $L \cap p^{\omega+n}G = \{0\}$. Likewise, $P \subseteq X \oplus H[p^n]$ yields that $P = X + (P \cap H[p^n]) = X + L$; indeed the modular law applies to get that $P = (X \oplus H[p^n]) \cap P = X + P \cap H[p^n]$ as stated. Consequently, we conclude that $P = p^\omega G + P = p^\omega G + L$. Thus $G/P = G/(p^\omega G + L)$ is Σ-cyclic.

We next will show that L is nice in G. Since $L \cap p^{\omega+n}G = \{0\}$, it readily follows via some technical efforts that $L \cap p^\omega G$ is nice in $p^\omega G$ and so nice in G. But $L + p^\omega G = P$ is also nice in G because $G/(p^\omega G + L)$ is separable, and these two conditions together imply that L is nice in G, as wanted (see, e.g., Section 79, Exercise 10 of [9]).

Furthermore, we claim that $p^\omega(G/L) = (p^\omega G + L)/L = P/L$ is countable. In fact, $P/L = P/(P \cap H[p^n]) \cong (P + H[p^n])/H[p^n] = (p^\omega G + H[p^n])/H[p^n] \cong p^\omega G/(p^\omega G \cap H[p^n])$. But $p^\omega G = X \oplus Y$ and since $Y \subseteq H$, one may have in view of the modular law that $p^\omega G \cap H = (X \oplus Y) \cap H = (X \cap H) \oplus Y = Y$. We therefore establish that $P/L \cong (X \oplus Y)/Y[p^n] \cong X \oplus (Y/Y[p^n]) \cong X \oplus p^n Y = X$, because $p^n Y = \{0\}$. As noticed above, X is countable, so that $p^\omega(G/L)$ is really countable as claimed. Finally, Lemma 4.1 allows us to infer that G is strongly ω_1-$p^{\omega+n}$-projective, as required. □

An immediate consequence is this one:

4.14. Proposition. *Suppose that G is a group whose $p^{\omega+n}G$ is countable. Then the following are equivalent:*

(a) G is strongly ω_1-$p^{\omega+n}$-projective;
(b) $G/p^{\omega+n}G$ is strongly ω_1-$p^{\omega+n}$-projective;
(c) $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective.

Proof. Follows by a direct application of Corollaries 4.4 and 4.5 as well as of Theorem 4.13. □

As a new valuable consequence of the First Reduction Criterion, we obtain an analog of Proposition 4.9 (see also Corollary 3.2):

4.15. Corollary. *Suppose $p^{\omega+n}G$ is countable. Then G is strongly n-simply presented if and only if G is strongly ω_1-$p^{\omega+n}$-projective.*

Proof. One direction "$\Leftarrow$" being trivial, we observe for the another one "$\Rightarrow$" that, appealing to [8], the quotient $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective. Next, the First Reduction Criterion can be applied to derive that G is strongly ω_1-$p^{\omega+n}$-projective, as formulated. □

An interesting consequence to the last statement is the following.

4.16. Corollary. *Strongly n-simply presented ω_1-$p^{\omega+n}$-projective groups are strongly ω_1-$p^{\omega+n}$-projective, and vice versa.*

Proof. The sufficiency being elementary, we will attack the necessity. Since by Theorem 1.2 (i) for each ω_1-$p^{\omega+n}$-projective group G we have that $p^{\omega+n}G$ is countable, Corollary 4.15 applies to infer that G is, in fact, strongly ω_1-$p^{\omega+n}$-projective. $\square$

4.17. Corollary. *Suppose G is a group such that $p^\omega G$ is countable. Then the following are equivalent:*

(1) G is strongly ω_1-$p^{\omega+1}$-projective;

(2) G is separably 1-simply presented;

(3) G is separably ω_1-$p^{\omega+1}$-projective.

Proof. The equivalence (1) $\iff$ (3) follows from directly Propositions 4.7 and 4.8. On the other hand the equivalence (2) $\iff$ (3) was proved in Proposition 4.9. $\square$

For $n = 1$ the alluded to above Corollary 4.15 can be slightly extended in the following way:

4.18. Corollary. *Suppose that G is a group with countable $p^{\omega+1}G$. Then the following three conditions are equivalent:*

(1) G is strongly 1-simply presented;

(2) G is separably 1-simply presented;

(3) G is strongly ω_1-$p^{\omega+1}$-projective.

Proof. For the fact that (1) is tantamount to (3) we employ Corollary 4.15.

To prove that (2) and (3) are equal, we first observe that separably 1-simply presented groups are strongly 1-simply presented and thus by what we have just shown, they are strongly ω_1-$p^{\omega+1}$-projective. So (2) implies (3). In order to verify the converse, we next apply the First Reduction Criterion to deduce that $G/p^{\omega+1}G$ is $p^{\omega+1}$-projective, whence in view of [11] this quotient must be separably 1-simply presented. Finally, again an appeal to [11] insures that G has to be separably 1-simply presented, as wanted. $\square$

Note that the last two corollaries fail for $n \geq 2$.

4.19. Corollary. *Let H be a subgroup of the strongly ω_1-$p^{\omega+n}$-projective group such that $H \cap p^{\omega+n}G = p^{\omega+n}H$. Then H is strongly ω_1-$p^{\omega+n}$-projective.*

In particular, isotype subgroups of strongly ω_1-$p^{\omega+n}$-projectives are strongly ω_1-$p^{\omega+n}$-projective.

Proof. Employing Theorem 4.13 we can write that $p^{\omega+n}G$ is countable and $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective. Thus $p^{\omega+n}H$ is countable as being a subgroup of $p^{\omega+n}G$. Moreover, $H/p^{\omega+n}H = H/(H \cap p^{\omega+n}G) \cong (H + p^{\omega+n}G)/p^{\omega+n}G \subseteq G/p^{\omega+n}G$ is $p^{\omega+n}$-projective as well. So, again the utilization of the First Reduction Criterion guarantees that H is strongly ω_1-$p^{\omega+n}$-projective, as expected. The final part is immediate. $\square$

We are now in a position and state and prove the second major result of this section.

4.20. Theorem. *(Second Reduction Criterion). For every $n \geq 1$ the group G is separably ω_1-$p^{\omega+n}$-projective if and only if*

(1) $p^\omega G$ is countable;

(2) $G/p^{\omega+n}G$ is separably ω_1-$p^{\omega+n}$-projective.

Proof. "$\Rightarrow$". That $p^\omega G$ is countable is evident in virtue of Lemma 4.2. With the aid of the same lemma write that $G/(P \oplus p^\omega G)$ is Σ-cyclic for some $P \leq G[p^n]$ with $P \cap p^\omega G = \{0\}$. But by the modular law we have

$$[(P + p^{\omega+n}G)/p^{\omega+n}G] \cap p^\omega(G/p^{\omega+n}G) = [(P + p^{\omega+n}G) \cap p^\omega G]/p^{\omega+n}G =$$

$$= [(P \cap p^\omega G) + p^{\omega+n} G]/p^{\omega+n} G = \{0\}.$$

Furthermore,

$$G/(P \oplus p^\omega G) \cong G/p^{\omega+n} G/(P \oplus p^\omega G)/p^{\omega+n} G =$$

$$= G/p^{\omega+n} G/[((P + p^{\omega+n} G)/p^{\omega+n} G) \oplus p^\omega (G/p^{\omega+n} G)]$$

is Σ-cyclic with $p^n[(P + p^{\omega+n} G)/p^{\omega+n} G] = \{0\}$ and $p^\omega (G/p^{\omega+n} G)$ countable. This verifies the necessity.

"$\Leftarrow$". Let Q be a subgroup of G containing $p^{\omega+n} G$ such that $Q/p^{\omega+n} G$ is p^n-bounded (i.e., $p^n Q \subseteq p^{\omega+n} G$), $Q \cap p^\omega G \subseteq p^{\omega+n} G$ and $G/p^{\omega+n} G/[(Q/p^{\omega+n} G) \oplus p^\omega (G/p^{\omega+n} G)] = G/p^{\omega+n} G/(Q + p^\omega G)/p^{\omega+n} G \cong G/(Q + p^\omega G)$ is Σ-cyclic. Suppose

$$Q/p^{\omega+n} G = \oplus_{i \in I} \langle x_i + p^{\omega+n} G \rangle$$

where $x_i \in Q$ and order$(x_i + p^{\omega+n} G) = p^{t_i} \leq p^n$ in $G/p^{\omega+n} G$, which is equivalent to $p^{t_i}(x_i + p^{\omega+n} G) = p^{\omega+n} G$, i.e. to $p^{t_i} x_i \in p^{\omega+n} G$, and t_i is the minimal natural number with this property. Now, for each $i \in I$, $p^{t_i} x_i \in p^{\omega+n} G = p^n(p^\omega G) = p^{t_i}(p^{\omega+n-t_i} G)$ whence $p^{t_i} x_i = p^{t_i} g_i$ for some $g_i \in p^\omega G$. Put

$$P = \oplus_{i \in I} \langle x_i - g_i \rangle$$

observing that $P \subseteq G$. Clearly, $p^n P = \{0\}$ because $p^n x_i = p^n g_i$. If now $y \in P \cap p^\omega G$, then one may write $y = a_1(x_{i_1} - g_{i_1}) + \cdots + a_k(x_{i_k} - g_{i_k})$ for some collection of indexes i_j and integers a_j, where $j = 1, \cdots, k$. This forces that

$$a_1 x_{i_1} + \cdots + a_k x_{i_k} + p^{\omega+n} G = y + a_1 g_{i_1} + \cdots + a_k g_{i_k} + p^{\omega+n} G \in$$

$$(Q/p^{\omega+n} G) \cap p^\omega (G/^{\omega+n} G) = (Q/p^{\omega+n} G) \cap (p^\omega G/p^{\omega+n} G) = (Q \cap p^\omega G)/p^{\omega+n} G = \{0\},$$

hence we have $a_1 x_{i_1} + \cdots + a_k x_{i_k} \in p^{\omega+n} G$ which ensures that $a_1(x_{i_1} + p^{\omega+n} G) = \cdots = a_k(x_{i_k} + p^{\omega+n} G) = p^{\omega+n} G$. Consequently, $p^{t_{i_j}}/a_j$ for every $j = 1, \cdots, k$ and hence

$$y = s_1 p^{t_1}(x_{i_1} - g_{i_1}) + \cdots + s_k p^{t_k}(x_{i_k} - g_{i_k}) =$$

$$s_1(p^{t_1} x_{i_1} - p^{t_1} g_{i_1}) + \cdots + s_k(p^{t_k} x_{i_k} - p^{t_k} g_{i_k}) = 0.$$

That is why $P \cap p^\omega G = \{0\}$ as expected. Finally, since $Q = \sum_{i \in I} \langle x_i + p^{\omega+n} G \rangle = \sum_{i \in I} \langle x_i \rangle + p^{\omega+n} G$, we infer that $P + p^\omega G = Q + p^\omega G$. But $G/(P \oplus p^\omega G) = G/(Q + p^\omega G)$ is Σ-cyclic and this substantiates the sufficiency in accordance with Lemma 4.2. □

Remark. It is worthwhile noticing that, unfortunately, the Second Reduction Criterion does not directly lead to the aforementioned fact from Proposition 4.9 that separably n-simply presented groups with countable first Ulm subgroup are themselves separably ω_1-$p^{\omega+n}$-projective. The reason for this contrast with the First Reduction Criterion is that separably n-simply presented groups of length $\leq \omega + n$ need not be separably ω_1-$p^{\omega+n}$-projective for any $n \geq 1$; they are just $p^{\omega+n}$-projective.

The following example illustrates that in point (2) of Theorem 4.20 the factor-group $G/p^{\omega+n}G$ cannot be replaced to be $p^{\omega+n}$-projective when $n \geq 2$ (compare also the difference with Proposition 4.8 when $n = 1$). This is so because separably ω_1-$p^{\omega+n}$-projectives of length $\leq \omega + n$ are necessarily $p^{\omega+n}$-projective but the converse fails whenever $n \geq 2$ and even for $n = 1$ provided the first Ulm subgroup is uncountable (see Proposition 4.10 too).

Example. Let A be the $p^{\omega+n}$-projective group which is not separably ω_1-$p^{\omega+n}$-projective for some $n \geq 2$, as constructed in [11], and let G be a group such that $G/p^{\omega+n}G \cong A$. We claim that G is not separably ω_1-$p^{\omega+n}$-projective because, otherwise, Corollary 4.4 would imply that so is $G/p^{\omega+n}G$ that is against our construction. The example is sustained.

However, since $G/p^{\omega+n}G$ is $p^{\omega+n}$-projective, the First Reduction Criterion, that is Theorem 4.13, assures that G is necessarily strongly ω_1-$p^{\omega+n}$-projective.

In [10] was appeared that summands of separably n-simply presented groups are again separably n-simply presented. The same idea works and for separably ω_1-$p^{\omega+n}$-projective groups, so that one may formulate without a proof the following.

4.21. Proposition. *A summand of a separably ω_1-$p^{\omega+n}$-projective group is also separably ω_1-$p^{\omega+n}$-projective.*

5. Concluding Discussion

Certainly, the major concept of strong ω_1-$p^{\omega+n}$-projectivity can be extended as follows:

Definition 5.1. A group G is called *weakly n-ω_1-$p^{\omega+n}$-projective* if there exists a subgroup $R \leq G[p^n]$ which is nice in G such that G/R is a subgroup of the direct sum of a countable group and a $p^{\omega+n}$-projective group.

It is worth noticing that, in view of Theorem 1.5 (a) from [12], G/R must be ω_1-$p^{\omega+n}$-projective. Also, the subgroup $p^{\omega+2n}G$ must be countable.

Besides, strongly n-simply presented groups of length $\leq \omega+2n$ and n-simply presented groups of length $\leq \omega + n$ are both strongly n-ω_1-$p^{\omega+n}$-projective by taking $R = p^{\omega+n}G$ or $R = p^{\omega}G$, respectively.

Another interesting variation in a more weak form of ω_1-$p^{\omega+n}$-projectivity is given in the following new concept:

Definition 5.2. A group G is said to be *nicely ω_1-$p^{\omega+n}$-projective* if it has a nice $p^{\omega+n}$-projective subgroup X such that G/X is countable.

Apparently, owing to ([12], Theorem 1.2 (c3)), nicely ω_1-$p^{\omega+n}$-projectives are themselves ω_1-$p^{\omega+n}$-projective.

The class of nicely ω_1-$p^{\omega+n}$-projectives is also worthy of investigation, which will be done in a subsequent article.

Corrigendum. In the proof of Proposition 2.3 from [7] there is a typo, namely the subgroup P of H should satisfy $p^{n+1}P = \{0\}$ instead of the written there equality $p^{\omega+n+1}P = \{0\}$.

Acknowledgments: First and foremost the author is very indebted to the colleague, Professor Patrick Keef, for the valuable communication. The author also would like to express his sincere thanks to the referees for their expert suggestions as well as to the Editor, Professor Yücel Tiras, for the time and efforts in processing this work.

References

[1] K. Benabdallah, J. Irwin and M. Rafiq, *A core class of abelian p-groups,* Sympos. Math. **13** (1974), 195–206.

[2] P. Danchev, *Countable extensions of torsion abelian groups,* Arch. Math. (Brno) **41** (3) (2005), 265–272.

[3] P. Danchev, *Primary abelian* n-Σ*-groups revisited,* Math. Pannonica **22** (1) (2011), 85–93.

[4] P. Danchev, *On weakly* ω_1-$p^{\omega+n}$*-projective abelian p-groups,* J. Indian Math. Soc. **80** (1-4) (2013), 33–46.

[5] P. Danchev, *On* ω_1*-weakly* p^{α}*-projective abelian p-groups,* Bull. Malays. Math. Sci. Soc. **37** (2014).

[6] P. Danchev and P. Keef, *Generalized Wallace theorems,* Math. Scand. **104** (1) (2009), 33–50.

[7] P. Danchev and P. Keef, *An application of set theory to* $\omega + n$*-totally* $p^{\omega+n}$*-projective primary abelian groups,* Mediterr. J. Math. **8** (4) (2011), 525–542.

[8] P. Danchev and P. Keef, *On n-simply presented primary abelian groups,* Houston J. Math. **38** (4) (2012), 1027–1050.

[9] L. Fuchs, Infinite Abelian Groups, Volumes **I** and **II**, Academic Press, New York and London 1970 and 1973.

[10] L. Fuchs and J. Irwin, *On elongations of totally projective p-groups by* $p^{\omega+n}$*-projective p-groups,* Czechoslovak Math. J. **32** (4) (1982), 511–515.

[11] P. Keef, *Elongations of totally projective groups and* $p^{\omega+n}$*-projective abelian groups,* Commun. Algebra **18** (12) (1990), 4377–4385.

[12] P. Keef, *On* ω_1-$p^{\omega+n}$*-projective primary abelian groups,* J. Alg. Numb. Th. Acad. **1** (1) (2010), 41–75.

[13] C. Megibben, *On high subgroups,* Pac. J. Math. **14** (4) (1964), 1353–1358.

[14] R. Nunke, *Purity and subfunctors of the identity,* Topics in Abelian Groups, Scott, Foresman and Co., 1962, 121–171.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 65 – 68

ON A REDUCTION FORMULA FOR THE KAMPÉ de FÉRIET FUNCTION

Yong Sup Kim*, Tibor K. Pogány† and Arjun K. Rathie‡

Received 19 : 05 : 2011 : Accepted 08 : 10 : 2012

Abstract

The aim of this short research note is to provide a reduction formula for the Kampé de Fériet function $F^{h:2;0}_{g:2;0}[-x, x]$ by employing a new summation formula for Clausen's series ${}_3F_2[1]$ obtained recently by the authors [Miskolc Math. Notes **10**(2), 145–153, 2009.]

Keywords: Clausen's series ${}_3F_2$, Euler's transformation for ${}_2F_2$, Kampé de Fériet function, Kummer–type I transformation for ${}_2F_2$, summation formula.

2000 AMS Classification: Primary 33C70; Secondary 33C15, 33C20, 33C65.

1. Introduction and results required

Recently Paris [9] established a Kummer–type I transformation formula for the generalized hypergeoemtric function ${}_2F_2[x]$, namely

$$(1.1)\qquad {}_2F_2\left[\begin{matrix} a, & c+1 \\ b, & c \end{matrix}\; ; x\right] = \mathrm{e}^x\, {}_2F_2\left[\begin{matrix} b-a-1, & f+1 \\ b, & f \end{matrix}\; ; -x\right] \qquad x \in \mathbb{C},$$

where

$$f = \frac{c(1+a-b)}{a-c}\,.$$

Equation (1.1) is seen to be analogous to the well–known and much employed Kummer's first transformation for the confluent hypergeometric function

$${}_1F_1\left[\begin{matrix} a \\ b \end{matrix}\; ; x\right] = \mathrm{e}^x\, {}_1F_1\left[\begin{matrix} b-a \\ b \end{matrix}\; ; -x\right].$$

*Department of Mathematics Education, Wonkwang University, Iksan 570–749, South Korea, Email: yspkim@wonkwang.ac.kr

†Faculty of Maritime Studies, University of Rijeka, 51000 Rijeka, Studentska 2, Croatia, Email: poganj@pfri.hr Corresponding Author.

‡Department of Mathematics, School of Mathematical and Physical Sciences, Central University of Kerala, Riverside Transit Campus, Padennakkad P.O. Nileshwar, Kasaragod - 671 328, Kerala, India, Email: akrathie@gmail.com

Paris' result (1.1) may be regarded as the generalization of the Exton's result [5], by letting $2c = a$ so that $f = 1 + a - b$, given by

$$ {}_2F_2\left[\begin{array}{c} a,\ 1+\frac{1}{2}a \\ \frac{1}{2}a \end{array};x\right] = \mathrm{e}^x\, {}_2F_2\left[\begin{array}{c} b-a-1,\ 2+a-b \\ b,\ 1+a-b \end{array};-x\right]. $$

Recently Kim *et al.* [8] have obtained a new summation formula for Clausen's ${}_3F_2[1]$ series given by

$$ {}_3F_2\left[\begin{array}{c} -n,\ b-a-1,\ f+1 \\ b,\ f \end{array};1\right] = \frac{(a)_n(c+1)_n}{(b)_n(c)_n}, \tag{1.2} $$

where $(a)_n = \Gamma(a+n)/\Gamma(n) = a(a+1)\cdots(a+n-1)$, $a \in \mathbb{C} \setminus Z_0^-$ stands for the Pochhammer symbol and f is the same as in (1.1). We note that by convention $(a)_0 = 1$.

By utilizing (1.2), Kim *et al.* [8] have obtained the following result:

$$ (1-x)^{-h}\, {}_3F_2\left[\begin{array}{c} h,\ b-a-1,\ f+1 \\ b,\ f \end{array};-\frac{x}{1-x}\right] = {}_3F_2\left[\begin{array}{c} h,\ a,\ c+1 \\ b,\ c \end{array};x\right]. $$

This result is also recorded in [10], in a slightly modified form. On the other hand, this relation may be regarded as a generalization of the following result due to Exton [5]:

$$ (1-x)^{-h}\, {}_3F_2\left[\begin{array}{c} h,\ a,\ 1+\frac{1}{2}a \\ b,\ \frac{1}{2}a \end{array};-\frac{x}{1-x}\right] = {}_3F_2\left[\begin{array}{c} h,\ b-a-1,\ 2+a-b \\ b,\ 1+a-b \end{array};x\right]. $$

On the other hand, just as the Gauss function ${}_2F_1$ was extended to generalized hypergeometric function ${}_pF_q$ by increasing the number of parameters in the numerator as well as in the denominator, the four Appell functions were introduced and generalized by Appell and Kampé de Fériet [1] who defined a general hypergeometric function in two variables. For further details see [12]. The notation defined and introduced originally by Kampé de Fériet for this double hypergeometric function of superior order was subsequently abbreviated by Burchnall and Chaundy [3]. We, however, recall here the definition of a more general double hypergeometric function (than the one defined by Kampé de Fériet) in a sligthly modified notation given by Srivastava and Panda [14, p. 423, Eq. (26)]. For this, let (H_h) denotes the sequence of parameters $(H_1, \cdots, H_h)$ and for nonnegative integers define the Pochhammer symbols $((H_h)) := (H_1)_n(H_2)_n \cdots (H_h)_n$, where when $n = 0$, the product is understood to reduce to unity. Therefore, the convenient generalization of the Kampé de Fériet function is defined as follows:

$$ F^{h:a;b}_{g:c;d}\left[\begin{array}{c} (H_h):\ (A_a);\ (B_b)\ ; \\ (G_g):\ (C_c);\ (D_d)\ ; \end{array}\ x,\ y\right] = \sum_{m,n\geq 0} \frac{((H_h))_{m+n}((A_a))_m((B_b))_n}{((G_g))_{m+n}((C_c))_m((D_d))_n}\,\frac{x^m}{m!}\,\frac{y^n}{n!}. \tag{1.3} $$

For more details about the convergence for the function (1.3) we refer to [1]. Various authors (see e.g. [1, 4, 5, 6, 7, 11, 12]) have discussed the reducibility of the Kampé de Fériet function.

The main objective of this short research note is to establish a reduction formula for the Kampé de Fériet function $F^{h:2;0}_{g:2;0}[-x, x]$ by employing the summation formula (1.2).

2. Main result

2.1. Theorem. *There holds true*

$$ F^{h:2;0}_{g:2;0}\left[\begin{array}{cccc} (H_h): & b-a-1, & f+1 & ;-; \\ (G_g): & b, & f & ;-; \end{array}\ -x,\ x\right] = {}_{h+2}F_{g+2}\left[\begin{array}{c} (H_h),\ a,\ c+1 \\ (G_g),\ b,\ c \end{array};x\right], \tag{2.1} $$

where f is given in (1.1). *Here the series* (2.1) *converges either for all $x \in \mathbb{C}$ for $g \geq h$; or inside the unit circle $|x| < 1$ when $g = h - 1$; or on the unit circle $|x| = 1$ when*

$$\Re\left\{\sum_{j=1}^{h-1} G_j - \sum_{j=1}^{h} H_j + b - a\right\} > 1\,.$$

Proof. In order to derive (2.1), we proceed as follows. Denoting the left–hand side of (2.1) by S and expressing the Kampé de Fériet function as a double series, we have

$$S = \sum_{m,n\geq 0} \frac{((H_h))_{m+n}\,(b-a-1)_m\,(f+1)_m}{((G_g))_{m+n}\,(b)_m\,(f)_m}\,\frac{(-1)^m x^{n+m}}{m!\,n!}\,.$$

Making use of the well–known Bailey–transform technique in summing up double infinite series [2]

$$\sum_{n\geq 0}\sum_{k\geq 0} A(k,n) = \sum_{n\geq 0}\sum_{k=0}^{n} A(k, n-k)\,,$$

we have, after some little algebra, using

$$(n-m)! = \frac{(-1)^m\,n!}{(-n)_m}\,,$$

that

$$S = \sum_{n\geq 0} \frac{((H_h))_n}{((G_g))_n}\,\frac{x^n}{n!}\,\sum_{m=0}^{n} \frac{(-n)_m (b-a-1)_m\,(f+1)_m}{(b)_m\,(f)_m\,m!}\,.$$

The inner–most finite series we recognize as a ${}_3F_2[1]$ expression, that is

$$S = \sum_{n\geq 0} \frac{((H_h))_n}{((G_g))_n}\,\frac{x^n}{n!}\,{}_3F_2\left[\begin{matrix} -n,\ b-a-1,\ f+1 \\ b,\quad f \end{matrix}\,;\,1\right].$$

Using (1.2) we have

$$S = \sum_{n\geq 0} \frac{((H_h))_n}{((G_g))_n}\cdot\frac{(a)_n\,(c+1)_n}{(b)_n\,(c)_n}\cdot\frac{x^n}{n!}\,,$$

which gives in fact the right–hand side of the series (2.1).

By conditions that hold for the generalized hypergeometric function we easily conclude the stated convergence constraints. □

3. Special cases

3.1. In (2.1), if we take $2c = a$, so that $f = 1 + a - b$, we get the following result due to Exton [5]:

$$F^{h:2;0}_{g:2;0}\left[\begin{matrix} (H_h): & b-a-1, & 2+a-b & ;-; \\ (G_g): & b, & 1+a-b & ;-; \end{matrix}\ -x,\,x\right] = {}_{h+2}F_{g+2}\left[\begin{matrix} (H_h),\ a,\ \frac{1}{2}a+1 \\ (G_g),\ \frac{1}{2}a,\quad b \end{matrix}\,;\,x\right],$$

where the series converges under the same conditions which hold for (2.1).

3.2. If we take $b = c+1$, so that $f = c$, we arrive at the following result:

$$F^{h:1;0}_{g:1;0}\left[\begin{array}{lll} (H_h): & c-a & ;-; \\ (G_g): & c & ;-; \end{array} -x,\, x\right] = {}_{h+1}F_{g+1}\left[\begin{array}{l} (H_h),\ a \\ (G_g),\ c \end{array}; x\right],$$

where the series converges under the same conditions which hold for (2.1), exception is the convergence for $g = h-1$ on the unit circle $|x| = 1$ which follows for

$$\Re\left\{\sum_{j=1}^{h-1} G_j - \sum_{j=1}^{h} H_j + c - a\right\} > 0\,.$$

3.3. Finally, if we take $(H) = (G)$ and $h = g = 0$, we arrive at Paris' result (1.1). In this case, the formula is valid in the whole complex plane $\mathbb{C}$.

Acknowledgement. The research work of Yong–Sup Kim is supported by Wonkwang University, Iksan, South Korea (2014).

References

[1] Appell, P. and Kampé de Fériet, J. *Fonctions hypergéométrique. Polynômes d'Hermite*, (Gautier–Villars, Paris, 1926).

[2] Bailey, W.N. *Generalized Hypergeometric Series*, Cambridge Tract, No. 32. (Cambridge University Press, Cambridge, 1935).

[3] Burchnall, J.L. and Chaundy, T.W. *Expansions of Appell's double hypergeometric functions*, Quart. J. Math. (Oxford Ser.) **11**, 249–270, 1940.

[4] Buschmann R.G. and Srivastava, H.M. *Some identities and reducibility of Kampé de Fériet function*, Math. Proc. Cambridge Philos. Soc. **91**, 435–440, 1982.

[5] Exton, H. *On the reducibility of Kampé de Fériet functions*, J. Comput. Appl. Math. **83**, 119–121, 1997.

[6] Exton H. and Krupnikov, E.D. *A register of computer–oriented reduction identities for the Kampé de Fériet function.* Draft manuscript. (Novosibirsk, Russia, 1998).

[7] Karlsson, P.W. *Some reduction formulae for power series and Kampé de Fériet functions*, Nederl. Akad. Wetensch. Indag. Math. **46**(1), 31–36, 1984.

[8] Kim,Y.–S., Pogány, T.K. and Rathie, A.K. *On a summation formula for the Clausen's series ${}_3F_2$ with applications*, Miskolc Math. Notes **10**(2), 145–153, 2009.

[9] Paris, R.B. *A Kummer–type transformation for a ${}_2F_2$ hypergeometric function*, J. Comput. Appl. Math. **173**, 379–382, 2005.

[10] Rathie, A.K. and Paris, R.B. *An extension of the Euler–type transformation for the ${}_3F_2$ series*, Far East J. Math. Sci.(FJMS) **27**(1), 43–48, 2007.

[11] Slater, L.J. *Generalized Hypergeometric Functions*, (Cambridge University Press, Cambridge, 1966).

[12] Srivastava H.M. and Karlsson, P.W. *Multiple Gaussian Hypergeometric Series*, (Halsted Press (Ellis Norwood Limited, Chichester), John Wiley & Sons, New York, 1985).

[13] Srivastava, H.M. and Manocha, H.L. *A Treatise on Generating Functions*, (Halsted Press (Ellis Norwood Limited, Chichester), John Wiley & Sons, New York, 1984).

[14] Srivastava, H.M. and Panda, R. *An integral representation for the product of two Jacobi polynomials*, J. London Math. Soc. (2) **12**, 419–425, 1976.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 69 – 83

GENERALIZED SKEW DERIVATIONS ON MULTILINEAR POLYNOMIALS IN RIGHT IDEALS OF PRIME RINGS

E. Albaş[*], N. Argaç[†], V. De Filippis[‡] and Ç. Demir[§]

Received 09 : 06 : 2011 : Accepted 18 : 12 : 2012

Abstract

Let R be a prime ring, $f(x_1, \ldots, x_n)$ a multilinear polynomial over C in n noncommuting indeterminates, I a nonzero right ideal of R, and $F : R \to R$ be a nonzero generalized skew derivation of R.
Suppose that $F(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) \in C$, for all $r_1, \ldots, r_n \in I$. If $f(x_1, \ldots, x_n)$ is not central valued on R, then either $char(R) = 2$ and R satisfies s_4 or one of the following holds:

(i) $f(x_1, \ldots, x_n)x_{n+1}$ is an identity for I;
(ii) $F(I)I = (0)$;
(iii) $[f(x_1, \ldots, x_n), x_{n+1}]x_{n+2}$ is an identity for I, there exist $b, c, q \in Q$ with q an invertible element such that $F(x) = bx - qxq^{-1}c$ for all $x \in R$, and $q^{-1}cI \subseteq I$. Moreover, in this case either $(b - c)I = (0)$ or $b - c \in C$ and $f(x_1, \ldots, x_n)^2$ is central valued on R.

Keywords: Identity, generalized skew derivation, automorphism, (semi-)prime ring.

2000 AMS Classification: 16W25, 16N60.

[*]Department of Mathematics, Science Faculty, Ege University, 35100, Bornova, Izmir, Turkey, Email: `emine.albas@ege.edu.tr`
[†]Department of Mathematics, Science Faculty, Ege University, 35100, Bornova, Izmir, Turkey, Email: `nurcan.argac@ege.edu.tr`
[‡]Di.S.I.A., Faculty of Engineering, University of Messina, 98166 Messina, Italy, Email: `defilippis@unime.it`
[§]Department of Mathematics, Science Faculty, Ege University, 35100, Bornova, Izmir, Turkey, Email: `cagri.demir@ege.edu.tr`

1. Introduction.

Throughout this paper, unless specially stated, K denotes a commutative ring with unit, R is always a prime K-algebra with center $Z(R)$, right Martindale quotient ring Q and extended centroid C. The definition, axiomatic formulations and properties of this quotient ring can be found in [2] (Chapter 2).

Many results in literature indicate how the global structure of a ring R is often tightly connected to the behaviour of additive mappings defined on R. A well known result of Posner [32] states that if d is a derivation of R such that $[d(x), x] \in Z(R)$, for any $x \in R$, then either $d = 0$ or R is commutative. Later in [3], Bresar proved that if d and δ are derivations of R such that $d(x)x - x\delta(x) \in Z(R)$, for all $x \in R$, then either $d = \delta = 0$ or R is commutative. In [29], Lee and Wong extended Bresar's result to the Lie case. They proved that if $d(x)x - x\delta(x) \in Z(R)$, for all x in some non-central Lie ideal L of R then either $d = \delta = 0$ or R satisfies s_4, the standard identity of degree 4.

Recently in [28], Lee and Zhou considered the case when the derivations d and δ are replaced respectively by the generalized derivations H and G, and proved that if $R \neq M_2(GF(2))$, H, G are two generalized derivations of R, and m, n are two fixed positive integers, then $H(x^m)x^n = x^nG(x^m)$ for all $x \in R$ if and only if the following two conditions hold: (1) There exists $w \in Q$ such that $H(x) = xw$ and $G(x) = wx$ for all $x \in R$; (2) either $w \in C$, or x^m and x^n are C-dependent for all $x \in R$.

More recently in [5], a similar situation is examined: more precisely it is proved that if $H(u^n)u^n + u^nG(u^n) \in C$, for all $u \in L$, a non-central Lie ideal of R, then there exists $a \in Q$ such that $H(x) = xa$, $G(x) = -ax$, or R satisfies the standard identity s_4. Moreover in this last case a complete description of H and G is given.

Finally, as a partial extension of the above results to the case of derivations and generalized derivations acting on multilinear polynomials, we have the following:

1.1. Fact. (Theorem 2 in [27]) Let R be a prime ring, $f(x_1, \ldots, x_n)$ a multilinear polynomial over C in n noncommuting indeterminates, and $d : R \to R$ a nonzero derivation of R. If $d(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) \in C$, for all $r_1, \ldots, r_n \in R$ and $f(x_1, \ldots, x_n)$ is not central valued on RC, then $char(R) = 2$ and R satisfies s_4.

1.2. Fact. (Lemma 3 in [1]) Let R be a prime ring, $f(x_1, \ldots, x_n)$ a noncentral multilinear polynomial over C in n noncommuting indeterminates, and $G : R \to R$ a nonzero generalized derivation of R. If $G(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) \in C$, for all $r_1, \ldots, r_n \in R$, then either $char(R) = 2$ and R satisfies s_4 or there exists $b \in C$ such that $G(x) = bx$ for all $x \in R$ and $f(x_1, \ldots, x_n)^2$ is central valued on R.

These facts in a prime K-algebra are natural tests which evidence that, if d is a derivation of R and G is a generalized derivation of R, then the sets $\{d(x)x \mid x \in S\}$ and $\{G(x)x \mid x \in S\}$ are rather large in R, where S is either a non-central Lie ideal of R, or the set of all the evaluations of a non-central multilinear polynomial over K.

In this paper we will continue the study of the set

$$\{F(f(x_1, \ldots, x_n))f(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in R\}$$

for a generalized skew derivation F of R instead of a generalized derivation, and for a multilinear polynomial $f(x_1, \ldots, x_n)$ in n noncommuting variables over C. For the sake of clearness and completeness we now recall the definition of a generalized skew derivation of R. Let R be an associative ring and α be an automorphism of R. An additive mapping $d : R \longrightarrow R$ is called a *skew derivation* of R if

$$d(xy) = d(x)y + \alpha(x)d(y)$$

for all $x, y \in R$. The automophism α is called an *associated automorphism* of d. An additive mapping $F : R \longrightarrow R$ is said to be a *generalized skew derivation* of R if there exists a skew derivation d of R with associated automorphism α such that

$$F(xy) = F(x)y + \alpha(x)d(y)$$

for all $x, y \in R$, and d is said to be an *associated skew derivation* of F and α is called an *associated automorphism* of F. For fixed elements a and b of R, the mapping $F : R \to R$ defined as $F(x) = ax - \sigma(x)b$ for all $x \in R$ is a generalized skew derivation of R. A generalized skew derivation of this form is called an *inner* generalized skew derivation. The definition of generalized skew derivations is a unified notion of skew derivation and generalized derivation, which have been investigated by many researchers from various view points (see [8, 9, 10], [11], [26]).

The main result of this paper is the following:

1. Theorem. Let R be a prime ring, $f(x_1, \dots, x_n)$ a multilinear polynomial over C in n noncommuting indeterminates, I a nonzero right ideal of R, and $F : R \to R$ a nonzero generalized skew derivation of R.

Suppose that $F(f(r_1, \dots, r_n))f(r_1, \dots, r_n) \in C$, for all $r_1, \dots, r_n \in I$. If the polynomial $f(x_1, \dots, x_n)$ is not central valued on R, then either $char(R) = 2$ and R satisfies s_4 or one of the following holds:

- (i) $f(x_1, \dots, x_n)x_{n+1}$ is an identity for I;
- (ii) $F(I)I = (0)$;
- (iii) $[f(x_1, \dots, x_n), x_{n+1}]x_{n+2}$ is an identity for I, there exist $b, c, q \in Q$ with q an invertible element such that $F(x) = bx - qxq^{-1}c$ for all $x \in R$, and $q^{-1}cI \subseteq I$. Moreover, in this case either $(b - c)I = (0)$ or $b - c \in C$ and $f(x_1, \dots, x_n)^2$ is central valued on R.

It is well known that automorphisms, derivations and skew derivations of R can be extended to Q. Chang in [8] extended the definition of a generalized skew derivation to the right Martindale quotient ring Q of R as follows: by a (right) generalized skew derivation we mean an additive mapping $F : Q \longrightarrow Q$ such that $F(xy) = F(x)y + \alpha(x)d(y)$ for all $x, y \in Q$, where d is a skew derivation of R and α is an automorphism of R. Moreover, there exists $F(1) = a \in Q$ such that $F(x) = ax + d(x)$ for all $x \in R$ (Lemma 2 in [8]).

2. X-inner Generalized Skew Derivations on Prime Rings.

In this section we consider the case when F is an X-inner generalized skew derivation induced by the elements $b, c \in R$, that is, $F(x) = bx - \alpha(x)c$ for all $x \in R$, where $\alpha \in Aut(R)$ is the associated automorphism of F. Here $Aut(R)$ denotes the group of automorphisms of R.

At the outset, we will study the case when $R = M_m(K)$ is the algebra of $m \times m$ matrices over a field K. Notice that the set $f(R) = \{f(r_1, \dots, r_n) : r_1, \dots, r_n \in R\}$ is invariant under the action of all inner automorphisms of R. Hence if we denote $r = (r_1, \dots, r_n) \in R \times \dots \times R = R^n$, then for any inner automorphism φ of $M_m(K)$, we have that $\underline{r} = (\varphi(r_1), \dots, \varphi(r_n)) \in R^n$ and $\varphi(f(r)) = f(\underline{r}) \in f(R)$.

Let us recall some results from [23] and [30]. Let T be a ring with 1 and let $e_{ij} \in M_m(T)$ be the matrix unit having 1 in the (i, j)-entry and zero elsewhere. For a sequence $u = (A_1, \dots, A_n)$ in $M_m(T)$ the value of u is defined to be the product $|u| = A_1A_2 \cdots A_n$ and u is nonvanishing if $|u| \neq 0$. For a permutation σ of $\{1, 2, \cdots, n\}$ we write $u^\sigma = (A_{\sigma(1)}, \dots, A_{\sigma(n)})$. We call u *simple* if it is of the form $u = (a_1e_{i_1j_1}, \dots, a_ne_{i_nj_n})$, where $a_i \in T$. A simple sequence u is called *even* if for some σ, $|u^\sigma| = be_{ii} \neq 0$, and *odd* if for some σ, $|u^\sigma| = be_{ij} \neq 0$, where $i \neq j$ and $b \in T$. We have:

2.1. Fact. (Lemma in [23]) Let T be a K-algebra with 1 and let $R = M_m(T)$, $m \geq 2$. Suppose that $h(x_1, \dots, x_n)$ is a multilinear polynomial over K such that $h(u) = 0$ for all odd simple sequences u. Then $h(x_1, \dots, x_n)$ is central valued on R.

2.2. Fact. (Lemma 2 in [30]) Let T be a K-algebra with 1 and let $R = M_m(T)$, $m \geq 2$. Suppose that $h(x_1, \dots, x_n)$ is a multilinear polynomial over K. Let $u = (A_1, \dots, A_n)$ be a simple sequence from R.

1. If u is even, then $h(u)$ is a diagonal matrix.
2. If u is odd, then $h(u) = ae_{pq}$ for some $a \in T$ and $p \neq q$.

2.3. Fact. Suppose that $f(x_1, \dots, x_n)$ is a multilinear polynomial over a field K not central valued on $R = M_m(K)$. Then by Fact 2.1 there exists an odd simple sequence $r = (r_1, \dots, r_n)$ from R such that $f(r) = f(r_1, \dots, r_n) \neq 0$. By Fact 2.2, $f(r) = \beta e_{pq}$, where $0 \neq \beta \in K$ and $p \neq q$. Since $f(x_1, \dots, x_n)$ is a multilinear polynomial and K is a field, we may assume that $\beta = 1$. Now, for distinct i and j, let $\sigma \in S_n$ be such that $\sigma(p) = i$ and $\sigma(q) = j$, and let ψ be the automorphism of R defined by $\psi(\sum_{s,t} \xi_{st} e_{st}) = \sum_{s,t} \xi_{st} e_{\sigma(s)\sigma(t)}$. Then $f(\psi(r)) = f(\psi(r_1), \dots, \psi(r_n)) = \psi(f(r)) = \beta e_{ij} = e_{ij}$.

In all that follows we always assume that $f(x_1, \dots, x_n)$ is not central valued on R.

2.4. Lemma. *Let $R = M_m(K)$ be the algebra of $m \times m$ matrices over the field K and $m \geq 2$, $f(x_1, \dots, x_n)$ a multilinear polynomial over K, which is not central valued on R. If there exist $b, c, q \in R$ with q an invertible matrix such that*

$$\Big(bf(r_1, \dots, r_n) - qf(r_1, \dots, r_n)q^{-1}c\Big)f(r_1, \dots, r_n) \in Z(R)$$

for all $r_1, \dots, r_n \in R$, then either $char(R) = 2$ and $m = 2$, or $q^{-1}c, b - c \in Z(R)$ and $f(x_1, \dots, x_n)^2$ is central valued on R, provided that $b \neq c$.

Proof. If $q^{-1}c \in Z(R)$ then the conclusion follows from Fact 1.2. Thus we may assume that $q^{-1}c$ is not a scalar matrix and proceed to get a contradiction. Say $q = \sum_{hl} q_{hl} e_{hl}$ and $q^{-1}c = \sum_{hl} p_{hl} e_{hl}$, for $q_{hl}, p_{hl} \in K$. By Fact 2.3, $e_{ij} \in f(R)$ for all $i \neq j$, then for any $i \neq j$

$$X = (be_{ij} - qe_{ij}q^{-1}c)e_{ij} \in Z(R).$$

By X, we have $qe_{ij}q^{-1}ce_{ij} = qp_{ji}e_{ij} \in Z(R)$. Then for any $1 \leq k \leq m$ $[qp_{ji}e_{ij}, e_{ik}] = 0$, that is $q_{ki}p_{ji} = 0$. Since q is invertible $q_{k_0 i} \neq 0$ for some k_0, we get $p_{ji} = 0$ for all $i \neq j$. Hence $q^{-1}c$ is a diagonal matrix in R. Let $i \neq j$ and $\varphi(x) = (1 + e_{ji})x(1 - e_{ji})$ be an automorphism of R. It is well known that $\varphi(f(r_i)) \in f(R)$, then

$$\Big(\varphi(b)u - \varphi(q)u\varphi(q^{-1}c)\Big)u \in Z(R)$$

for all $u \in f(R)$. By the above argument, $\varphi(q^{-1}c)$ is a diagonal matrix, that is the (j, i)-entry of $\varphi(q^{-1}c)$ is zero. By calculations it follows $p_{ii} = p_{jj}$, and we get the contradiction that $q^{-1}c$ is central in R. $\square$

2.5. Lemma. *Let R be a prime ring, $f(x_1, \dots, x_n)$ be a non-central multilinear polynomial over C. If there exist $b, c, q \in R$ with q an invertible element such that*

$$(bf(r_1, \dots, r_n) - qf(r_1, \dots, r_n)q^{-1}c)f(r_1, \dots, r_n) \in C$$

for all $r_1, \dots, r_n \in R$, then either $char(R) = 2$ and R satisfies s_4, or $q^{-1}c, b - c \in Z(R)$ and $f(x_1, \dots, x_n)^2$ is central valued on R, provided that $b \neq c$.

Proof. Consider the generalized polynomial

$$\Phi(x_1,\dots,x_{n+1}) = \Big[\Big(bf(x_1,\dots,x_n) - qf(x_1,\dots,x_n)q^{-1}c\Big)f(x_1,\dots,x_n), x_{n+1}\Big]$$

which is a generalized polynomial identity for R. If $\{1, q^{-1}c\}$ is linearly C-dependent, then $q^{-1}c \in C$. In this case R satisfies

$$\Phi(x_1,\dots,x_{n+1}) = \Big[\Big((b-c)f(x_1,\dots,x_n)\Big)f(x_1,\dots,x_n), x_{n+1}\Big]$$

and we are done by Fact 1.2.

Hence we here assume that $\{1, q^{-1}c\}$ is linearly C-independent. In this case $\Phi(x_1,\dots,x_{n+1})$ is a non-trivial generalized polynomial identity for R and by [12] $\Phi(x_1,\dots,x_{n+1})$ is a non-trivial generalized polynomial identity for Q. By Martindale's theorem in [31], Q is a primitive ring having nonzero socle with the field C as its associated division ring. By [20] (p. 75) Q is isomorphic to a dense subring of the ring of linear transformations of a vector space V over C, containing nonzero linear transformations of finite rank. Assume first that $dim_C V = k$ a finite integer. Then $Q \cong M_k(C)$ and the conclusion follows from Lemma 2.4. Therefore we may assume that $dim_C V = \infty$. As in Lemma 2 in [33], the set $f(R) = \{f(r_1,\dots,r_n) : r_i \in R\}$ is dense in R and so from $\Phi(r_1,\dots,r_{n+1}) = 0$ for all $r_1,\dots,r_{n+1} \in R$, we have that Q satisfies the generalized identity

$$\big[\big(bx_1 - qx_1q^{-1}c\big)x_1, x_2\big].$$

In particular for $x_1 = 1$, $[b-c, x_2]$ is an identity for Q, that is $b - c \in C$, say $b = c + \lambda$ for some $\lambda \in C$. Thus Q satisfies

$$\big[\big((c+\lambda)x_1 - qx_1q^{-1}c\big)x_1, x_2\big]$$

and by replacing x_1 with $y_1 + t_1$ we have that

$$\Big[\Big((c+\lambda)y_1 - qy_1q^{-1}c\Big)t_1, x_2\Big] + \Big[\Big((c+\lambda)t_1 - qt_1q^{-1}c\Big)y_1, x_2\Big]$$

is an identity for Q. Once again for $y_1 = 1$ it follows that Q satisfies

$$\big[\lambda t_1 + (c+\lambda)t_1 - qt_1q^{-1}c, x_2\big]$$

and for $x_2 = t_1$

$$\big[ct_1 - qt_1q^{-1}c, t_1\big].$$

By Lemma 3.2 in [17] (or [18] Theorem 1) and since R cannot satisfy any polynomial identity ($dim_C V = \infty$), it follows the contradiction $q^{-1}c \in C$. □

2.6. Proposition. *Let R be a prime ring, $f(x_1,\dots,x_n)$ a non-central multilinear polynomial over C in n non-commuting variables, $b, c \in R$ and $\alpha \in Aut(R)$ such that $F(x) = bx - \alpha(x)c$ for all $x \in R$. If $F(f(r_1,\dots,r_n))f(r_1,\dots,r_n) \in C$, for all $r_1,\dots,r_n \in R$, and F is nonzero on R, then either $char(R) = 2$ and R satisfies s_4, or $f(x_1,\dots,x_n)^2$ is central valued on R and there exists $\gamma \in C$ such that $F(x) = \gamma x$, for all $x \in R$. When this last case occurs, we have:*

- *(i) if α is X-outer then $\gamma = b$ and $c = 0$;*
- *(ii) if $\alpha(x) = qxq^{-1}$ for all $x \in R$ and for some invertible element $q \in Q$, then $\gamma = b - c$ and $q^{-1}c \in C$.*

Proof. In case α is an X-inner automorphism of R, there exists an invertible element $q \in Q$ such that $\alpha(x) = qxq^{-1}$ for all $x \in R$ and the conclusion follows from Lemma 2.5. So we may assume here that α is X-outer. Since by [14] R and Q satisfy the same generalized identities with automorphisms, then

$$\Phi(x_1,\dots,x_{n+1}) = \big[\big(bf(x_1,\dots,x_n) - \alpha(f(x_1,\dots,x_n))c\big)f(x_1,\dots,x_n), x_{n+1}\big]$$

is satisfied by Q, moreover Q is a centrally closed prime C-algebra. Note that if $c = 0$ we are done by Fact 1.2. Thus we may assume $c \neq 0$. In this case, by [13] (main Theorem), $\Phi(x_1, \ldots, x_{n+1})$ is a non-trivial generalized identity for R and for Q. By Theorem 1 in [21], RC has non-zero socle and Q is primitive. Moreover, since α is an outer automorphism and any $(x_i)^\alpha$-word degree in $\Phi(x_1, \ldots, x_n)$ is equal to 1, then by Theorem 3 in [14], Q satisfies the identity

$$\left[\left(bf(x_1, \ldots, x_n) - f^\alpha(y_1, \ldots, y_n)c\right)f(x_1, \ldots, x_n), x_{n+1}\right],$$

where $f^\alpha(X_1, \ldots, X_n)$ is the polynomial obtained from f by replacing each coefficient γ of f with $\alpha(\gamma)$. By Fact 1.2 we conclude that either $char(R) = 2$ and R satisfies s_4 or $b, c \in C$ and $f(x_1, \ldots, x_n)^2$ is central valued on R. Moreover, in this last case we also have that Q satisfies

$$c\left[f(y_1, \ldots, y_n)f(x_1, \ldots, x_n), x_{n+1}\right].$$

Since $c \neq 0$ we have $[f(y_1, \ldots, y_n)f(x_1, \ldots, x_n), x_{n+1}]$ is a polynomial identity for Q. Thus there exists a suitable field K such that Q and the $l \times l$ matrix ring $M_l(K)$ satisfy the same polynomial identities by Lemma 1 in [22]. In particular, $M_l(K)$ satisfies $[f(y_1, \ldots, y_n)f(x_1, \ldots, x_n), x_{n+1}]$. Hence, since $f(x_1, \ldots, x_n)$ is not central valued on $M_l(K)$ (and hence $l \geq 2$), by Fact 2.3 we have that for all $i \neq j$ there exist $r_1, \ldots, r_n, s_1, \ldots, s_n \in M_l(K)$ such that $f(r_1, \ldots, r_n) = e_{ij}$ and $f(s_1, \ldots, s_n) = e_{ji}$. As a consequence we get $0 = [e_{ij}e_{ji}, x_{n+1}] = [e_{ii}, x_{n+1}]$, which is a contradiction for a suitable choice of $x_{n+1} \in M_l(K)$ (for example $x_{n+1} = e_{ij}$). $\square$

2.7. Fact. (Theorem 1 in [15]) Let R be a prime ring, D be an X-outer skew derivation of R and α be an X-outer automorphism of R. If $\Phi(x_i, D(x_i), \alpha(x_i))$ is a generalized polynomial identity for R, then R also satisfies the generalized polynomial identity $\Phi(x_i, y_i, z_i)$, where x_i, y_i and z_i are distinct indeterminates.

We close this section by collecting the results we obtained so far in the following

2.8. Proposition. *Let R be a prime ring, $f(x_1, \ldots, x_n)$ a non-central multilinear polynomial over C in n non-commuting variables, $F : R \to R$ a nonzero X-inner generalized skew derivation of R.*

If $F(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) \in C$, for all $r_1, \ldots, r_n \in R$, then either $char(R) = 2$ and R satisfies s_4, or $f(x_1, \ldots, x_n)^2$ is central valued on R and there exists $\gamma \in C$ such that $F(x) = \gamma x$, for all $x \in R$.

Proof. We can write $F(x) = bx + d(x)$ for all $x \in R$ where $b \in Q$ and d is a skew derivation of R (see [8]). We denote $f(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} \gamma_\sigma x_{\sigma(1)} \cdots x_{\sigma(n)}$ with $\gamma_\sigma \in C$. By Theorem 2 in [15] R and Q satisfy the same generalized polynomial identities with a single skew derivation, then Q satisfies

$$\left[\left(bf(x_1, \ldots, x_n) + d(f(x_1, \ldots, x_n))\right)f(x_1, \ldots, x_n), x_{n+1}\right]. \tag{2.1}$$

Since F is X-inner then d is X-inner, that is there exist $c \in Q$ and $\alpha \in Aut(Q)$ such that $d(x) = cx - \alpha(x)c$, for all $x \in R$. Hence $F(x) = (b + c)x - \alpha(x)c$ and we conclude by Proposition 2.6. $\square$

2.9. Corollary. *Let R be a prime ring, $f(x_1, \ldots, x_n)$ a non-vanishing multilinear polynomial over C in n non-commuting variables, $F : R \to R$ a non-zero X-inner generalized skew derivation of R. If $F(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) = 0$, for all $r_1, \ldots, r_n \in R$, then $char(R) = 2$ and R satisfies s_4.*

3. Generalized Skew Derivations on Right Ideals.

We premit the following:

3.1. Fact. (Main Theorem in [1]) Let R be a prime ring, I a nonzero right ideal of R, $f(x_1,\dots,x_n)$ a multilinear polynomial over C in n non-commuting indeterminates, which is not an identity for R, and $g: R \to R$ a nonzero generalized derivation of R with the associated derivation $d: R \to R$, that is $g(x) = ax + d(x)$, for all $x \in R$ and a fixed $a \in Q$.

Suppose that $g(f(r_1,\dots,r_n))f(r_1,\dots,r_n) \in C$, for all $r_1,\dots,r_n \in I$. Then either $char(R) = 2$ and R satisfies s_4 or $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I, or there exist $b, c \in Q$ such that $g(x) = bx + xc$ for all $x \in R$ and one of the following holds:

(*i*) $b, c \in C$ and $f(x_1,\dots,x_n)^2$ is central valued on R;
(*ii*) there exists $\lambda \in C$ such that $b = \lambda - c$ and $f(x_1,\dots,x_n)$ is central valued on R;
(*iii*) $(b + c)I = (0)$ and I satisfies the identity $[f(x_1,\dots,x_n), x_{n+1}]x_{n+2}$;
(*iv*) $(b + c)I = (0)$ and there exists $\gamma \in C$ such that $(c - \gamma)I = (0)$.

3.2. Fact. (Theorem 1 in [1]) Under the same situation as in above Fact, we notice that in case $g(f(r_1,\dots,r_n))f(r_1,\dots,r_n) = 0$, for all $r_1,\dots,r_n \in I$, the conclusions (*i*) and (*ii*) cannot occur. Hence we have that either $char(R) = 2$ and R satisfies s_4 or $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I, or there exist $b, c \in Q$ such that $g(x) = bx + xc$ for all $x \in R$ and one of the following holds:

(*i*) $(b + c)I = (0)$ and I satisfies the identity $[f(x_1,\dots,x_n), x_{n+1}]x_{n+2}$;
(*ii*) $(b + c)I = (0)$ and there exists $\gamma \in C$ such that $(c - \gamma)I = (0)$.

3.3. Proposition. *Let R be a prime ring, $f(x_1,\dots,x_n)$ a non-central multilinear polynomial over C in n non-commuting indeterminates, I a nonzero right ideal of R, $F: R \to R$ an X-outer generalized skew derivation of R. If*

$$F(f(r_1,\dots,r_n))f(r_1,\dots,r_n) \in C, \tag{3.1}$$

for all $r_1,\dots,r_n \in I$, then either $char(R) = 2$ and R satisfies $s_4(x_1,\dots,x_4)$, or $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I.

Proof. As above we write $F(x) = bx + d(x)$ for all $x \in R$, $b \in Q$ and d is an X-outer skew derivation of R. Let $\alpha \in Aut(Q)$ be the automorphism which is associated with d. Notice that in case α is the identity map on R, then d is a usual derivation of R and so F is a generalized derivation of R. Therefore by Fact 3.1 we obtain the required conclusions. Hence in what follows we always assume that $\alpha \neq 1 \in Aut(R)$.

We denote by $f^d(x_1,\dots,x_n)$ the polynomial obtained from $f(x_1,\dots,x_n)$ by replacing each coefficient γ_σ with $d(\gamma_\sigma)$. Notice that

$$\begin{aligned} d\big(\gamma_\sigma x_{\sigma(1)}\cdots x_{\sigma(n)}\big) &= d(\gamma_\sigma)x_{\sigma(1)}\cdots x_{\sigma(n)} \\ &\quad + \alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(x_{\sigma(1)}\cdots x_{\sigma(j)})d(x_{\sigma(j+1)})x_{\sigma(j+2)}\cdots x_{\sigma(n)} \end{aligned}$$

so that

$$\begin{aligned} d(f(x_1,\dots,x_n)) &= f^d(x_1,\dots,x_n) \\ &\quad + \sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(x_{\sigma(1)}\cdots x_{\sigma(j)})d(x_{\sigma(j+1)})x_{\sigma(j+2)}\cdots x_{\sigma(n)}. \end{aligned}$$

Since IQ satisfies (3.1), then for all $0 \neq u \in I$, Q satisfies

$$\Big[\Big(bf(ux_1,\dots,ux_n)+f^d(ux_1,\dots,ux_n)\Big)f(ux_1,\dots,ux_n),x_{n+1}\Big]$$
$$+\Big[\Big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(ux_{\sigma(1)}\dots ux_{\sigma(j)})d(ux_{\sigma(j+1)})ux_{\sigma(j+2)}\dots ux_{\sigma(n)}\Big)f(ux_1,\dots,ux_n),x_{n+1}\Big].$$

By Theorem 1 in [15], Q satisfies

$$\Big[\big(bf(ux_1,\dots,ux_n)+f^d(ux_1,\dots,ux_n)\big)f(ux_1,\dots,ux_n),x_{n+1}\Big]$$
$$+\Big[\big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(ux_{\sigma(1)}\dots ux_{\sigma(j)})d(u)x_{\sigma(j+1)}\dots ux_{\sigma(n)}\Big)f(ux_1,\dots,ux_n),x_{n+1}\Big]$$
$$+\Big[\Big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(ux_{\sigma(1)})\dots ux_{\sigma(j)})\alpha(u)y_{\sigma(j+1)}ux_{\sigma(j+2)}\dots ux_{\sigma(n)}\Big)f(ux_1,\dots,ux_n),x_{n+1}\Big].$$

In particular Q satisfies

(3.2)
$$\Big[\Big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(ux_{\sigma(1)}\dots ux_{\sigma(j)})\alpha(u)y_{\sigma(j+1)}ux_{\sigma(j+2)}\dots ux_{\sigma(n)}\Big)f(ux_1,\dots,ux_n),x_{n+1}\Big].$$

Here we suppose that either $char(R)\neq 2$ or R does not satisfy s_4, moreover $f(x_1,\dots,x_n)x_{n+1}$ is not an identity for I, if not we are done. Hence suppose there exist $a_1,\dots,a_{n+1}\in I$ such that $f(a_1,\dots,a_n)a_{n+1}\neq 0$. We proceed to get a number of contradictions.

Since $0\neq\alpha(u)$ is a fixed element of Q, we notice that (3.2) is a non-trivial generalized polynomial identity for Q, then Q has nonzero socle H which satisfies the same generalized polynomial identities of Q (see [12]). In order to prove our result, we may replace Q by H, and by Lemma 1 in [19], we may assume that Q is a regular ring. Thus there exists $0\neq e=e^2\in IQ$ such that $\sum_{i=1}^{n+1}a_iQ=eQ$, and $a_i=ea_i$ for each $i=1,\dots,n+1$. Notice that eQ satisfies the same generalized identities with skew derivations and automorphisms of I. So that we may assume $e\neq 1$, if not $eQ=Q$ and the conclusion follows from Proposition 2.6.

Assume that α is X-outer. Thus, by Fact 2.7 and (3.2), Q satisfies

(3.3)
$$\Big[\Big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\sum_{j=0}^{n-1}\alpha(e)t_{\sigma(1)}\cdots\alpha(e)t_{\sigma(j)}\alpha(e)y_{\sigma(j+1)}ex_{\sigma(j+2)}\cdots ex_{\sigma(n)}\Big)f(ex_1,\dots,ex_n),x_{n+1}\Big]$$

and in particular

$$\Big[\Big(\sum_{\sigma\in S_n}\alpha(\gamma_\sigma)\alpha(e)y_{\sigma(1)}\cdots\alpha(e)y_{\sigma(n)}\Big)f(ex_1,\dots,ex_n),x_{n+1}\Big]. \tag{3.4}$$

We also denote by $f^\alpha(x_1,\dots,x_n)$ the polynomial obtained from $f(x_1,\dots,x_n)$ by replacing each coefficient γ_σ with $\alpha(\gamma_\sigma)$. Therefore we may rewrite (3.4) as follows:

$$\Big[f^\alpha(\alpha(e)r_1,\dots,\alpha(e)r_n)f(es_1,\dots,es_n),X\Big]=0 \tag{3.5}$$

for all $r_1,\dots,r_n,s_1,\dots,s_n,X\in Q$. Choose in (3.5) $X=Y(1-\alpha(e))$, then we get

$$f^\alpha(\alpha(e)r_1,\dots,\alpha(e)r_n)f(es_1,\dots,es_n)Y(1-\alpha(e))=0$$

and by the primeness of Q and since $e \neq 1$, it follows that Q satisfies

$$f^{\alpha}(\alpha(e)y_1, \ldots, \alpha(e)y_n)f(ex_1, \ldots, ex_n)$$

that is $f^{\alpha}(\alpha(e)Q)f(eQ) = (0)$, where $\alpha(e)Q$ and eQ are both right ideals of Q and f^{α} and f are distinct polynomials over C (since $\alpha \neq 1$). In this situation, applying the result in [16] (see the proof of Lemma 3, pp. 181), it follows that either $f^{\alpha}(\alpha(e)Q)\alpha(e) = (0)$ or $f(eQ) = (0)$. Since this last case cannot occur, we have that $f^{\alpha}(\alpha(e)r_1, \ldots, \alpha(e)r_n)\alpha(e) = 0$ for all $r_1, \ldots, r_n \in Q$. Hence

$$0 = \alpha^{-1}\Big(f^{\alpha}(\alpha(e)r_1, \ldots, \alpha(e)r_n)\alpha(e)\Big) = f(e\alpha^{-1}(r_1), \ldots, e\alpha^{-1}(r_n))e$$

and since α^{-1} is an automorphism of Q, it follows that $f(es_1, \ldots, es_n)e = 0$, for all $s_1, \ldots, s_n \in Q$, which is again a contradiction.

Finally consider the case when there exists an invertible element $q \in Q$ such that $\alpha(x) = qxq^{-1}$, for all $x \in Q$. Thus from (3.2) we have that Q satisfies

(3.6)
$$\left[\Big(\sum_{\sigma\in S_n} \alpha(\gamma_\sigma) \sum_{j=0}^{n-1} q(ex_{\sigma(1)} \cdots ex_{\sigma(j)})eq^{-1}y_{\sigma(j+1)}ex_{\sigma(j+2)} \cdots ex_{\sigma(n)}\Big)f(ex_1, \ldots, ex_n), x_{n+1}\right].$$

Since $\alpha(\gamma_\sigma) = \gamma_\sigma$ and by replacing $y_{\sigma(i)}$ with $qx_{\sigma(i)}$, for all $\sigma \in S_n$ and for all $i = 1, \ldots, n$, it follows that Q satisfies

$$\left[\Big(\sum_{\sigma\in S_n} \gamma_\sigma qex_{\sigma(1)} \cdots ex_{\sigma(j)}ex_{\sigma(j+1)}ex_{\sigma(j+2)} \cdots ex_{\sigma(n)}\Big)f(ex_1, \ldots, ex_n), x_{n+1}\right] \tag{3.7}$$

that is

$$\left[\Big(qf(ex_1, \ldots, ex_n)\Big)f(ex_1, \ldots, ex_n), x_{n+1}\right]. \tag{3.8}$$

By Fact 3.1 it follows that one of the following holds:

1. $char(Q) = 2$ and Q satisfies s_4;
2. $f(x_1, \ldots, x_n)x_{n+1}$ is an identity for eQ;
3. $q \in C$;
4. $qeQ = (0)$.

Since in any case we get a contradiction, we are done. □

3.4. Lemma. *Let R be a prime ring, $f(x_1, \ldots, x_n)$ a non-central multilinear polynomial over C in n non-commuting indeterminates, I a nonzero right ideal of R, $b, c \in Q$ and $\alpha \in Aut(R)$ be an automorphism of R such that $F(x) = bx - \alpha(x)c$, for all $x \in R$. Assume that $F(f(r_1, \ldots, r_n))f(r_1, \ldots, r_n) \in C$, for all $r_1, \ldots, r_n \in I$. If R does not satisfy any non-trivial generalized polynomial identity then $F(I)I = (0)$.*

Proof. Let u be any nonzero element of I. By the hypothesis R satisfies the following:

$$\left[\Big(b(f(ux_1, \ldots, ux_n)) - \alpha(f(ux_1, \ldots, ux_n))c\Big)f(ux_1, \ldots, ux_n), x_{n+1}\right].$$

Also here we denote by $f^{\alpha}(x_1, \ldots, x_n)$ the polynomial obtained from $f(x_1, \ldots, x_n)$ by replacing each coefficient γ_σ of $f(x_1, \ldots, x_n)$ with $\alpha(\gamma_\sigma)$. Thus R satisfies

$$\left[\Big(bf(ux_1, \ldots, ux_n) - f^{\alpha}(\alpha(u)\alpha(x_1), \ldots, \alpha(u)\alpha(x_n))c\Big)f(ux_1, \ldots, ux_n), x_{n+1}\right]. \tag{3.9}$$

In case α is X-outer, by Theorem 3 in [14] and (3.9) we have that R satisfies

$$\left[\left(b(f(ux_1,\dots,ux_n)) - f^{\alpha}(\alpha(u)y_1,\dots,\alpha(u)y_n)c\right)f(ux_1,\dots,ux_n),x_{n+1}\right]$$

and in particular R satisfies both

$$\left[bf(ux_1,\dots,ux_n)^2,x_{n+1}\right] \tag{3.10}$$

and

$$\left[f^{\alpha}(\alpha(u)y_1,\dots,\alpha(u)y_n)cf(ux_1,\dots,ux_n),x_{n+1}\right]. \tag{3.11}$$

Since (3.10) and (3.11) must be trivial generalized polynomial identities for R, by [12] it follows that $bu = 0$ and $cu = 0$ that is $F(I)I = (0)$.

Consider now the case $\alpha(x) = qxq^{-1}$ for all $x \in R$, for some invertible element $q \in Q$. Since by (3.9)

$$\left[\left(bf(ux_1,\dots,ux_n) - qf(ux_1,\dots,ux_n)q^{-1}c\right)f(ux_1,\dots,ux_n),x_{n+1}\right] \tag{3.12}$$

is a trivial generalized polynomial identity for R, again by [12] we have that $bu = \lambda qu$, for some $\lambda \in C$. Thus we may write (3.12) as follows

$$\left[qf(ux_1,\dots,ux_n)(\lambda - q^{-1}c)f(ux_1,\dots,ux_n),x_{n+1}\right]. \tag{3.13}$$

Once again (3.13) is a trivial identity for R, moreover $qu \neq 0$. This implies that $(\lambda - q^{-1}c)u = 0$ and hence $(\lambda_u - q^{-1}c)u = 0$ for all $u \in I$ and for some $\lambda_u \in C$. Then u and $q^{-1}cu$ are C-dependent for all $u \in I$. By a standard argument we conclude that $(\lambda - q^{-1}c)I = (0)$ for some $\lambda \in C$, and thus $F(I)I = (0)$. □

3.5. Lemma. *Let R be a prime ring, $f(x_1,\dots,x_n)$ a non-central multilinear polynomial over C in n non-commuting indeterminates, I a nonzero right ideal of R, $b, c \in Q$ and $\alpha \in Aut(R)$ be an X-outer automorphism of R such that $F(x) = bx - \alpha(x)c$, for all $x \in R$. If $F(f(r_1,\dots,r_n))f(r_1,\dots,r_n) \in C$, for all $r_1,\dots,r_n \in I$, then either $char(R) = 2$ and R satisfies s_4 or one of the following holds:*

- *(i) $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I;*
- *(ii) $F(I)I = (0)$;*
- *(iii) $cI = (0)$, $b \in C$ and $f(x_1,\dots,x_n)^2$ is central valued on R.*

Proof. Firstly we notice that in case $cI = (0)$, then $bf(r_1,\dots,r_n)^2 \in C$, for all $r_1,\dots,r_n \in I$. Thus by Fact 3.1 it follows that either $cI = (0)$, $b \in C$ and $f(x_1,\dots,x_n)^2$ is central valued on R, or $cI = bI = (0)$ that is $F(I)I = (0)$. Hence in the following we assume $cI \neq (0)$. By previous Lemma we may assume that R satisfies some non-trivial generalized polynomial identity. As above let u be any nonzero element of I. By the hypothesis R satisfies the following:

$$\left[\left(bf(ux_1,\dots,ux_n) - f^{\alpha}(\alpha(u)\alpha(x_1),\dots,\alpha(u)\alpha(x_n))c\right)f(ux_1,\dots,ux_n),x_{n+1}\right]. \tag{3.14}$$

Since α is X-outer, by Theorem 3 in [14], R satisfies

$$\left[\left(bf(ux_1,\dots,ux_n) - f^{\alpha}(\alpha(u)y_1,\dots,\alpha(u)y_n)c\right)f(ux_1,\dots,ux_n),x_{n+1}\right] \tag{3.15}$$

and in particular R as well as Q satisfy the component

$$\left[f^{\alpha}(\alpha(u)y_1,\dots,\alpha(u)y_n)cf(ux_1,\dots,ux_n),x_{n+1}\right]. \tag{3.16}$$

By [31] Q is a primitive ring having nonzero socle H with the field C as its associated division ring. Moreover H and Q satisfy the same generalized polynomial identities with automorphisms (Theorem 1 in [14]). Therefore H satisfies (3.14) and so we may replace Q by H. Suppose there exist $a_1,\dots,a_{n+2}\in I$ such that $f(a_1,\dots,a_n)a_{n+1}\neq 0$ and $ca_{n+2}\neq 0$. Since Q is a regular GPI-ring, there exists an idempotent element $e\in IQ$ such that $eQ=\sum_{i=1}^{n+2}a_iQ$ and $a_i=ea_i$, for any $i=1,\dots,n+2$. Therefore, by (3.14), Q satisfies

$$\left[\left(bf(ex_1,\dots,ex_n)-f^{\alpha}(\alpha(e)\alpha(x_1),\dots,\alpha(e)\alpha(x_n))c\right)f(ex_1,\dots,ex_n),x_{n+1}\right]. \tag{3.17}$$

Moreover assume $e\neq 1$, if not $eQ=Q$ and by Proposition 2.6 we get $b\in C$, $c=0$ and $f(x_1,\dots,x_n)^2$ is central valued on R. Since α is X-outer, as above by (3.17) Q satisfies

$$\left[\left(bf(ex_1,\dots,ex_n)-f^{\alpha}(\alpha(e)y_1,\dots,\alpha(e)y_n)c\right)f(ex_1,\dots,ex_n),x_{n+1}\right].$$

In particular Q satisfies

$$\left[f^{\alpha}(\alpha(e)y_1,\dots,\alpha(e)y_n)cf(ex_1,\dots,ex_n),x_{n+1}(1-\alpha(e))\right]$$

that is Q satisfies

$$f^{\alpha}(\alpha(e)y_1,\dots,\alpha(e)y_n)cf(ex_1,\dots,ex_n)x_{n+1}(1-\alpha(e))$$

and since Q is prime and $e\neq 0,1$, it follows $f^{\alpha}(\alpha(e)r_1,\dots,\alpha(e)r_n)cf(es_1,\dots,es_n)=0$, for all $r_1,\dots,r_n,s_1,\dots,s_n\in Q$. Since $f(ea_1,\dots,ea_n)ea_{n+1}\neq 0$ and $cea_{n+2}\neq 0$ and by using the result in [16], it follows that $f^{\alpha}(\alpha(e)y_1,\dots,\alpha(e)y_n)$ is an identity for Q. This implies that $f(e\alpha^{-1}(y_1),\dots,e\alpha^{-1}(y_n))$ is also an identity for Q. Moreover it is clear that α^{-1} is X-outer, therefore $f(ex_1,\dots,ex_n)$ is an identity for Q, a contradiction. $\square$

3.6. Lemma. *Let R be a prime ring, $f(x_1,\dots,x_n)$ a non-central multilinear polynomial over C in n non-commuting indeterminates, I a nonzero right ideal of R, $b,c,q\in Q$ such that $F(x)=bx-qxq^{-1}c$, for all $x\in R$. If*

$$F(f(r_1,\dots,r_n))f(r_1,\dots,r_n)=0,$$

for all $r_1,\dots,r_n\in I$, then either $char R=2$ and R satisfies s_4 or one of the following holds:

- (i) *$f(x_1,\dots,x_n)x_{n+1}$ is an identity for I;*
- (ii) *$[f(x_1,\dots,x_n),x_{n+1}]x_{n+2}$ is an identity for I, $(b-c)I=(0)$ and $q^{-1}cI\subseteq I$;*
- (iii) *$F(I)I=(0)$.*

Proof. Here I satisfies

$$\left(bf(x_1,\dots,x_n)-qf(x_1,\dots,x_n)q^{-1}c\right)f(x_1,\dots,x_n) \tag{3.18}$$

and left multiplying by q^{-1}, I satisfies

$$\left(q^{-1}b(f(x_1,\dots,x_n))-(f(x_1,\dots,x_n)q^{-1}c\right)f(x_1,\dots,x_n). \tag{3.19}$$

Since we assume $f(x_1,\dots,x_n)$ is not central valued on R, by Fact 3.2 we have that either $char R=2$ and R satisfies the standard identity s_4, or $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I, or one of the following holds:

1. there exists $\gamma\in C$ such that $q^{-1}bx=\gamma x=q^{-1}cx$, for all $x\in I$ (this is the case $F(I)I=(0)$).
2. $q^{-1}(b-c)I=(0)$, that is $(b-c)I=(0)$, moreover $[f(x_1,\dots,x_n),x_{n+1}]x_{n+2}$ is an identity for I.

In this last case, by (3.19) it follows that I satisfies

$$\left(bf(ux_1,\dots,ux_n)-qf(ux_1,\dots,ux_n)q^{-1}b\right)f(ux_1,\dots,ux_n) \tag{3.20}$$

and moreover, since I satisfies the polynomial identity $[f(x_1,\dots,x_n),x_{n+1}]x_{n+2}$, in view of Proposition in [25], $I=eQ$ for some idempotent e in the socle of Q. Here we write $f(x_1,\dots,x_n)=\sum t_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)x_i$, where any t_i is a multilinear polynomial in $n-1$ variables and x_i never appears in t_i. Of course, if $t_i(ex_1,\dots,ex_{i-1},ex_{i+1},\dots,ex_n)e$ is an identity for Q, then $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I and we are done. Thus assume there exists $i\in\{1,\dots,n\}$ such that $t_i(er_1,\dots,er_{i-1},er_{i+1},\dots,er_n)e\neq 0$ for some $r_1,\dots,r_n\in I$. In particular,

$$f(ex_1,\dots,ex_{i-1},ex_i(1-e),ex_{i+1},\dots,ex_n)=t_i(ex_1,\dots,ex_n)ex_i(1-e)$$

and by (3.20) Q satisfies

$$\begin{aligned}&bt_i(ex_1,\dots,ex_n)ex_i(1-e)t_i(ex_1,\dots,ex_n)ex_i(1-e)\\&\qquad-qt_i(ex_1,\dots,ex_n)ex_i(1-e)q^{-1}bt_i(ex_1,\dots,ex_n)ex_i(1-e)\end{aligned}$$

that is Q satisfies

$$\left(-qt_i(ex_1,\dots,ex_n)ex_i(1-e)q^{-1}b\right)t_i(ex_1,\dots,ex_n)ex_i(1-e) \tag{3.21}$$

and left multiplying by $(1-e)q^{-1}bq^{-1}$, we easily have that Q satisfies

$$(1-e)q^{-1}bt_i(ex_1,\dots,ex_n)eX(1-e)q^{-1}bt_i(ex_1,\dots,ex_n)eX(1-e). \tag{3.22}$$

By Lemma 2 in [32] and since $e\neq 1$, it follows that

$$(1-e)q^{-1}bt_i(ex_1,\dots,ex_{i-1},ex_{i+1},\dots,ex_n)e$$

is an identity for Q, that is $(1-e)q^{-1}bet_i(x_1e,\dots,x_{i-1}e,x_{i+1}e,\dots,x_ne)$ is an identity for Q. In this case, since $t_i(x_1e,\dots,x_{i-1}e,x_{i+1}e,\dots,x_ne)$ is not an identity for Q, we get in view of the result in [16], $(1-e)q^{-1}be=0$, that is $q^{-1}bI\subseteq I$ and also $q^{-1}cI\subseteq I$. $\square$

3.7. Theorem. *Let R be a prime ring, $f(x_1,\dots,x_n)$ a multilinear polynomial over C in n non-commuting variables, I a non-zero right ideal of R, $F:R\to R$ be a non-zero generalized skew derivation of R. Suppose that*

$$F(f(r_1,\dots,r_n))f(r_1,\dots,r_n)\in C,$$

for all $r_1,\dots,r_n\in I$. If $f(x_1,\dots,x_n)$ is not central valued on R, then either $char(R)=2$ and R satisfies s_4 or one of the following holds:

- *(i) $f(x_1,\dots,x_n)x_{n+1}$ is an identity for I;*
- *(ii) $F(I)I=(0)$;*
- *(iii) $[f(x_1,\dots,x_n),x_{n+1}]x_{n+2}$ is an identity for I, there exist $b,c,q\in Q$ with q invertible such that $F(x)=bx-qxq^{-1}c$ for all $x\in R$, and $q^{-1}cI\subseteq I$; moreover in this case either $(b-c)I=(0)$ or $b-c\in C$ and $f(x_1,\dots,x_n)^2$ is central valued on R provided that $b\neq c$.*

Proof. In view of all previous Lemmas and Propositions, we may assume $I\neq R$ and $F(x)=bx-qxq^{-1}c$, for all $x\in R$. Moreover we may assume that there exist $s_1,\dots,s_n\in I$ such that $F(f(s_1,\dots,s_n))f(s_1,\dots,s_n)\neq 0$. Therefore

$$(bf(x_1,\dots,x_n)-qf(x_1,\dots,x_n)q^{-1}c)f(x_1,\dots,x_n)$$

is a central generalized polynomial identity for I. Thus R is a PI-ring and so RC is a finite dimensional central simple C-algebra (the proof of this fact is the same of Theorem

1 in [7]). By Wedderburn-Artin theorem, $RC \cong M_k(D)$ for some $k \geq 1$ and D a finite-dimensional central division C-algebra. By Theorem 2 in [24]

$$(bf(x_1,\ldots,x_n) - qf(x_1,\ldots,x_n)q^{-1}c)f(x_1,\ldots,x_n) \in C$$

for all $x_1,\ldots,x_n \in IC$. Without loss of generality we may replace R with RC and assume that $R = M_k(D)$. Let E be a maximal subfield of D, so that $M_k(D) \otimes_C E \cong M_t(E)$ where $t = k \cdot [E : C]$. Hence $(bf(r_1,\ldots,r_n) - qf(r_1,\ldots,r_n)q^{-1}c)f(r_1,\ldots,r_n) \in C$, for any $r_1,\ldots,r_n \in I \otimes E$ (Lemma 2 in [24] and Proposition in [29]). Therefore we may assume that $R \cong M_t(E)$ and $I = eR = (e_{11}R + \cdots + e_{ll}R)$, where $t \geq 2$ and $l \leq t$.

Suppose that $t \geq 2$, otherwise we are done and denote $q = \sum_{r,s} q_{rs}e_{rs}$ and $q^{-1}c = \sum_{r,s} c_{rs}e_{rs}$, for $q_{rs}, c_{rs} \in E$. As in Lemma 3.6 we write

$$f(x_1,\ldots,x_n) = \sum t_i(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)x_i$$

and there exists some $t_i(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)x_i$ which is not an identity for I. In particular $qt_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)ex_i$ is not an identity for R, because q is invertible. Hence, again for

$$f(ex_1,,\ldots,ex_{i-1},ex_i(1-e),ex_{i+1},\ldots,ex_n) = t_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)ex_i(1-e)$$

and by our hypothesis, we have that

$$qt_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)ex_i(1-e)q^{-1}ct_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)ex_i(1-e)$$

is an identity for R, and by the primeness of R it follows that

$$(1-e)q^{-1}ct_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)e$$

is an identity for R. By [16] and since $t_i(ex_1,\ldots,ex_{i-1},ex_{i+1},\ldots,ex_n)ex_i$ is not an identity for R, the previous identity says that $(1-e)q^{-1}ce = 0$. Thus $q^{-1}cI \subseteq I$.
In case $[f(x_1,\ldots,x_n),x_{n+1}]x_{n+2}$ is an identity for I, then by our assumption we get $(b-c)f(r_1,\ldots,r_n)^2 \in C$ for all $r_1,\ldots,r_n \in I$. In view of Fact 3.1, either $(b-c)I = (0)$ and we are done, or $b-c \in C$ and $f(x_1,\ldots,x_n)^2$ is central valued on R, provided that $b \neq c$.

Consider finally the case $[f(x_1,\ldots,x_n),x_{n+1}]x_{n+2}$ is not an identity for I. By Lemma 3 in [6], for any $i \leq l$, $j \neq i$, the element e_{ij} falls in the additive subgroup of RC generated by all valuations of $f(x_1,\ldots,x_n)$ in I. Since the matrix $(be_{ij} - qe_{ij}q^{-1}c)e_{ij}$ has rank at most 1, then it is not central. Therefore $qe_{ij}q^{-1}ce_{ij} = 0$, i.e. $q_{ki}(q^{-1}c)_{ji} = 0$ for all k and for all $j \neq i$. Since q is invertible, there exists some $q_{ki} \neq 0$, therefore $(q^{-1}c)_{ji} = 0$ for all $j \neq i$.

Consider the following automorphism of R:

$$\lambda(x) = (1+e_{ij})x(1-e_{ij}) = x + e_{ij}x - xe_{ij} - e_{ij}xe_{ij}$$

for any $i,j \leq l$, and note that $\lambda(I) \subseteq I$ is a right ideal of R satisfying

$$\left[\left(\lambda(b)f(x_1,\ldots,x_n) - \lambda(q)f(x_1,\ldots,x_n)\lambda(q^{-1}c)\right)f(x_1,\ldots,x_n), x_{n+1}\right].$$

If we denote $\lambda(q^{-1}c) = \sum_{rs} c'_{rs}e_{rs}$, the above argument says that $c'_{rs} = 0$ for all $s \leq l$ and $r \neq s$. In particular the (i,j)-entry of $\lambda(q^{-1}c)$ is zero. This implies that $c_{ii} = c_{jj} = \alpha$, for all $i,j \leq l$. Therefore $q^{-1}cx = \alpha x$ for all $x \in I$. This leads to $(b-c)f(r_1,\ldots,r_n)^2 \in C$ for all $r_1,\ldots,r_n \in I$ and we conclude by the same argument above. □

For the sake of completeness, we would like to conclude this paper by showing the explicit meaning of the conclusion $F(I)I = (0)$, more precisely we state the following:

3.8. Remark. Let R be a prime ring, I be a non-zero right ideal of R and $F : R \to R$ be a non-zero generalized skew derivation of R. If $F(I)I = (0)$ then there exist $a, b \in Q$ and $\alpha \in Aut(R)$ such that $F(x) = (a + b)x - \alpha(x)b$ for all $x \in R$, $aI = (0)$ and one of the following holds:

- (i) $bI = (0)$;
- (ii) there exist $\lambda \in C$ and an invertible element $q \in Q$ such that $\alpha(x) = qxq^{-1}$, for all $x \in R$, and $q^{-1}by = \lambda y$, for all $y \in I$.

Proof. As previously remarked we can write $F(x) = ax + d(x)$ for all $x \in R$, where $a \in Q$ and d is a skew derivation of R (see [8]). Let $\alpha \in Aut(R)$ be the automorphism associated with d, in the sense that $d(xy) = d(x)y + \alpha(x)d(y)$, for all $x, y \in R$. Thus, by the hypothesis, for all $x, y \in I$,

$$(3.23) \quad (ax + d(x))y = 0.$$

For all $x, y, z \in I$ we have:

$$0 = F(xz)y = (ax + d(x))zy + \alpha(x)d(z)y$$

and by (3.23) we obtain $\alpha(x)d(z)y = 0$ for all $x, y, z \in I$. Moreover $\alpha(I)$ is a non-zero right ideal of R, so that it follows

$$(3.24) \quad d(z)y = 0$$

for all $y, z \in I$. Once again by (3.23) we get $azy = 0$ for all $z, y \in I$, that is $aI = (0)$.

Finally in (3.24) replace z with xs, for any $x \in I$ and $s \in R$, then:

$$(3.25) \quad 0 = d(xs)y = d(x)sy + \alpha(x)d(s)y$$

for all $x, y \in I$, $s \in R$. In case d is X-outer, it follows that $d(x)sy + \alpha(x)ty = 0$, for all $x, y \in I$ and $s, t \in R$ (Theorem 1 in [15]). In particular $\alpha(x)ty = 0$, which implies the contradiction $\alpha(x) = 0$ for all $x \in I$. Therefore we may assume that d is X-inner, that is there exists $b \in Q$ such that $d(r) = br - \alpha(r)b$, for all $r \in R$ and by (3.24)

$$(3.26) \quad (bx - \alpha(x)b)y = 0$$

for all $x, y \in I$. Consider first the case α is X-outer and replace x with xr, for any $r \in R$. Then $(bxr - \alpha(x)\alpha(r)b)y = 0$ and, by Theorem 3 in [14], $(bxr - \alpha(x)sb)y = 0$ for all $x, y \in I$ and $r, s \in R$. In particular $bIRI = (0)$, which implies $bI = (0)$ and we are done.

On the other hand, if there exists an invertible element $q \in Q$ such that $\alpha(r) = qrq^{-1}$, for all $r \in R$, from (3.26) we have $(bx - qxq^{-1}b)y = 0$, for all $x, y \in I$. Left multiplying by q^{-1}, it follows $[q^{-1}b, x]y = 0$, and by Lemma in [4] there exists $\lambda \in C$ such that $q^{-1}bx = \lambda x$ for all $x \in I$. $\square$

References

[1] Ç. Demir, N. Argaç, *Prime rings with generalized derivations on right ideals*, Algebra Colloq. **18** (Spec. 1), 987-998, 2011.

[2] K.I. Beidar, W.S. Martindale, III, A.V. Mikhalev, *Rings with generalized identities*(Monographs and Textbooks in Pure and Applied Mathematics, 196, 1996).

[3] M. Bresar, *Centralizing mappings and derivations in prime rings*, J. Algebra **156** (2), 385-394, 1993.

[4] M. Bresar, *One-sided ideals and derivations of prime rings*, Proc. Amer. Math. Soc. **122** (4), 979-983, 1994.

[5] L. Carini, V. De Filippis, *Identities with generalized derivations on prime rings and Banach algebras*, Algebra Colloq., **19** (Spec. 1), 971-986, 2012.

[6] C.M. Chang, *Power central values of derivations on multilinear polynomials*, Taiwanese J. Math. **7** (2), 329-338, 2003.

[7] C.M. Chang, T.K. Lee, *Annihilators of power values of derivations in prime rings*, Comm. in Algebra **26** (7), 2091-2113, 1998.

[8] J.C. Chang, *On the identitity $h(x) = af(x)+g(x)b$*, Taiwanese J. Math. **7** (1), 103-113, 2003.

[9] J. C. Chang, *Generalized skew derivations with annihilating Engel conditions*, Taiwanese J. Math. **12** (7), 1641-1650, 2008.

[10] J. C. Chang, *Generalized skew derivations with nilpotent values on Lie ideals*, Monatsh. Math. **161** (2), 155-160, 2010.

[11] H. W. Cheng and F. Wei, *Generalized skew derivations of rings*, Adv. Math. (China), **35** (2), 237-243, 2006.

[12] C.L. Chuang, *GPIs having coefficients in Utumi quotient rings*, Proc. Amer. Math. Soc. **103** (3), 723-728, 1988.

[13] C.L. Chuang, *Differential identities with automorphisms and antiautomorphisms I*, J. Algebra **149** (2), 371-404, 1992.

[14] C.L. Chuang, *Differential identities with automorphisms and antiautomorphisms II*, J. Algebra **160** (1), 130-171, 1993.

[15] C.L. Chuang, T.K. Lee, *Identities with a single skew derivation*, J. Algebra **288** (1), 59-77, 2005.

[16] C.L. Chuang, T.K. Lee, *Rings with annihilator conditions on multilinear polynomials*, Chinese J. Math. **24** (2), 177-185, 1996.

[17] M.C. Chou, C.K. Liu, *An Engel condition with skew derivations*, Monatsh. Math. **158** (3), 259-270, 2009.

[18] V. De Filippis, F. Wei, *Posner's second theorem for skew derivations on left ideals*, Houston J. Math. **38** (2), 373-395, 2012.

[19] C. Faith, Y. Utumi, *On a new proof of Litoff's Theorem*, Acta Math. Acad. Sci. Hungar **14**, 369-371, 1963.

[20] N. Jacobson, *Structure of rings* (American Mathematical Society Colloquium Publications, Vol. 37. Revised edition American Mathematical Society, Providence, R.I. 1964).

[21] V.K. Kharchenko, *Generalized identities with automorphisms*, Algebra and Logic **14** (2), 132-148, 1975.

[22] C. Lanski, *An Engel condition with derivation*, Proc. Amer. Math. Soc. **118** (3), 731-734, 1993.

[23] T.K. Lee, *Derivations with invertible values on a multilinear polynomial*, Proc. Amer. Math. Soc. **119** (4), 1077-1083, 1993.

[24] T.K. Lee, *Left annihilators characterized by GPIs*, Trans. Amer. Math. Soc. **347** (8), 3159-3165, 1995.

[25] T.K. Lee, *Power reduction property for generalized identities of one-sided ideals*, Algebra Colloq. **3** (1), 19-24, 1996.

[26] T.K. Lee, *Generalized skew derivations characterized by acting on zero products*, Pacific J. Math. **216** (2), 293-301, 2004.

[27] T.K. Lee, W.K. Shiue *Derivations cocentralizing polynomials*, Taiwanese J. Math. **2** (4), 457-467, 1998.

[28] T.K. Lee, Y. Zhou, *An identity with generalized derivations*, J. Algebra and Appl. **8** (3), 307-317, 2009.

[29] P.H. Lee, T.L. Wong, *Derivations cocentralizing Lie ideals*, Bull. Inst. Math. Acad. Sinica **23** (1), 1-5, 1995.

[30] U. Leron, *Nil and power central polynomials in rings*, Trans. Amer. Math. Soc. **202**, 97-103, 1975.

[31] W.S. Martindale III, *Prime rings satisfying a generalized polynomial identity*, J. Algebra **12**, 576-584, 1969.

[32] E.C. Posner, *Derivations in prime rings*, Proc. Amer. Math. Soc., **8**, 1093-1100, 1957.

[33] T.L. Wong, *Derivations with power central values on multilinear polynomials*, Algebra Colloq. **3** (4), 369-378, 1996.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 85 – 95

GENERALIZED NOTION OF WEAK MODULE AMENABILITY

Abasalt Bodaghi*

Received 16 : 06 : 2011 : Accepted 28 : 02 : 2013

Abstract

In the present paper, we introduce a new notion of weak module amenability for Banach algebras which is related to module homomorphisms. Among other results, we investigate the relationship between this concept for a Banach algebra $\mathcal{A}$ which is a Banach $\mathfrak{A}$-bimodule with compatible actions, and the quotient Banach algebra $\mathcal{A}/J$ where J is the closed ideal of $\mathcal{A}$ generated by elements of the form $(a\cdot\alpha)b-a(\alpha\cdot b)$ for $a\in\mathcal{A}$ and $\alpha\in\mathfrak{A}$. We then study this concept for an inverse semigroup S, where some examples on $\ell^1(S)$ and $C^*(S)$ are given.

Keywords: Banach modules; Module derivation; Weak amenability; Weak module amenability; Inverse semigroup.

2000 AMS Classification: 46H25.

1. Introduction

Let S be a (discrete) semigroup. The semigroup algebra $\ell^1(S)$ is the Banach algebra consisting of all absolutely summable complex-valued functions on S, with the convolution product and the ℓ^1-norm; $\|f\|_1=\sum_{s\in S}|f(s)|$ $(f\in\ell^1(S))$. We will use δ_s to denote the point mass function at s; $\delta_s(t)=1$ if $t=s$ and $=0$ elsewhere. Using point masses we may represent a function f on S as $f=\sum_{s\in S}f(s)\delta_s$. Here we recall that an *inverse semigroup* is a discrete semigroup S such that for each $s\in S$, there is a unique element $s^*\in S$ with $ss^*s=s$ and $s^*ss^*=s^*$. The set of elements of the form s^*s are called *idempotents* of S and denoted by E.

The concept of amenability for a Banach algebra $\mathcal{A}$ was introduced by B. E. Johnson in [18]. A Banach algebra $\mathcal{A}$ is *amenable* if every bounded derivation from $\mathcal{A}$ into any dual Banach $\mathcal{A}$-module is inner, equivalently if $H^1(\mathcal{A},X^*)=\{0\}$ for every Banach $\mathcal{A}$-module X, where $H^1(\mathcal{A},X^*)$ is the *first Hochschild cohomology group* of $\mathcal{A}$ with coefficients in X^*, the first dual space of X. Also, a Banach algebra $\mathcal{A}$ is *weakly amenable* if $H^1(\mathcal{A},\mathcal{A}^*)=\{0\}$. Bade, Curtis and Dales introduced the notion of weak amenability in [5]. They considered this concept only for commutative Banach algebras. After that

* Department of Mathematics, Garmsar Branch, Islamic Azad University, Garmsar, Iran, Email: abasalt.bodaghi@gmail.com

Johnson defined the weak amenability for arbitrary Banach algebras [19] and showed that for a locally compact group G, $L^1(G)$ is weakly amenable [20]. This fact fails for semigroups though. For example, if S is the bicyclic inverse semigroup, then $\ell^1(S)$ is not weakly amenable [9].

Homomorphisms on Banach algebras play an important role in Functional Analysis. Papers [8] and [21] defined and investigated two concepts of the amenability for Banach algebras by using homomorphisms which are different from weak amenability and amenability. In [1], Amini introduced the concept of module amenability of a Banach algebra $\mathcal{A}$ which is a Banach module over another Banach algebra $\mathfrak{A}$ with compatible actions. Later this notion of amenability is generalized by the author in [7]. The notion of weak module amenability of Banach algebras is defined in [4] and studied in [2]. In fact, the author and Amini investigated the concept of weak module amenability in [2] and obtained some results on the seond dual of a Banach algebra. In [6], the author showed that for an arbitrary inverse semigroup S with a set of idempotents E, the semigroup algebra $\ell^1(S)$ as an $\ell^1(E)$-module with trivial left action is always weakly module amenable. The abelian case for S was proved earlier in [4]. These papers motivated us to generalize of the concept of weak module amenability by homomorphisms.

Let $\mathcal{A}$ and $\mathfrak{A}$ be Banach algebras such that $\mathcal{A}$ is a Banach $\mathfrak{A}$-bimodule with compatible actions. Then every $\mathfrak{A}$-module homomorphism σ (not necessarily $\mathbb{C}$-linear) on $\mathcal{A}$ induces a linear continuous homomorphism $\widehat{\sigma}$ on $\mathcal{A}/J$, where J is a closed ideal of $\mathcal{A}$. In section three, we generalize the concept of weak module amenability of Banach algebras by using $\mathfrak{A}$-module homomorphisms. On the other hand, for each pair $\mathfrak{A}$-module homomorphism σ and τ on $\mathcal{A}$, we define (σ,τ)-weak module amenability of Banach algebras and among other results, we study the relation between (σ,τ)-weak module amenability of $\mathcal{A}$ and $(\widehat{\sigma},\widehat{\tau})$-weak amenability of $\mathcal{A}/J$, where J is the closed ideal of $\mathcal{A}$ generated by elements of the form $(a\cdot\alpha)b - a(\alpha\cdot b)$, for $a\in\mathcal{A}$ and $\alpha\in\mathfrak{A}$ (see also [8]).

In the last part of this paper, we show that under some conditions, $\ell^1(S)$ is (σ,τ)-weakly module amenable for all $\ell^1(E)$-module homomorphisms σ and τ on $\ell^1(S)$. Finally by applying our results, we give an example that $\ell^1(S)$ $[C^*(S)]$ is (σ,σ)-weakly module amenable as an $\ell^1(E)$-bimodule [as an $C^*(E)$-bimodule]. These examples show that this new concept and module amenability on Banach algebras do not coincide.

2. Preliminaries and Notations

Throughout this paper, $\mathcal{A}$ and $\mathfrak{A}$ are Banach algebras such that $\mathcal{A}$ is a Banach $\mathfrak{A}$-bimodule with compatible actions as follows:

$$\alpha\cdot(ab) = (\alpha\cdot a)b, \quad (ab)\cdot\alpha = a(b\cdot\alpha) \quad (a,b\in\mathcal{A},\alpha\in\mathfrak{A}).$$

Let X be a Banach $\mathcal{A}$-bimodule and a Banach $\mathfrak{A}$-bimodule with the following compatible actions:

$$\alpha\cdot(a\cdot x) = (\alpha\cdot a)\cdot x,\ a\cdot(\alpha\cdot x) = (a\cdot\alpha)\cdot x,\ (\alpha\cdot x)\cdot a = \alpha\cdot(x\cdot a) \quad (a\in\mathcal{A},\alpha\in\mathfrak{A},x\in X)$$

and similar for the right or two-sided actions. Then we say that X is a Banach $\mathcal{A}$-$\mathfrak{A}$-module. Moreover, if $\alpha\cdot x = x\cdot\alpha$ for all $\alpha\in\mathfrak{A}, x\in X$, then X is called a *commutative* $\mathcal{A}$-$\mathfrak{A}$-module. If X is a commutative Banach $\mathcal{A}$-$\mathfrak{A}$-module, then so is X^*, where the actions of $\mathcal{A}$ and $\mathfrak{A}$ on X^* are defined as follows:

$$\langle f\cdot\alpha, x\rangle = \langle f, \alpha\cdot x\rangle,\ \langle f\cdot a, x\rangle = \langle f, a\cdot x\rangle,$$

$$\langle\alpha\cdot f, x\rangle = \langle f, x\cdot\alpha\rangle,\ \langle a\cdot f, x\rangle = \langle f, x\cdot a\rangle \quad (a\in\mathcal{A},\alpha\in\mathfrak{A},x\in X,f\in X^*).$$

One should remember that $\mathcal{A}$ is not an $\mathcal{A}$-$\mathfrak{A}$-module in general because $\mathcal{A}$ does not satisfy the compatibility condition $a\cdot(\alpha\cdot b) = (a\cdot\alpha)\cdot b$ for $\alpha\in\mathfrak{A}, a,b\in\mathcal{A}$. But when $\mathcal{A}$ is a commutative $\mathfrak{A}$-module and acts on itself by multiplication from both sides, then it is also a Banach $\mathcal{A}$-$\mathfrak{A}$-module.

Let E and F be Banach algebras. We denote by $\mathrm{Hom}(E,F)$ the metric space of all bounded homomorphisms from E into F, with the metric derived from the bounded linear operators from E into F, and denote $\mathrm{Hom}(E,E)$ by $\mathrm{Hom}(E)$.

Now let $\mathcal{A}$ and $\mathcal{B}$ be $\mathfrak{A}$-bimodules. Then a $\mathfrak{A}$-module homomorphism from $\mathcal{A}$ to $\mathcal{B}$ is a bounded map $T:\mathcal{A}\longrightarrow\mathcal{B}$ with $T(a\pm b)=T(a)\pm T(b)$, and is multiplicative, that is $T(ab)=T(a)T(b)$ for all $a,b\in\mathcal{A}$, and

$$T(\alpha\cdot a)=\alpha\cdot T(a),\ T(a\cdot\alpha)=T(a)\cdot\alpha,\quad (a,\in\mathcal{A},\alpha\in\mathfrak{A}).$$

We denote by $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A},\mathcal{B})$, the space of all such homomorphisms and denote $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A},\mathcal{A})$ by $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. Note that when $\mathfrak{A}=\mathbb{C}$, the set of complex numbers, then $\mathrm{Hom}_{\mathbb{C}}(\mathcal{A},\mathcal{B})=\mathrm{Hom}(\mathcal{A},\mathcal{B})$. Although the elements of $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A},\mathcal{B})$ are not necessarily linear, their boundedness still implies their norm continuity.

Let $\mathcal{A}$ and $\mathfrak{A}$ be as above and X be a Banach $\mathcal{A}$-$\mathfrak{A}$-module. Recall that the mapping $D:\mathcal{A}\longrightarrow X$ is bounded if there exists $M>0$ such that $\|D(a)\|\le M\|a\|$ for all $a\in\mathcal{A}$. Suppose that φ and ψ are in $\mathrm{Hom}_{\mathfrak{A}}(A)$. A bounded map $D:\mathcal{A}\longrightarrow X$ is called a *module (φ,ψ)-derivation* if

$$D(\alpha\cdot a)=\alpha\cdot D(a),\quad D(a\cdot\alpha)=D(a)\cdot\alpha\quad (a\in\mathcal{A},\alpha\in\mathfrak{A})$$

and

$$D(a\pm b)=D(a)\pm D(b),\ \ D(ab)=D(a)\cdot\varphi(b)+\psi(a)\cdot D(b)\quad (a,b\in\mathcal{A}).$$

If X is a commutative $\mathcal{A}$-$\mathfrak{A}$-module, then each $x\in X$ defines a module (φ,ψ)-derivation $D_x(a)=x\cdot\varphi(a)-\psi(a)\cdot x$ on $\mathcal{A}$. These are called *module (φ,ψ)-inner derivations*. Derivations of these forms are studied in [7]. A Banach algebra $\mathcal{A}$ is called *module (φ,ψ)-amenable* (as an $\mathfrak{A}$-module) if for any commutative Banach $\mathcal{A}$-$\mathfrak{A}$-module X, each module (φ,ψ)-derivation $D:\mathcal{A}\longrightarrow X^*$ is (φ,ψ)-inner [7]. We use the notations $Z_{\mathfrak{A}}(\mathcal{A},(X_{(\varphi,\psi)})^*)$ for the space of all module (φ,ψ)-derivations $D:\mathcal{A}\longrightarrow X^*$, $B_{\mathfrak{A}}(\mathcal{A},(X_{(\varphi,\psi)})^*)$ for those which are inner (φ,ψ)-derivations, and $H_{\mathfrak{A}}(\mathcal{A},(X_{(\varphi,\psi)})^*)$ for the quotient space which we call the first relative (to $\mathfrak{A}$) (φ,ψ)-cohomology group of $\mathcal{A}$ with coefficients in X^*. Hence $\mathcal{A}$ is module (φ,ψ)-amenable if and only if $H_{\mathfrak{A}}(\mathcal{A},(X_{(\varphi,\psi)})^*)=\{0\}$ for all commutative Banach $\mathcal{A}$-$\mathfrak{A}$-module X. Indeed, for any $\phi,\psi\in\mathrm{Hom}(\mathcal{A})$, a Banach algebra $\mathcal{A}$ is (ϕ,ψ)-weakly amenable if $H^1(\mathcal{A},(\mathcal{A}_{(\phi,\psi)})^*)=\{0\}$ (for details see [8]).

3. (σ,τ)-weak module amenability of Banach algebras

Let Y be a subspace $\mathcal{A}^*$ as a vector space which is $\mathcal{A}$-submodule and commutative Banach $\mathfrak{A}$-submodule. From now on, such subspaces are called commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$.

3.1. Definition. Let $\mathcal{A}$ be a Banach $\mathfrak{A}$-module and $\sigma,\tau\in\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. Then $\mathcal{A}$ is called *(σ,τ)-weakly module amenable* (as an $\mathfrak{A}$-module) if for any commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule Y of $\mathcal{A}^*$, each module derivation from $\mathcal{A}$ to $Y_{(\sigma,\tau)}$ is inner.

In other words, in the above definition the module actions on $\mathcal{A}$ are considered as follows:

$$a\cdot x:=\sigma(a)x,\quad x\cdot a=x\tau(a)\quad (a,x\in\mathcal{A}).$$

Thus, the module actions $\mathcal{A}$ on $Y\subseteq\mathcal{A}^*$ are as follows:

$$\langle a\cdot y,b\rangle=\langle y,b\tau(a)\rangle,\quad \langle y\cdot a,b\rangle=\langle y,\sigma(a)b\rangle\quad (a,b\in\mathcal{A},y\in Y).$$

Note that if σ and τ are the identity maps, then (σ,τ)-weak module amenability becomes weak module amenability (see [2]).

Consider the closed ideal J of $\mathcal{A}$ generated by elements of the form $(a \cdot \alpha)b - a(\alpha \cdot b)$ for $\alpha \in \mathfrak{A}, a, b \in \mathcal{A}$. The ideal J is both $\mathcal{A}$-submodule and $\mathfrak{A}$-submodules of $\mathcal{A}$. Hence the quotient Banach algebra $\mathcal{A}/J$ is a Banach $\mathcal{A}$-$\mathfrak{A}$-module with compatible actions when $\mathcal{A}$ acts on $\mathcal{A}/J$ canonically. Now, if $\mathcal{A}/J$ is a commutative Banach $\mathfrak{A}$-module and σ, τ are epimorphisms in $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$, then $\mathcal{A}$ is (σ, τ)-weakly module amenable if and only if every module derivation from $\mathcal{A}$ to $(\mathcal{A}/J)^*$ is inner. In fact for each $\alpha \in \mathfrak{A}, a, b \in \mathcal{A}, y \in Y$, we have

$$\begin{aligned}
\langle y, (\sigma(a) \cdot \alpha)\tau(b) - \sigma(a)(\alpha \cdot \tau(b))\rangle &= \langle y, (\sigma(a) \cdot \alpha)\tau(b)\rangle - \langle y, \sigma(a)(\alpha \cdot \tau(b))\rangle \\
&= \langle b \cdot y, \sigma(a) \cdot \alpha\rangle - \langle y \cdot a, \alpha \cdot \tau(b)\rangle \\
&= \langle \alpha \cdot (b \cdot y), \sigma(a)\rangle - \langle (y \cdot a).\alpha, \tau(b)\rangle \\
&= \langle (b \cdot y) \cdot \alpha, \sigma(a)\rangle - \langle \alpha \cdot (y \cdot a), \tau(b)\rangle \\
&= \langle b \cdot (y \cdot \alpha), \sigma(a)\rangle - \langle (\alpha \cdot y) \cdot a, \tau(b)\rangle \\
&= \langle y \cdot \alpha, \sigma(a)\tau(b)\rangle - \langle \alpha \cdot y, \sigma(a)\tau(b)\rangle \\
&= \langle y \cdot \alpha - \alpha \cdot y, \sigma(a)\tau(b)\rangle = 0.
\end{aligned}$$

Thus for $\alpha \in \mathfrak{A}, a, b \in \mathcal{A}, y \in Y$ with $\sigma(a_0) = a$ and $\tau(b_0) = b$, we get

$$\langle y, (a \cdot \alpha)b - a(\alpha \cdot b)\rangle = \langle y, (\sigma(a_0) \cdot \alpha)\tau(b_0) - \sigma(a_0)(\alpha \cdot \tau(b_0))\rangle = 0.$$

By continuity of D, we see $D(a) \subseteq J^{\perp} = (\mathcal{A}/J)^*$. It immediately follows from the above definition that a module amenable Banach algebra $\mathcal{A}$ is (σ, τ)-weakly module amenable for all $\sigma, \tau \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. As we will see later in section four with some examples, the converse is false. Here and subsequently, we denote $\overbrace{\sigma \circ \sigma ... \circ \sigma}^{n-times}$ by σ^n for all $n \in \mathbb{N}$.

3.2. Proposition. *Let $\mathcal{A}$ be a Banach $\mathfrak{A}$-bimodule and $\sigma, \tau, \mu \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. If μ is an epimorphism and $\mathcal{A}$ is $(\sigma \circ \mu, \tau \circ \mu)$-weakly module amenable, then $\mathcal{A}$ is (σ, τ)-weakly module amenable. The converse is true if μ^2 is the identity map.*

Proof. Let Y be a commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$ and let $D : \mathcal{A} \to Y_{(\sigma,\tau)}$ be a module (σ, τ)-derivation. Then $D \circ \mu$ is a module $(\sigma \circ \mu, \tau \circ \mu)$-derivation. So there exists $y \in Y_{(\sigma\circ\mu,\tau\circ\mu)}$ such that for each $a \in \mathcal{A}$, $D(a) = y \cdot (\sigma \circ \mu)(a) - (\tau \circ \mu)(a) \cdot y$. Given $b \in \mathcal{A}$. Then there exists $a \in \mathcal{A}$ such that $\mu(a) = b$ and hence

$$D(b) = D(\mu(a)) = y \cdot \sigma(\mu(a)) - \tau(\mu(a)) \cdot y = y \cdot \sigma(b) - \tau(b) \cdot y.$$

Thus D is (σ, τ)-inner.

Conversely, suppose that $D : \mathcal{A} \to Y_{(\sigma\circ\mu,\tau\circ\mu)}$ is a module $(\sigma \circ \mu, \tau \circ \mu)$-derivation. It is easy to show that $\widetilde{D} = D \circ \mu^{-1}$ is in $Z_{\mathfrak{A}}(\mathcal{A}, (Y_{(\sigma,\tau)}))$. Thus there exists $y \in Y_{(\sigma,\tau)}$ so that for each $a \in \mathcal{A}$, $D(a) = y \cdot \sigma(a) - \tau(a) \cdot y$. We have

$$D(a) = D(\mu^{-1}(\mu(a))) = \widetilde{D}(\mu(a)) = y \cdot (\sigma \circ \mu)(a) - (\tau \circ \mu)(a) \cdot y,$$

for all $a \in \mathcal{A}$. Therefore D is $(\sigma \circ \mu, \tau \circ \mu)$-inner. □

3.3. Corollary. *Let $\mathcal{A}$ be a Banach $\mathfrak{A}$-module and $\sigma \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. Then the following statements hold:*

(i) *If σ is an epimorphism and $\mathcal{A}$ is (σ^n, σ^n)-weakly module amenable for some $n \in \mathbb{N}$, then $\mathcal{A}$ is weakly module amenable;*

(ii) *If $\mathcal{A}$ is weakly module amenable and σ^2 is the identity map, then $\mathcal{A}$ is (σ, σ)-weakly module amenable.*

3.4. Proposition. *Let $\sigma, \tau \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ such that σ be an epimorphism and let the restriction of σ on the set $\{ab - ba \mid a, b \in \mathcal{A}\}$ be the identity map. If $\mathcal{A}$ is (τ, τ)-weakly module amenable, then $\mathcal{A}$ is $(\sigma \circ \tau, \sigma \circ \tau)$-weakly module amenable.*

Proof. Let Y be a commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$ and let $D:\mathcal{A}\to Y_{(\sigma\circ\tau,\sigma\circ\tau)}$ be a module $(\sigma\circ\tau,\sigma\circ\tau)$-derivation. Define $\widetilde{D}:\mathcal{A}\to Y_{(\tau,\tau)}$ via $\langle\widetilde{D}(a),b\rangle:=\langle D(a),\sigma(b)\rangle$. It is easy to check that $\widetilde{D}$ is a module (τ,τ)-derivation and thus there exists $y\in Y_{(\tau,\tau)}$ such that $\widetilde{D}(a)=y\cdot\tau(a)-\tau(a)\cdot y$ for every $a\in\mathcal{A}$. Take $x\in\mathcal{A}$. Since σ is an epimorphism, there exists $b\in\mathcal{A}$ such that $x=\sigma(b)$. Then for each $a\in\mathcal{A}$, we get

$$\begin{aligned}\langle D(a),x\rangle&=\langle\widetilde{D}(a),b\rangle=\langle y\cdot\tau(a)-\tau(a)\cdot y,b\rangle\\&=\langle y,\sigma(\tau(a)b-b\tau(a))\rangle\\&=\langle y\cdot\sigma\circ\tau(a)-\sigma\circ\tau(a)\cdot y,x\rangle.\end{aligned}$$

It follows that D is an $(\sigma\circ\tau,\sigma\circ\tau)$-inner derivation. □

3.5. Corollary. *Let $\sigma\in\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ such that σ is an epimorphism and let the restriction of σ on $\widetilde{\mathcal{A}}=\{ab-ba\mid a,b\in\mathcal{A}\}$ be the identity map. If $\mathcal{A}$ is weakly module amenable, then $\mathcal{A}$ is (σ^n,σ^n)-weakly module amenable for all $n\in\mathbb{N}$.*

Recall that $\mathfrak{A}$ has a bounded approximate identity for $\mathcal{A}$ if there is a bounded net $\{\alpha_j\}$ in $\mathfrak{A}$ such that $\|\alpha_j\cdot a-a\|\to0$ and $\|a\cdot\alpha_j-a\|\to0$, for each $a\in\mathcal{A}$.

3.6. Proposition. *Let $\mathcal{A}$ be a Banach $\mathfrak{A}$-module and $\sigma,\tau\in\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. If $\mathfrak{A}$ has a bounded approximate identity, then (σ,τ)-weak amenability of $\mathcal{A}$ implies its (σ,τ)-weak module amenability.*

Proof. Let Y be a commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$ and let $D:\mathcal{A}\to Y_{(\sigma,\tau)}$ be a module (σ,τ)-derivation. If $\{\alpha_j\}$ is a bounded approximate identity for $\mathfrak{A}$, then by the Cohen factorization theorem [11], it is a bounded approximate identity for $\mathcal{A}$. Thus for each $a\in\mathcal{A}$ there are $\beta\in\mathfrak{A}$ and $b\in\mathcal{A}$ such that $a=\beta\cdot b$. Hence for each $a\in\mathcal{A}$ and $\rho\in\mathbb{C}$, we deduce that

$$\sigma(\rho a)=\sigma(\rho(\beta\cdot b))=\lim_j\sigma(\rho(\alpha_j\beta)\cdot b)=\lim_j\sigma(\rho\alpha_j\cdot a)=\lim_j\rho\alpha_j\cdot\sigma(a)=\rho\sigma(a).$$

Therefore σ is $\mathbb{C}$-linear. Similarly, $\tau\in\mathrm{Hom}(\mathcal{A})$. To complete of the proof, it is enough to show that D is $\mathbb{C}$-linear. Again, by the Cohen factorization theorem for each $a\in\mathcal{A}$ there are $\gamma\in\mathfrak{A}$ and $y\in Y$ such that $D(a)=\gamma\cdot y$. Then

$$\begin{aligned}D(\rho a)&=D(\rho(\beta\cdot b))=\lim_j D(\rho(\alpha_j\beta)\cdot b)\\&=\lim_j D(\rho\alpha_j\cdot a)=\lim_j\rho\alpha_j\cdot D(a)\\&=\lim_j\rho\alpha_j\cdot(\gamma\cdot y)=\rho(\gamma\cdot y)=\rho D(a).\end{aligned}$$

for all $a\in\mathcal{A}$ and $\rho\in\mathbb{C}$. □

3.7. Proposition. *Let $\mathcal{A}$ be a commutative Banach algebra and a commutative Banach $\mathfrak{A}$-bimodule. Suppose that $\sigma\in\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ such that $\sigma^2=\sigma$, and the range of σ is a closed ideal of $\mathcal{A}$. If $\mathcal{A}$ is weakly module amenable and $\mathfrak{A}$ has a bounded approximate identity for $\mathcal{A}$, then $\mathcal{A}$ is (σ,σ)-weakly module amenable.*

Proof. Let Y be a Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$ and let $D:\mathcal{A}\to Y_{(\sigma,\sigma)}$ be a module (σ,σ)-derivation. It is easily verified that the mapping $\overline{D}:\mathcal{A}\to Y$ is defined by $\langle\overline{D}(a),b\rangle:=\langle D(a),\sigma(b)\rangle$, is a module derivation. Thus there exists $y\in Y$ such that $\overline{D}(a)=y\cdot a-a\cdot y$. Since $\mathcal{A}=\ker(\sigma)\oplus\mathrm{Im}(\sigma)$, it follows from [4, Proposition 2.1] that $\mathcal{A}/\mathrm{Im}(\sigma)\cong\ker(\sigma)$ is a weakly module amenable Banach algebra. For every $a\in\mathcal{A}$, we put $a=a_1+a_2$ in which $a_1\in\ker(\sigma)$ and $a_2\in\mathrm{Im}(\sigma)$. By [4, Proposition 2.4] and the

Cohen factorization theorem, $(\ker(\sigma))^2$ is dense in $\ker(\sigma)$. Hence, there is a bounded net $(a_l b_l)_l \subset (\ker(\sigma))^2$ such that $a_l b_l \to a_1$, and

$$D(a_1) = \lim_l D(a_l b_l) = \lim_l (D(a_l) \cdot \sigma(b_l) - \sigma(a_l) \cdot D(b_l)) = 0.$$

This shows that $D(a) = D(\sigma(a))$ for all $a \in \mathcal{A}$. Now, suppose that $b \in \mathcal{A}$ such that $b = b_1 + b_2$ where $b_1 \in \ker(\sigma)$ and $b_2 \in \mathrm{Im}(\sigma)$. Take $a \in \mathcal{A}$ and the bounded nets $(a_{l_1} b_{l_2})_l \subset (\ker(\sigma))^2$ and $(a_{k_1} b_{k_2})_k \subset \mathcal{A}^2$ such that $a_{l_1} b_{l_2} \to b_1$ and $a_{k_1} b_{k_2} \to a$. Then, we have

$$\begin{aligned}
\langle D(a), b_1 \rangle &= \lim_l \lim_k \langle D(a_{k_1} b_{k_2}), a_{l_1} b_{l_2} \rangle \\
&= \lim_l \lim_k \langle D(a_{k_1}) \cdot \sigma(b_{k_2}) + \sigma(b_{k_1}) \cdot D(b_{k_2}), a_{l_1} b_{l_2} \rangle \\
&= \lim_l \lim_k \langle D(a_{k_1}), \sigma(b_{k_2}) a_{l_1} b_{l_2} \rangle + \lim_l \lim_k \langle D(b_{k_2}), a_{l_1} b_{l_2} \sigma(b_{k_1}) \rangle = 0.
\end{aligned}$$

The last equality follows from the fact that $\sigma(b_{k_2}) a_{l_1} b_{l_2}$ and $a_{l_1} b_{l_2} \sigma(b_{k_1})$ are in $\ker(\sigma) \cap \mathrm{Im}(\sigma) = \{0\}$. Also,

$$\begin{aligned}
\langle D(a), b_2 \rangle &= \langle D(a), \sigma(b_2) \rangle = \langle D(\sigma(a)), \sigma(b_2) \rangle \\
&= \langle \overline{D}(\sigma(a)), b_2 \rangle = \langle y \cdot \sigma(a) - \sigma(a) \cdot y, b_2 \rangle \\
&= \langle y, \sigma(a) b_2 - b_2 \sigma(a) \rangle = \langle \overline{D}(-b_2), \sigma(a) \rangle \\
&= \langle D(-\sigma(b_2)), \sigma^2(a) \rangle = \langle \overline{D}(-\sigma(b_2)), \sigma(a) \rangle \\
&= \langle y \cdot \sigma(a) - \sigma(a) \cdot y, b_2 \rangle.
\end{aligned}$$

The above computations show that $D \in B_{\mathfrak{A}}(\mathcal{A}, Y_{(\sigma,\sigma)})$. Therefore $\mathcal{A}$ is (σ,σ)-weakly module amenable. $\square$

Let $\mathcal{A}$ and $\mathfrak{A}$ be as in the previous section and X be a Banach $\mathcal{A}$-$\mathfrak{A}$-module with the compatible actions, and J be the corresponding closed ideals of $\mathcal{A}$. Let $\sigma \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$. Then for each $a, b \in \mathcal{A}$ and $\alpha \in \mathfrak{A}$, we have

$$\sigma((a \cdot \alpha) b - a(\alpha \cdot b)) = (\sigma(a) \cdot \alpha) \sigma(b) - \sigma(a)(\alpha \cdot \sigma(b)) \in J.$$

Since J is a closed ideal of $\mathcal{A}$ and σ is continuous, $\sigma(J) \subseteq J$. Therefore, the mapping $\widehat{\sigma} : \mathcal{A}/J \longrightarrow \mathcal{A}/J$ is defined by $\widehat{\sigma}(a + J) = \sigma(a) + J$ is well defined.

Recall that a left Banach $\mathcal{A}$-module X is called a *left essential* $\mathcal{A}$-module if the linear span of $\mathcal{A} \cdot X = \{a \cdot x : a \in \mathcal{A}, x \in X\}$ is dense in X. Right essential $\mathcal{A}$-modules and (two-sided) essential $\mathcal{A}$-bimodules are defined similarly. We remark that if $\mathcal{A}$ is an essential left (right) $\mathfrak{A}$-module, then every $\mathfrak{A}$-module homomorphism σ is also a linear homomorphism. If $a \in \mathcal{A}$, then there is a sequence $(b_n) \subseteq \mathcal{A} \cdot \mathfrak{A}$ such that $\lim_n b_n = a$. Assume that $b_n = \sum_{m=1}^{K_n} \alpha_{n,m} a_{n,m}$ for some finite sequences $(a_{n,m})_{m=1}^{m=K_n} \subseteq \mathcal{A}$ and $(\alpha_{n,m})_{m=1}^{m=K_n} \subseteq \mathfrak{A}$. Let $t \in \mathbb{C}$. Then

$$\begin{aligned}
\sigma(t b_n) &= \sigma(t \textstyle\sum_{m=1}^{K_n} \alpha_{n,m} \cdot a_{n,m}) = \sum_{m=1}^{K_n} \sigma((t\alpha_{n,m}) \cdot a_{n,m}) \\
&= \textstyle\sum_{m=1}^{K_n} (t\alpha_{n,m}) \cdot \sigma(a_{n,m}) = \sum_{m=1}^{K_n} t\sigma(\alpha_{n,m} \cdot a_{n,m}) = t\sigma(b_n),
\end{aligned}$$

and so by the continuity of σ, $\sigma(ta) = t\sigma(a)$. By definition of $\widehat{\sigma}$, it is also $\mathbb{C}$-linear.

We say the Banach algebra $\mathfrak{A}$ acts trivially on $\mathcal{A}$ from left (right) if for each $\alpha \in \mathfrak{A}$ and $a \in \mathcal{A}$, $\alpha \cdot a = \phi(\alpha) a$ ($a \cdot \alpha = \phi(\alpha) a$), where ϕ is a continuous linear functional on $\mathfrak{A}$. The following lemma is proved in [3, Lemma 3.1].

3.8. Lemma. *Let $\mathcal{A}$ be a Banach algebra and Banach $\mathfrak{A}$-module with compatible actions, and J_0 be a closed ideal of $\mathcal{A}$ such that $J \subseteq J_0$. If $\mathcal{A}/J_0$ has a left or right identity $e + J_0$,* *then for each $\alpha \in \mathfrak{A}$ and $a \in \mathcal{A}$ we have $a \cdot \alpha - \alpha \cdot a \in J_0$, i.e., $\mathcal{A}/J_0$ is a commutative Banach $\mathfrak{A}$-module.*

The concept of $(\widehat{\sigma}, \widehat{\tau})$-weak amenability of $\mathcal{A}/J$ has been investigated in [8]. Relating to this, we now prove the main result in this section which gives the sufficient conditions for being (σ, τ)-weakly module amenable of a Banach algebra.

3.9. Theorem. *Let $\mathcal{A}$ be a Banach $\mathfrak{A}$-module with trivial left action, and let σ, τ be in* $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ *and $\mathcal{A}/J$ has an identity. If $\mathcal{A}$ is a right essential $\mathfrak{A}$-module, then $(\widehat{\sigma}, \widehat{\tau})$-weak amenability of $\mathcal{A}/J$ implies (σ, τ)-weak module amenability of $\mathcal{A}$. The converse is true if σ and τ are epimorphisms.*

Proof. Let Y be a commutative Banach $\mathcal{A}$-$\mathfrak{A}$-submodule of $\mathcal{A}^*$, and let $D : \mathcal{A} \to Y_{(\sigma,\tau)}$ be a module (σ, τ)-derivation. For $y \in Y, a, b \in \mathcal{A}$ and $\alpha \in \mathfrak{A}$, we get

$$\begin{aligned}((a \cdot \alpha)b - a(\alpha \cdot b)) \cdot y &= (a \cdot \alpha) \cdot (b \cdot y) - a \cdot ((\alpha \cdot b) \cdot y)\\ &= a \cdot (\alpha \cdot (b \cdot y)) - a \cdot (\alpha \cdot (b \cdot y)) = 0.\end{aligned}$$

Hence, $J \cdot Y = \{0\}$. Similarly, we have $Y \cdot J = \{0\}$. Therefore, the following module actions are well-defined

$$(a + J) \cdot y := a \cdot y, \quad y \cdot (a + J) := y \cdot a \quad (y \in Y, a \in \mathcal{A}).$$

Thus Y is a Banach $\mathcal{A}/J$-$\mathfrak{A}$-module. Define $\widetilde{D} : \mathcal{A}/J \longrightarrow Y \subseteq J^{\perp} = ((\mathcal{A}/J)_{(\widehat{\sigma},\widehat{\tau})})^*$ via $\widetilde{D}(a + J) = D(a)$. For each $\alpha \in \mathfrak{A}$ and $a, b \in \mathcal{A}$ we have

$$\begin{aligned}D((a \cdot \alpha)b - a(\alpha \cdot b)) &= D((a \cdot \alpha)b) - D(a(\alpha \cdot b))\\ &= D(a \cdot \alpha) \cdot \sigma(b) + \tau(a \cdot \alpha) \cdot D(b)\\ &- (D(a) \cdot \sigma(\alpha \cdot b) - \tau(a) \cdot D(\alpha \cdot b))\\ &= (D(a) \cdot \alpha) \cdot \sigma(b) - D(a) \cdot (\alpha \cdot \sigma(b))\\ &+ (\tau(a) \cdot \alpha) \cdot D(b) - \tau(a) \cdot (\alpha \cdot D(b)) = 0.\end{aligned}$$

It means that D vanishes on J. Therefore $\widetilde{D}$ is well-defined. For each a, b in $\mathcal{A}$ we have

$$\begin{aligned}\tilde{D}(ab + J) = D(ab) &= D(a) \cdot \sigma(b) + \tau(a) \cdot D(b)\\ &= \tilde{D}(a + J) \cdot (\sigma(b) + J) + (\tau(a) + J) \cdot \tilde{D}(b + J)\\ &= \tilde{D}(a + J) \cdot \widehat{\sigma}(b + J) + \widehat{\tau}(a + J) \cdot \tilde{D}(b + J).\end{aligned}$$

Since $\mathcal{A}$ is a right essential $\mathfrak{A}$-module, $\widehat{\sigma}$ and $\widehat{\tau}$ are homomorphism. Thus $\widehat{\sigma}, \widehat{\tau} \in \mathrm{Hom}(\mathcal{A}/J)$. Now, it follows from the above discussion that $\tilde{D}$ is also $\mathbb{C}$-linear, and so it is $(\widehat{\sigma}, \widehat{\tau})$-inner. Hence there exists $y \in Y$ such that

$$D(a) = \tilde{D}(a + J) = y \cdot \widehat{\sigma}(a + J) - \widehat{\tau}(a + J) \cdot y = y \cdot \sigma(a) - \tau(a) \cdot y.$$

Therefore D is a module (σ, τ)-inner derivation.

Conversely, suppose that $\sigma, \tau \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ are epimorphisms, and $D : \mathcal{A}/J \longrightarrow ((\mathcal{A}/J)_{(\widehat{\sigma},\widehat{\tau})})^*$ is a $(\widehat{\sigma}, \widehat{\tau})$-derivation. We define $\tilde{D} : \mathcal{A} \longrightarrow ((\mathcal{A}/J)_{(\sigma,\tau)})^*$ by $\tilde{D}(a) = D(a + J)$, for all $a \in \mathcal{A}$. Lemma 3.8 shows that when $\mathfrak{A}$ acts on $\mathcal{A}$ trivially from left or right, then $\mathcal{A}/J$ is a commutative $\mathfrak{A}$-module and thus $Y = J^{\perp} \subseteq \mathcal{A}^*$. Hence $\tilde{D}$ could be considered as a map from $\mathcal{A}$ to Y. Now, for each $\alpha \in \mathfrak{A}$ and $a \in \mathcal{A}$ we have

$$\tilde{D}(\alpha \cdot a) = D(\alpha \cdot a + J) = D(\phi(\alpha)a + J) = \phi(\alpha)D(a + J) = \alpha \cdot \tilde{D}(a)$$

and

$$\tilde{D}(a \cdot \alpha) = D(a \cdot \alpha + J) = D(\phi(\alpha)a + J) = \phi(\alpha)D(a + J) = \tilde{D}(a) \cdot \alpha.$$

Also, for $a, b \in \mathcal{A}$ we obtain $\tilde{D}(ab) = \tilde{D}(a) \cdot \sigma(b) + \tau(a) \cdot \tilde{D}(b)$. Thus $\tilde{D}$ is a (σ, τ)-module derivation. Due to (σ, τ)-weak module amenability of $\mathcal{A}$, there exists $y \in Y \cong$

$((\mathcal{A}/J)_{(\sigma,\tau)})^*$ such that $\tilde{D}(a) = \sigma(a) \cdot y - y \cdot \tau(a)$, and so $D(a+J) = \widehat{\sigma}(a+J) \cdot y - y \cdot \widehat{\tau}(a+J)$. □

The Banach algebras with compatible $\mathfrak{A}$-module structure could be considered as objects of a category $\mathfrak{C}_{\mathfrak{A}}$ whose morphisms are bounded $\mathfrak{A}$-module maps. We are interested in the case where $\mathfrak{A}$ is an injective object in $\mathfrak{C}_{\mathfrak{A}}$, that is for any objects $A, B \in \mathfrak{C}_{\mathfrak{A}}$ and monomorphism $\theta : B \longrightarrow A$ and morphism $\mu : B \longrightarrow \mathfrak{A}$, there exists a morphism $\widetilde{\mu} : A \longrightarrow \mathfrak{A}$ such that $\mu = \widetilde{\mu} \circ \theta$. This is the case when $\mathfrak{A} = \mathbb{C}$ (Hahn Banach Theorem).

3.10. Proposition. *Let $\mathcal{A}$ be a commutative $\mathfrak{A}$-module and let σ, τ be in $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ such that $\sigma(a)b = a\tau(b)$ for all $a, b \in \mathcal{A}$. Also let $\mathfrak{A}$ be injective and has a bounded approximate identity. If $\mathcal{A}$ is (σ,τ)-weakly module amenable, then span $(\mathcal{A}\mathfrak{A}\mathcal{A})$ is dense in $\mathcal{A}$.*

Proof. Let B be the linear span of $(\mathcal{A}\mathfrak{A}\mathcal{A})$. Suppose that $\overline{B} \neq \mathcal{A}$. Take $a_0 \in \mathcal{A}\backslash\overline{B}$ and $f_1 \in \mathcal{A}^*$ such that $f_1(a_0) = 1$ and $f_1|_{\overline{B}} = 0$. Since a_0 is not in $\overline{B}$, similar to the proof of [2, lemma 2.1] we can construct an epimorphism $f_2 : \mathcal{A} \longrightarrow \mathfrak{A}$ such that $f_2|_{\overline{B}} = 0$ and $f_2(a_0) = 1$. Define $D : \mathcal{A} \longrightarrow ((\mathcal{A})_{(\sigma,\tau)})^*$ via $D(a) = f_2(a) \cdot f_1$ for all $a \in \mathcal{A}$. Then D is (σ,τ)-module derivation and hence there exists $g \in (\mathcal{A}_{(\sigma,\tau)})^*$ such that $D(a) = g \cdot \sigma(a) - \tau(a) \cdot g$, for all $a \in \mathcal{A}$. Thus, we have

$$\begin{aligned} 1 &= f_2(a_0)f_1(a_0) = \langle D(a_0), a_0\rangle \\ &= \langle g \cdot \sigma(a_0) - \tau(a_0) \cdot g, a_0\rangle \\ &= \langle g, \sigma(a_0)a_0 - \tau(a_0)a_0\rangle = 0, \end{aligned}$$

which is a contradiction. □

3.11. Corollary. *With the hypotheses of the above Proposition, $\mathcal{A}$ is $(0,0)$-weakly module amenable if and only if span $(\mathcal{A}\mathfrak{A}\mathcal{A})$ is dense in $\mathcal{A}$.*

Proof. Let $D : \mathcal{A} \to (\mathcal{A}_{(0,0)})^*$ be a $(0,0)$-module derivation. Then we have $D(\mathcal{A}\mathfrak{A}\mathcal{A}) = \{0\}$. Since D is continuous, we have $D = 0$. So D is $(0,0)$-inner. Conversely, let $\mathcal{A}$ be $(0,0)$-weakly amenable. Then by Proposition (3.10), $\overline{\mathcal{A}\mathfrak{A}\mathcal{A}} = \mathcal{A}$. □

3.12. Remark. Let $\mathcal{A}$ be a commutative $\mathfrak{A}$-module and let $\sigma, \tau \in \mathrm{Hom}_{\mathfrak{A}}(\mathcal{A})$ such that $\sigma(a)b = a\tau(b)$ for all $a, b \in \mathcal{A}$. Then the second adjoints σ'' and τ'' belong to $\mathrm{Hom}_{\mathfrak{A}}(\mathcal{A}^{**})$ and are also w^*-w^*-continuous. We thus can show that $\sigma''(F)\Box G = F\Box\tau''(G)$, where $\Box$ is the first Arens product on the second dual $\mathcal{A}^{**}$ (for more information about this product see [10]). Now, if $\mathcal{A}^{**}$ is (σ'', τ'')-weakly amenable then by Proposition 3.10, $\overline{\mathcal{A}^{**}\mathfrak{A}\mathcal{A}^{**}} = \mathcal{A}^{**}$. It follows from the proof of [2, Proposition 3.6] that $\overline{\mathcal{A}\mathfrak{A}\mathcal{A}} = \mathcal{A}$. Therefore $\mathcal{A}$ is $(0,0)$-weakly amenable by Corollary 3.11.

4. (σ, τ)-weak module amenability of semigroup algebras

Let S be an (discrete) inverse semigroup with the set of idempotents E_S (or E), where the order of E is defined by

$$e \leq d \Longleftrightarrow ed = e \quad (e, d \in E).$$

It is easy to show that E is a (commutative) subsemigroup of S [17, Theorem V.1.2]. In particular $\ell^1(E)$ could be regarded as a subalgebra of $\ell^1(S)$, and thereby $\ell^1(S)$ is a Banach algebra and a Banach $\ell^1(E)$-module with compatible actions [1]. We consider the following module actions $\ell^1(E)$ on $\ell^1(S)$:

$$\delta_e \cdot \delta_s = \delta_s, \ \delta_s \cdot \delta_e = \delta_{se} = \delta_s * \delta_e \quad (s \in S, e \in E). \tag{4.1}$$

If ϕ is a continuous linear function on $\ell^1(E)$, then for each $e \in E$ we have $\phi(\delta_e) = 1$. So for each $f = \sum_{e\in E} f(e)\delta_e \in \ell^1(E)$ and $g = \sum_{s\in S} g(s)\delta_s \in \ell^1(S)$, we get

$$\begin{aligned} f \cdot g &= (\sum_{e\in E} f(e)\delta_e) \cdot (\sum_{s\in S} g(s)\delta_s) = \sum_{s\in S, e\in E} f(e)g(s)\delta_e \cdot \delta_s \\ &= \sum_{s\in S, e\in E} f(e)g(s) \cdot \delta_s = (\sum_{e\in E} f(e))(\sum_{s\in S} g(s)\delta_s) = \phi(f)g. \end{aligned}$$

Therefore multiplication from left is trivial. In this case, the ideal J (see section 3) is the closed linear span of $\{\delta_{set} - \delta_{st} :\ s, t \in S, e \in E\}$. We consider an equivalence relation on S as follows:

$$s \approx t \Longleftrightarrow \delta_s - \delta_t \in J \quad (s, t \in S).$$

For an inverse semigroup S, the quotient $S/\approx$ is a discrete group (see [3] and [23]). As in [24, Theorem 3.3], we may observe that $\ell^1(S)/J \cong \ell^1(S/\approx)$. We consider the following module actions $\ell^1(E)$ on $\ell^1(S)/J \cong \ell^1(S/\approx)$:

$$\delta_e \cdot (\delta_s + J) = \delta_s + J,\ (\delta_s + J) \cdot \delta_e = \delta_{se} + J \quad (s \in S, e \in E).$$

Indeed $\delta_s - \delta_{se} \in J$ if and only if $\delta_{st} - \delta_{set} \in J$, for all $s, t \in S, e \in E$. Therefore $\ell^1(S/\approx)$ is a commutative $\ell^1(E)$-bimodule. For each $\sigma \in \mathrm{Hom}_{\ell^1(E)}(\ell^1(S))$, we define $\widehat{\sigma}$ in $\mathrm{Hom}(\ell^1(S/\approx))$ by $\widehat{\sigma}(\delta_{[s]}) = \delta_{[\sigma(s)]}$ and extend by linearity, where $[s]$ denote the equivalence class of s in $S/\approx$ (see the explanations after Proposition 3.7). We see that all conditions of Theorem 3.9 hold for $\sigma, \tau \in \mathrm{Hom}_{\ell^1(E)}(\ell^1(S))$ which are also epimorphism. Now, if $\ell^1(S)$ is (σ, τ)-weakly module amenable then $\ell^1(S/\approx)$ is $(\widehat{\sigma}, \widehat{\tau})$-weakly amenable. We are now going to prove the main result in this section.

4.1. Theorem. *Let S be an inverse semigroup with the set of idempotents E. Then for each σ and τ in $\mathrm{Hom}_{\ell^1(E)}(\ell^1(S))$, the semigroup algebra $\ell^1(S)$ is (σ, τ)-weakly module amenable as an $\ell^1(E)$-module, with trivial left action.*

Proof. Suppose firstly that σ or τ is zero map. Since $S/\approx$ is a discrete group, the group algebra $\ell^1(S/\approx)$ has an identity, and thus $\ell^1(S/\approx)$ is $(\widehat{\sigma}, 0)$ and $(0, \widehat{\sigma})$-weakly amenable by [8, Example 4.2]. With the actions considered in (4.1), for each $f \in \ell^1(S)$, we have

$$f = \sum_{s\in S} f(s)\delta_s = \sum_{s\in S} f(s)\delta_s * \delta_{s^*s} = \sum_{s\in S} f(s)\delta_s \cdot \delta_{s^*s}.$$

Consequently f belongs to the closed linear span of $\ell^1(S) \cdot \ell^1(E) = \{\delta_s \cdot \delta_e : e \in E, s \in S\}$. This shows that $\ell^1(S)$ is a right essential $\ell^1(E)$-module. For $\mathcal{A} = \ell^1(S)$ and $\mathfrak{A} = \ell^1(E)$, the result of this case follows from Theorem 3.9. For the case that both σ and τ are non-zero homomorphisms, it is proved in [14, Theorem 2.5] that for any locally compact group G, the group algebra $L^1(G)$ is (φ, ψ)-weakly amenable for all $\varphi, \psi \in \mathrm{Hom}(L^1(G))$. In particular, $\ell^1(S/\approx)$ is $(\widehat{\sigma}, \widehat{\tau})$-weakly amenable. Now, Theorem 3.9 again shows that $\ell^1(S)$ is (σ, τ)-weakly module amenable. □

Note that for an amenable inverse semigroup S, $\ell^1(S)$ is module $\ell^1(E)$-amenable [1, Theorem 3.1] and so, it is module (σ, τ)-amenable [7, Corollary 2.3]. We close this section by two examples.

4.2. Example. *Let S be a commutative inverse semigroup. Then $\ell^1(S)$ is a commutative Banach algebra and commutative Banach $\ell^1(E)$-module with the following actions:*

$$\delta_e \cdot \delta_s = \delta_s \cdot \delta_e = \delta_{es} \quad (s \in S, e \in E).$$

We consider the mapping σ as follows:

$$\sigma : \ell^1(S) \longrightarrow \ell^1(S);\ \sum_{s\in S} f(s)\delta_s \mapsto \sum_{s\in S} \overline{f(s)}\delta_{s^*} \quad (s \in S),$$

where $\overline{f(s)}$ is the complex conjugate of $f(s)$. Obviously $\sigma \in Hom_{\ell^1(E)}(\ell^1(S))$. Also, σ is also $\mathbb{C}$-linear and σ^2 is the identity map. It is shown in [4, Theorem 3.1] *that $\ell^1(S)$ is weakly module amenable. Now it follows from Corollary 3.3 that $\ell^1(S)$ is (σ,σ)-weakly module amenable. Note that if S is not amenable, $\ell^1(S)$ is not module amenable* [1, Theorem 3.1].

4.3. Example. *Let S be an inverse semigroup with the set of idempotents E. Let $C^*(S)$ be the enveloping C^*-algebra of $\ell^1(S)$ (see* [13]*). Then by continuity, the action of $\ell^1(E)$ on $\ell^1(S)$ extends to an action of $C^*(E)$ on $C^*(S)$. The C^*-algebra $C^*(E)$ has a bounded approximate identity, and so it is $(\sigma, 0)$ and $(0,\sigma)$-weakly module amenable by Proposition 3.6 and* [8, Example 4.2]*, for all $\sigma \in Hom_{C^*(E)}(C^*(S))$. Now, suppose that σ^2 is the identity map (see Example 4.2). Since $C^*(S)$ is weakly amenable* [16, Theorem 1.10]*, $C^*(S)$ is (σ,σ)-weakly module amenable by Corollary 3.3. However, if $C^*(S)$ is nuclear then it is amenable* [15]*. By* [1, Proposition 2.1]*, $C^*(S)$ is module amenable as an $C^*(E)$-module. Therefore $C^*(S)$ is module (σ,τ)-amenable, for all $\sigma,\tau \in Hom_{C^*(E)}(C^*(S))$ by* [7, Corollary 2.3].

Acknowledgements. The author express his sincere thanks to the referee for the careful and detailed reading of the manuscript and very helpful suggestions that improved the manuscript substantially.

References

[1] M. Amini, *Module amenability for semigroup algebras*, Semigroup Forum **69**, 243–254, 2004.
[2] M. Amini and A. Bodaghi, *Module amenability and weak module amenability for second dual of Banach algeras*, Chamchuri J. Math. **2**, No. 1, 57–71, 2010.
[3] M. Amini, A. Bodaghi and D. Ebrahimi Bagha, *Module amenability of the second dual and module topological center of semigroup algebras*, Semigroup Forum **80**, 302–312, 2010.
[4] M. Amini and D. Ebrahimi Bagha, *Weak Module Amenability for semigroup algebras*, Semigroup Forum **71**, 18–26, 2005.
[5] W. G. Bade, P. C. Curtis and H. G. Dales, *Amenability and weak amenability for Beurling and Lipschits algebra*, Proc. London Math. Soc. **55** (3), 359–377, 1987.
[6] A. Bodaghi, *Semigroup algebras and their weak module amenability*, J. Appl. Func. Anal. **7**, No. 4, 332–338, 2012.
[7] A. Bodaghi, *Module (φ,ψ)-amenability of Banach algeras*, Arch. Math (Brno) **46**, No. 4, 227–235, 2010.
[8] A. Bodaghi, M. Eshaghi Gordji and A. R. Medghalchi, *A generalization of the weak amenability of Banach algebras*, Banach J. Math. Anal. **3**, No. 1, 131–142, 2009.
[9] S. Bowling and J. Duncan, *Order cohomology of Banach semigroup algebras*, Semigroup Forum **56**, 130–145, 1998.
[10] H. G. Dales, *Banach Algebras and Automatic Continuity*, Oxford University Press, Oxford, 2000.
[11] R. S. Doran and J. Wichmann, *Approximate Identities and Factorization in Banch Modules*, Lecture Notes in Mathematics 768, Springer, Berlin, 1979.
[12] J. Duncan and I. Namioka, *Amenability of inverse semigroups and their semigroup algebras*, Proc. Roy. Soc. Edinburgh **80A**, 309-321, 1988.
[13] J. Duncan and A. L. T. Paterson, *C^*-algebras of inverse semigroups*, Proc. Roy. Soc. Edinburgh Soc. **28**, 41–58, 1985.
[14] M. Eshaghi Gordji and A. Jabbari, *Generalization of weak amenability of group algebras*, preprint.
[15] U. Haagerup, *All nuclear C^*-algebras are amenable*, Invent. Math. **74**, 305–319, 1983.
[16] U. Haagerup and N. J. Laustsen, *Weak amenability of C^*-algebras and theorem of Goldstien*, In Banach algebra 97. 223–243.
[17] J. M. Howie, *An Introduction to Semigroup Theory*, Academic Press, London, 1976.

[18] B. E. Johnson, *Cohomology in Banach Algebras*, Memoirs Amer. Math. Soc. **127**, American Math. Soc., Providence, 1972.

[19] B. E. Johnson, *Derivation from $L^1(G)$ into $L^1(G)$ and $L^\infty(G)$*, Lecture Note in Math. **1359**, 191–198, 1988.

[20] B. E. Johnson, *Weak amenability of group Algebras*, Bull. London Math. Soc. **23** (3), 281–284, 1991.

[21] M. S. Moslehian and A. N. Motlagh, *Some notes on (σ, τ)-amenability of Banach algebras* , Stud. Univ. Babes-Bolyai. Math. **53**, No. 3, 57–68, 2008.

[22] A. L. T. Paterson, *Amenability*, American Math. Soc., Providence, Rhode Island, 1988.

[23] H. Pourmahmood-Aghababa, *(Super)Module amenability, module topological center and semigroup algebras*, Semigroup Forum. **81**, 344–356, 2010.

[24] R. Rezavand, M. Amini, M. H. Sattari and D. Ebrahimi Bagha, *Module Arens regularity for semigroup algebras*, Semigroup Forum. **77**, 300–305, 2008.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 97 – 115

POWER STUDY OF CIRCULAR ANOVA TEST AGAINST NONPARAMETRIC ALTERNATIVES

Feridun Tasdan * and Ozgur Yeniay †

Received 24 : 12 : 2012 : Accepted 17 : 10 : 2013

Abstract

This study compares circular ANOVA against bootstrap test, uniform scores test and Rao's test of homogeneity which are considered nonparametric alternatives. Circular ANOVA is one-way analysis of variance method to test the equality of mean directions in circular data analysis, but it requires some assumptions. The main assumption for circular ANOVA is that all r-independent samples must come from von Mises distribution with equal directional means and equal concentration parameters. On the other hand, nonparametric alternatives are distribution free methods and, therefore, does not require having von Mises distribution or equality of parameters. Literature of circular statistics is very limited on the comparison of these tests; therefore, a power simulation study is performed to compute the power of circular ANOVA against the nonparametric alternatives under assumptions of von Mises and non-von Mises populations. Power simulation study shows that bootstrap and uniform scores tests perform slightly better than circular ANOVA if the common concentration parameter, κ, is less than 1 under the assumption of von Mises distribution. If $\kappa \geq 2$, then bootstrap and circular ANOVA perform better than the other alternatives. Rao's test of homogeneity requires very large samples in order to reach the same power levels of competitive tests in this study. Finally, uniform scores tests performs better than circular ANOVA and bootstrap test if the sample sizes are small and the data comes from mixed von Mises distributions or wrapped Cauchy.

Keywords: Keywords:Bootstrap, Circular Data, Circular ANOVA, von Mises Distribution, Seasonal Wind Directions, Uniform Scores Test, Rao's Test.

2000 AMS Classification:

*Department of Mathematics Western Illinois University Macomb, IL 61455
Email: (F. Tasdan) F-Tasdan@wiu.edu

†Department of Statistics Hacettepe University Beytepe, Ankara, Turkey.
Email: (O. Yeniay) yeniay@hacettepe.edu.tr

1. Introduction

The history of circular data problems, which can be seen in biology, geography, medicine, meteorology, oceanography and many other fields, goes back to the 1950s, but we have seen more publications in the last 25 years. Several textbooks and many papers have been published in recent years about the circular data problems. [11], [16], [6], [9],[2] are excellent resources for circular data problems. Technological developments in computers and programming made it possible to analyze large or complicated circular data problems. There are several computer programs currently available for the analysis of circular data problems. One of these is R program with circular package, which is jointly developed by [1]. It is called "circular" in R package repository. In fact, some of the results in this study are obtained from this circular package.

Circular data is obtained by measuring directions or arrival times of subjects with respect to a reference point on the unit circle. This reference point or the choice of the origin is arbitrary and the final conclusions should not depend on it. For example, North can be taken as a reference point (considered as 0 degrees) on the unit circle. Therefore, circular data will have a domain of $[0, 2\pi)$ in radians or $[0, 360)$ in degrees depending on the definition of the problem. If the arrival times of patients to an emergency room are the main interest, then the data can be recorded in 24 hour clock notation (domain of [0:00, 24:00)) and can later be converted to the angles on the unit circle.

Moreover, two or more sample circular data problems have been increasingly common in recent years. Watson and Williams ([17]) introduced a test for the equality of r-population means in circular data problems. This test can be considered an equivalent of one-way ANOVA in the traditional linear data problems. In later years, [11] and [14] modified the Watson-Williams test for certain conditions, which are given in Section 2. Nonparametric tests are also developed for two or more sample circular data problems. The test of homogeneity of r-populations is proposed by [11] and [18]. It is called uniform scores test or Mardia-Watson-Wheeler test in the literature. The test is based on ranks of the combined samples, but it is very sensitive to the existence of ties. [2] suggested that Mardia-Watson-Wheeler test should not be used if there are many ties in the data, but a few ties could be broken by a randomization or average methods. [13] introduced a nonparametric test called "Rao's test of homogeneity" for the equality of r-populations (homogeneity of populations). The details of the test are given in Section 3. Also, a bootstrap based test for the equality of r-population means is available and promoted by [6] especially if the sample sizes are less than 10 or assumptions do not meet in circular ANOVA test. The next section will give some insight about the multi-sample method called circular ANOVA in circular statistics.

2. Circular ANOVA

Circular ANOVA (One-Way Analysis of Variance) has been proposed by [17] and later modified by [11] based on suggestions by [14]. The theory of circular ANOVA is discussed extensively by [6], [11] and [9] on pages 125-128. In an another important paper, [8] also discuss the drawbacks of suggestions by [14]. The first assumption of the circular ANOVA is that all random samples should come from von Mises distribution with a common concentration parameter κ such that $H_0 : \kappa_1 = \kappa_2 = .. = \kappa_r = \kappa$ (test of homogeneity of kappa). If the assumptions of having von Mises distribution and the test of homogeneity of the kappa parameters fail, then [6] proposes nonparametric approaches for the analysis of two or more samples in circular data. If the sample sizes $n_1, .., n_r$ are less than 25, the bootstrap approach is heavily emphasized by [6]. There are several options (analogous to Levene's test in linear data) available for testing that all κ parameters are equal. We will introduce one of them in the next section when we

Source	DF	SS	MS	F
Between Samples	$r-1$	$\sum R_i - R$	$(\sum R_i - R)/(r-1) = I$	$F_t = I/II$
Within Samples	$N-r$	$N - \sum R_i$	$(N - \sum R_i)/(N-r) = II$	
Total	$N-1$	$N-R$		

perform the large sample example with R's circular package. There is a necessity that either the common concentration parameter κ is given or must be estimated from the data. So, [6] proposes $\hat{\kappa} = median\{\hat{\kappa_1}, \hat{\kappa_2}, .., \hat{\kappa_r}\}$ as an estimator of κ if it is unknown. Depending on the value of the common concentration parameter, there are several alternative approaches for circular ANOVA. [6] categorizes these approaches in three sections: $\kappa \geq 2$, $1 < \kappa < 2$, and $\kappa \leq 1$.

First, assume that $\kappa \geq 2$ and state the hypothesis that

$$H_0 : \mu_1 = \mu_2 = .. = \mu_r \text{ vs } H_1 : \text{At least two are distinct.}$$

Let θ_{ij} (for $i = 1, .., r$ and $j = 1, .., n_i$) shows angular observations coming from a circular distribution on the unit circle. Let R be the resultant length of all $N(N = n_1 + n_2 + .. + n_r)$ observations. The variable R can be computed by using all observations $(\theta_1, ..., \theta_N)$ or [6] provided the following formula that uses individual sample resultant lengths $(R_1, R_2, ..R_r)$ and mean directions $(\overline{\theta_i})$. Let

$$R = [(\sum_{i=1}^{r} R_i \cos(\overline{\theta_i}))^2 + (\sum_{i=1}^{r} R_i \sin(\overline{\theta_i}))^2]^{0.5} \tag{2.1}$$

The test statistic for circular ANOVA is defined by

$$F_t = (N-r)(\sum_{i=1}^{r} R_i - R)/[(r-1)(N - \sum_{i=1}^{r} R_i)] \tag{2.2}$$

where F_t has an F distribution with r-1 and N-r degrees of freedoms. We reject the test if $F_t > F_{r-1,N-r}$. One advantage of this test is that the F critical values can be found in many statistics books. [11] defined a circular ANOVA table summarizes the result:

If $1 < \kappa < 2$, [14] proposes a modified test that uses correction a factor and it is defined as $F_t' = [1 + 3/(8 * \hat{\kappa})]F_t$. If $\kappa \leq 1$, then [11] proposes an approximate likelihood ratio test which is defined below,

$$-2\log_e \lambda \doteq \frac{2}{N}\{(\sum_{i=1}^{r} R_i)^2 - R^2\} = U. \tag{2.3}$$

where for a large N, U has an approximate chi-square (χ^2) distribution with $r-1$ degrees of freedom when H_0 is true. The expression for λ can be derived from the equation (2.3). Details of this approximation can be seen in [11] on page 164.

3. Nonparametric Tests

Recall that circular ANOVA is discussed in Section 2 and requires multiple assumptions: (i) r-samples are coming from (at least approximately) von Mises distribution, (ii) the concentration parameters (κ) are equal, (iii) the value of the common concentration parameter is larger than $\hat{\kappa} > 1$. In many real life situations, one or more of these assumptions may not be satisfied. Therefore, alternative tests for circular ANOVA must be considered in order to avoid those assumptions or replace circular ANOVA if the assumptions are not satisfied. Bootstrap test is one approach that avoids these assumptions listed above. Mardia-Watson-Wheeler test (also called uniform scores test) and Rao's test of homogeneity are also nonparametric tests that they do not require having von Mises distribution assumption or the equality of parameters. One disadvantage for Rao's test is that it requires sufficiently large sample sizes. These nonparametric alternatives are discussed in the following sections.

3.1. Bootstrap Test. The bootstrap method was first introduced by [3] and became popular in recent years due to technological advances in the computer sciences. With the bootstrap method, the original sample is treated as the population and a resampling procedure is performed on it. This is done by randomly drawing a sample of size n from the original sample (size n) with replacement. [4] introduced many bootstrap methods as an alternative to parametric methods. [5] and [7] studied bootstrap method for circular data problems extensively. An algorithm based on bootstrap test for circular data has also been discussed by [12]. They showed that the bootstrap based hypothesis testing method to test the equality of peak months for fish populations could be used by considering the months as circular variables. In comparison to the circular ANOVA, bootstrap test approach uses the bootstrap estimate of the test statistic (F statistic) from the combined samples of circular data. In each bootstrap step, bootstrap estimate of the test statistics ($F^\star$) is found and compared with the original test statistic which is computed from the original samples. Then an estimated significance value (p-value) of the bootstrap test is calculated by first finding the number of bootstrap test statistic which is greater than the original test statistics and dividing the result with the number of bootstrap runs (B replications). If the estimated significance value is less than or equal to level of significance, it means that there is a significant difference among the population mean directions and, therefore, $H_0 : \mu_1 = \mu_2 = ... = \mu_r$ is rejected.

The following bootstrap test algorithm can be defined in order to obtain the bootstrap significance value or p-value. The algorithm is somewhat similar to [6]'s definition of the bootstrap test for two or more samples but the main difference is that [6] does not combine the samples whereas the proposed bootstrap test combines the samples to create one large sample and draws a bootstrap sample from this combined sample, then partitions it into $n_1, n_2, ..n_r$ sub-samples randomly. Of course, bootstrap test is performed under H_0. Therefore, combining r-samples to create one large sample and re-sampling from this large sample is used in the proposed algorithm.

An algorithm for the construction of bootstrap test and finding p-value as follow:

(1) Let θ_{ij} for $i = 1, .., n_j$, and $j = 1, .., r$ be the angular measurements from $n_1, ..n_r$ samples. Calculate F_t test statistics using the original samples with "aov.circular" function in R.
(2) Draw a bootstrap sample of size $N = n_1 + n_2 + .. + n_r$ from the combined sample of θ_{ij} with replacement. Assign first n_1 observation to first level 1, then n_2 observations to level 2, and the last n_r observations to level r. This way $n_1, n_2, .., n_r$ observations are assigned to $1, 2, .., r$ samples respectively. Calculate the test statistics $F_b^\star$ using these samples.
(3) Repeat the last two steps for $b = 1, ..., B$.
(4) There are now $F_1^\star, .., F_B^\star$ estimated bootstrap test statistics.
(5) Find the number of $F_b^\star \geq F_t$ and then divide the result by B. The result gives $\hat{p} = \#\{F_b^\star \geq F_t\}/B$.
(6) Compare $\hat{p}$ by the level of significance α. If $\hat{p} \leq \alpha$, reject $H_0 : \mu_1 = \mu_2 = ... = \mu_r$. Otherwise, do not reject H_0.

3.2. Uniform Scores Test. A nonparametric test for the equality of two circular distributions is first presented by [18]. A few years later, two-sample case has been extended to k-sample case by [10]. For this reason, k-sample uniform scores test has also been called as Mardia-Watson-Wheeler test in the literature. The null and alternative hypothesis of the test is

$$H_0\text{:All samples come from the same population}$$
$$H_1\text{:At least two are distinct.}$$

Let θ_{ij} (for $i = 1, .., r$ and $j = 1, .., n_i$) show the combined samples of $n_1, n_2, .., n_r$, where each sample consists of angular observations on the circle. The testing procedure assigns ranks to all θ_{ij} and finds a uniform score or circular rank for each θ_{ij} as,

$$d_{ij} = \frac{2\pi(r_{ij})}{N} \text{ for } i = 1, .., r \text{ and } j = 1, .., n_i$$

where r_{ij} is the rank of jth observation from ith sample and $N = n_1+, .., +n_r$. A starting point should be set on the circle in order to find the ranks which can be assigned clock wise or counter clock wise on the circle. In fact, the test is invariant under all rotations as shown by [11], therefore the initial rank could be given to the smallest angle in the data. The test statistics is defined as

$$W = 2\sum_{i=1}^{r}(C_i^2 + S_i^2)/n_i \tag{3.1}$$

where

$$C_i = \sum_{j=1}^{n_i}\cos(d_{ij}) \quad \textit{and} \quad S_i = \sum_{j=1}^{n_i}\sin(d_{ij}). \tag{3.2}$$

are the components of resultant vector for each sample. We should keep in mind that $\sum_{i=1}^{r} C_i = 0$ and $\sum_{i=1}^{r} S_i = 0$, where they could be used to check if the computations are correct in the formulas above. The test statistic, W, has an approximate chi-square (χ^2) distribution with degrees of freedom of 2(r-1) as shown by [10]. Therefore, if $W > \chi^2_{\alpha,2(r-1)}$, H_0 is rejected in favor of H_1. [6] suggests that this test is applicable if $n_i > 10$ for $i = 1, .., r$. Otherwise, a permutation test should be applied.

3.3. Rao's Test of Homogeneity. [13] proposed a test of homogeneity that it is considered large sample alternative of circular ANOVA test. The test is available from R circular package. The requirements to apply Rao's test of homogeneity tests is that the data must be unimodal and the sample size must be sufficiently large.

Let θ_{ij} (for $i = 1, .., r$ and $j = 1, .., n_i$) show the combined samples of $n_1, n_2, .., n_r$. Let X_i and Y_i denote the means of cosine and sine values for ith sample of size n_i such that

$$X_i = \frac{\sum_{j=1}^{n_i} cos\theta_{ij}}{n_i} \text{ and } Y_i = \frac{\sum_{j=1}^{n_i} sin\theta_{ij}}{n_i}$$

and $T_i = \frac{Y_i}{X_i}$ with asymptotic estimated variance of s_i^2 in which the details can be found in [13]. The test statistics, H, is defined as

$$H = \sum_{i=1}^{r} \frac{T_i^2}{s_i^2} - \left(\sum_{i=1}^{r} \frac{T_i^2}{s_i^2}\right)^2 / \left(\sum_{i=1}^{r} \frac{1}{s_i^2})\right)$$

Under H_0 and some general conditions, the test statistics H has a χ^2 distribution with $df = r-1$. For large values of H, the null hypothesis H_0 is rejected which implies different mean directions.

4. Large Sample Example

4.1. Application of Circular ANOVA. The city of Ankara is the capital of Turkey and has a population 4.4 million according to Turkish Institute of Statistics. The city has an elevation of 3077 feet (938 meters) and located at the central part of Turkey. Turkish State Meteorological Services (TSMS) has regional stations that collect and distribute weather related data in Turkey. The literature review did not reveal any studies about the analysis of the seasonal wind directions for the city of Ankara. This study will be the first in this regard. The data provided by TSMS consisted of daily wind directions of Ankara for the year of 2010. First, using the data provided, descriptive summary results were obtained for each seasons (winter, spring, summer and fall). Table 1 shows the descriptive statistics for four seasons. To see the seasonal differences, the data is divided

TABLE 1. Descriptive Statistics for Seasonal Wind Directions in Ankara

Parameters	Winter	Spring	Summer	Fall
Sample Size	90	92	92	91
Mean Direction(degrees)	108.38	140.93	111.48	116.79
Mean Resultant Length	0.6182	0.6458	0.7086	0.6727
Circ. Variance	0.3818	0.3542	0.2914	0.3273
Circ. Std. Deviation	0.9808	0.9464	0.8396	0.8963
Median Direction(degrees)	100.5	140.5	103	107

into four seasons(winter, spring, summer and fall), and rose diagrams(equivalent of histogram) are graphed for each season. Figure 1 shows the seasonal distribution of the wind directions for the year of 2010 in Ankara. In Figure 2, QQ plots of von Mises distribution for each season is shown. It is safe to assume that seasonal wind directions of Ankara (at least for the year of 2010) follow von Mises distribution.

Before performing a circular ANOVA test, we needed to find MLE of κ parameter for all four seasons. The common κ is estimated by $\hat{\kappa} = 1.693012$ with all the samples combined together. If we use [6]'s approach by finding the median of the four seasons, we find that $\hat{\kappa} = 1.754571$. Both results are very much comparable and on the interval $1 < \hat{\kappa} < 2$. See Table 2.

Assumption of the homogeneity of concentration parameters (κ) must be tested in the next step. The circular ANOVA test proposed by [17] assumes that all r concentration parameters are equal to the common concentration parameter κ. So,

$$H_0 : \kappa_1 = \kappa_2 = \kappa_3 = \kappa_4 = \kappa \text{ vs } H_1\text{: At least two are distinct.}$$

This must be tested before starting circular ANOVA method. The following results are obtained from R software using the package called "circular" and using "rao.test" function. The hypothesis test checks the equality of the concentration parameters, the results are from R software (See Table 3).

As we see from the result, the p-value of the test is 0.6171 which is greater than a level of significance of $\alpha = 0.05$ or even 0.10. Therefore, it is safe to assume that all concentration parameters are equal. Since the estimated common concentration, $\hat{\kappa}$, is between 1 and 2, we must use the modified F-test in circular ANOVA according to [6].

TABLE 2. κ parameter estimates for all four seasons. Table also includes common κ estimates which are the last two values

Winter	Spring	Summer	Fall	Common $\hat{\kappa}$	Fisher's $\hat{\kappa}$
1.585405	1.679172	2.024359	1.829970	1.6930	1.7545

TABLE 3. Test of Homogeneity of Kappa Parameter

df	ChiSq	P-value
3	1.79	0.6171

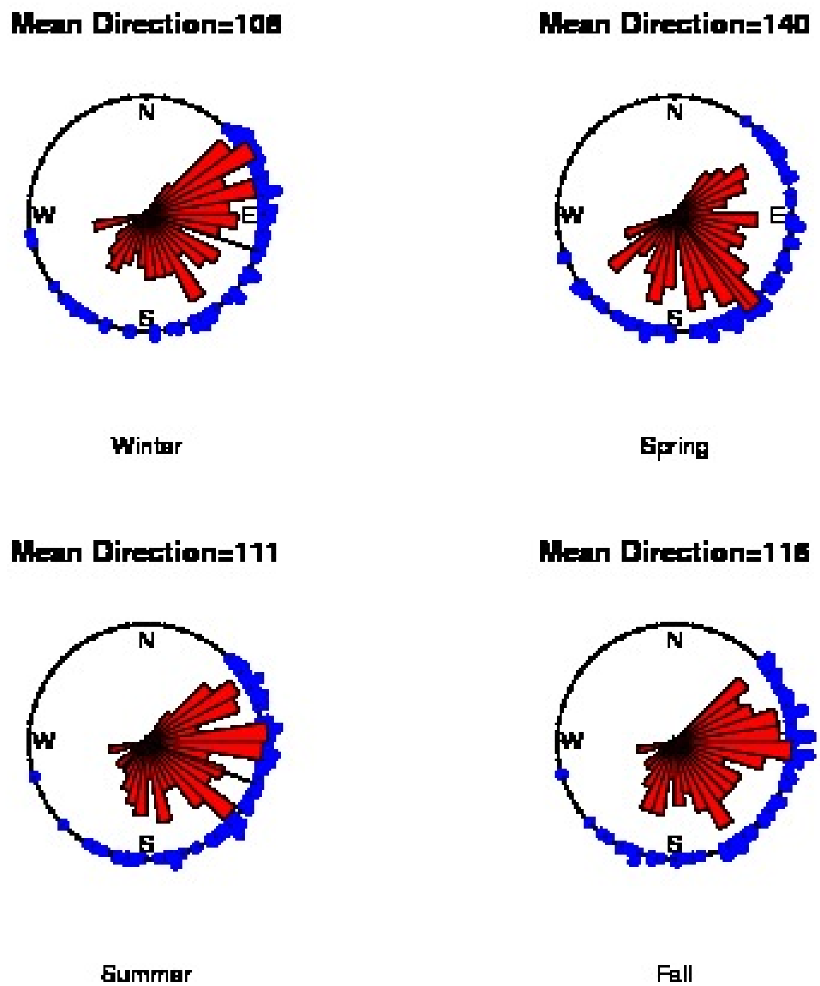


FIGURE 1. Seasonal Rose Diagrams For Ankara's Wind Data

The modified version is proposed by [11] which is based on Stephen's approximation; as suggested by [14].

After confirming the validity of the assumptions before circular ANOVA, we are now ready to run the circular ANOVA test in R. We would like to see if there is a significant difference in the mean wind directions of winter, spring, summer and fall seasons for the city of Ankara. So, we set

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4, \text{ versus } H_1\text{: At least two are distinct.}$$

The circular package in R has aov.circular option that performs circular ANOVA test. The circular ANOVA program in R has two options. First, the analysis can be performed by using F-test if the common kappa parameter (κ) is greater than 1 (if the κ parameter is between 1 and 2, then a modified F test must be performed). The second option

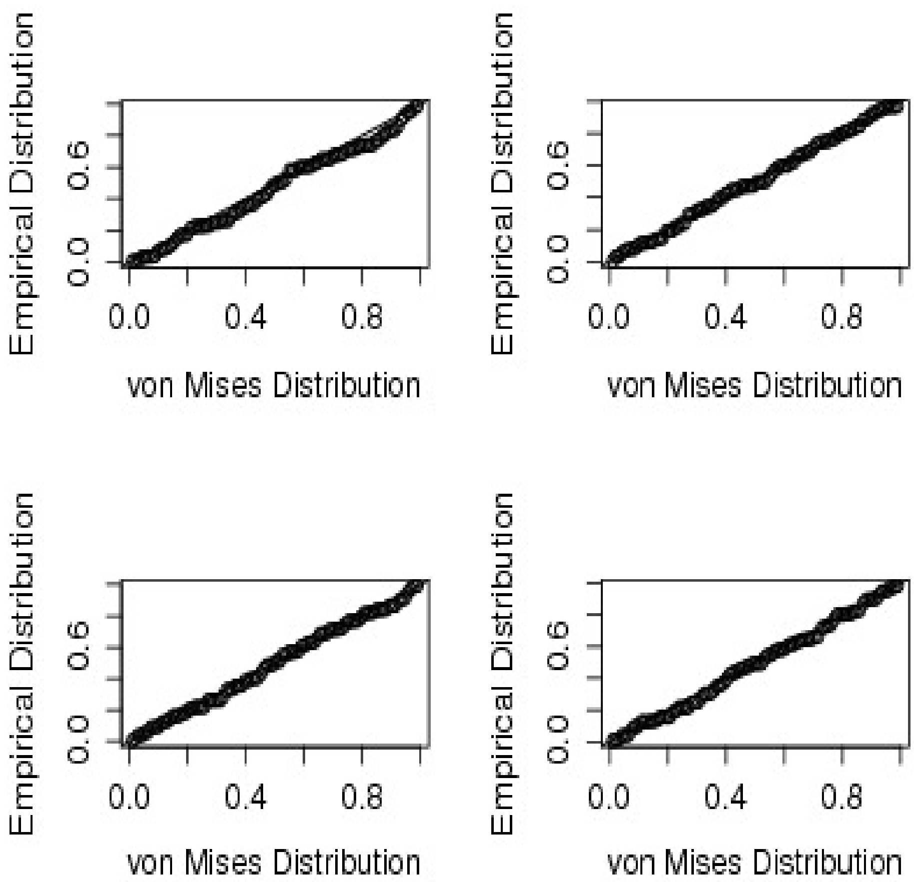


Figure 2. von Mises QQ plots of Wind Directions from Winter, Spring, Summer and Fall

performs Likelihood Ratio Test if the common kappa, κ, parameter is less than 1. Since the estimated common concentration parameter, $\hat{\kappa} = 1.69$, a modified F test is used in circular ANOVA. The result of the circular ANOVA is shown below in Table 4. Table 4

Table 4. Test of Circular ANOVA using R

Source	df	SumSquare	MeanSquare	F	Pvalue
Between	3	5.5446	1.8487	6.516	0.000266
Within	356	123.371	0.3465		
Total	359	128.917	0.3591		

implies that H_0 is rejected and, therefore, there is a significant difference among the seasonal winds directions of Ankara since the p-value of the test is 0.000266. This means that there was a seasonal difference among four seasons for the year of 2010. Visual analysis of Table 1 and Figure 1 indicates that the mean wind direction of spring season is 140^0 and looks significantly different than the other three seasons. In the next step, we will perform the circular ANOVA again without the spring season data in order to see the effect of the spring season on the analysis. The results can be seen below in Table 5. It appears that there is no significant difference among three seasons (winter, summer,

Table 5. Test of Circular ANOVA without Spring data

Source	df	SumSquare	MeanSquare	F	Pvalue
Between	2	0.3238	0.1619	0.575	0.5634
Within	267	90.8774	0.3404		
Total	269	91.2012	0.3390		

and fall) since the p-value is 0.5634. This means that the spring season has significantly different mean wind direction for the months of March, April and May. Figure 1 shows the circular plots and rose diagrams for each season, and the mean direction for spring is significantly different at $\alpha = 0.05$. The result of the circular ANOVA and also bootstrap approach could lead to new studies related to seasonal wind directions in different parts of Turkey.

4.2. Application of Nonparametric Tests. Nonparametric tests did not need prior investigation of the circular data in order to check assumptions as in the case of circular ANOVA. So, we implemented bootstrap, uniform scores test and Rao's test of homogeneity in R using circular package. Bootstrap and uniform scores test are not available in R's circular package. Therefore, a function has been written in R for those two tests. Rao's test of homogeneity is called rao.test in R via circular package. Rao's test of homogeneity gives p-value of 0.0214 for the test of H_0 which assumes all seasonal mean directions are equal. So, Rao's test implies that there is a significant difference in the seasonal wind directions of Ankara. Similar to the circular ANOVA, spring wind directions are excluded and Rao's test is applied again using winter, summer and fall data. The result shows that Rao's test gives a p-value of 0.6216 which implies no significant difference in the remaining seasons. When we run the uniform scores test on Ankara's seasonal wind data, it gives a p-value of 0.0014 which implies significant difference among the seasonal wind directions. If we repeat the test without spring season, then uniform scores test gives a p-value of 0.64 which is not significant or no difference in the mean wind directions. Bootstrap test finds a p-value of 0.0005 which is very significant and implies a difference in the seasonal mean wind directions of Ankara. If we remove the spring season from the data and run the bootstrap test again, we obtain a p-value of 0.6055. Therefore, we note that circular ANOVA and alternative nonparametric tests confirm each other and reach the same decision for Ankara's seasonal wind data.

5. Small Sample Example

5.1. Application of Circular ANOVA. Circular ANOVA and nonparametric alternatives are demonstrated under a small sample example (all samples are less than 25). The example consists of seasonal wind directions of Gorleston, England from [11]. The data have winter, spring, summer and fall wind directions, which are collected between 11:00 and 12:00AM on Sundays in 1968. Descriptive Statistics for the data shown below in Table 7. The main focus is again "is there any significant seasonal difference in the wind directions?". For this purpose, we again set $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ vs H_1: At least two are distinct. [11] also investigated this example and assumed that the concentration parameters of the seasonal winds are equal. [11] estimated the concentration parameter

Table 6. Descriptive Statistics for Seasonal Wind Directions in Gorleston, England

Parameters	Winter	Spring	Summer	Fall
Sample Size	12	12	13	12
Mean Direction(degrees)	272	330	57	232
Mean Resultant Length	0.4265	0.1776	0.2975	0.2656
Circ. Variance	0.5735	0.8224	0.7025	0.7344
Circ. Std. Deviation	1.3054	1.8589	1.5570	1.6282
Median Direction(degrees)	288	360	30	255.6

Table 7. Likelihood Ratio Test of Homogeneity of Seasonal Wind Directions in Gorleston, England

df	ChiSq	Pvalue
3	3.459	0.3261

from the combined samples and found it as $\hat{\kappa} = 0.24$. Moreover, it is true that all $\hat{\kappa_i} < 1$ for $i = 1, .., r$. Therefore, [11] suggests Likelihood Ratio Test (LRT) type test statistics for this problem because of too small (less than 1) concentration parameter estimate. See example 6.11 on page 165 of Mardia ([11]). Using "aov.circular" (with LRT option) in R, we find the following results: The chi-square critical value for df=3, and $\alpha = 0.05$ is 7.81 from a chi-square table. The p-value of the test is 0.3261. Thus, the result from LRT test option concludes that the seasonal wind directions are not significantly different at $\alpha = 0.05$.

5.2. Application of Nonparametric Tests. Nonparametric tests from Section 3 is executed in R to get the significance probability of the tests (p-values). In fact, [6] made a remark that the summer seasonal directions for Gorleston data appear to be different that the rest of the data and excluded it from his application of Gorleston data. Similarly, [11] used the same data set to run the uniform scores test(Mardia-Watson-Wheeler test) to investigate the homogeneity of population distributions and found that uniform score test rejects H_0 with a p-value of 0.0409. So, uniform scores test finds significant difference among seasonal wind directions. On the other hand, Rao's test of homogeneity finds a p-value of 0.9095 which does not reject H_0 that claims all mean directions are equal. One explanation of this difference in Rao's test is that it requires large samples in order to reach the nominal type-I error rate as seen in Section 6. So, as indicated by [6] and [11], the uniform score test was able to identify the significance of seasonal wind directions for Gorleston, England. Finally, bootstrap test obtains a p-value of 0.2045 for H_0 and it implies no significance difference among the seasonal wind directions.

6. Power Study

Performance of nonparametric tests are compared against the circular ANOVA by a power simulation study. Three different distribution models are considered: von Mises (ideal case for Circular ANOVA test), wrapped Cauchy and mixed von Mises with rate of mixtures of 90% and 70%, respectively. Mixed von Mises is analogues the contaminated normal distribution which is commonly used in traditional statistics to investigate data models with contaminations or outliers. We assumed that there are four random samples (for example, wind directions in four seasons) and the equality of the mean directions of four populations is the null hypothesis. So, we consider the following alternative hypothesis in order to compute the power of circular ANOVA against the nonparametric test:

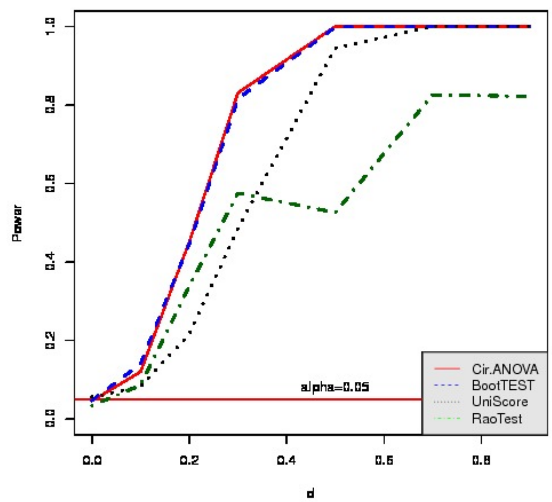
$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$
$$H_1 : \mu_1 + d = \mu_2,\ \mu_1 + 2d = \mu_3,\ \mu_1 + 3d = \mu_4$$

where d is a constant (shift value) that controls the alternative hypothesis. If d=0, then $H_0 = H_1$ and the tests compared in this study should reach nominal value of type-I error rate (which is set to $\alpha = 0.05$). First, Monte Carlo simulation is performed (B=1000 replications) by generating four independent random samples ($n_1 = n_2 = n_3 = n_4 = 25$) from von Mises distribution with parameters $\mu = \pi$ and $\kappa = 2$. Monte Carlo simulation finds the number of times the tests rejects H_0 under the assumption that H_1 is true for each

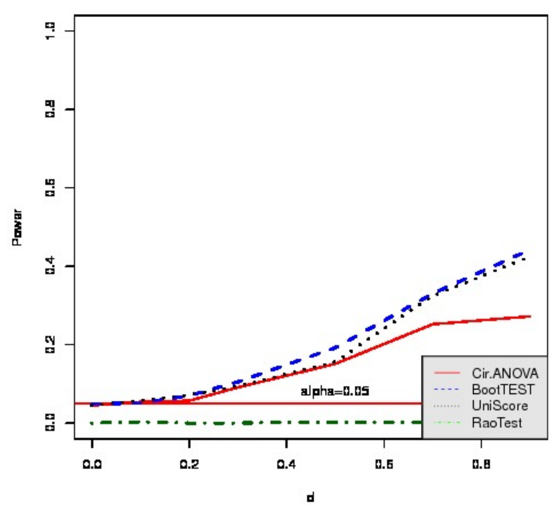
d=(0,0.1,0.2,0.3,0.5,0.7,0.9)

Then, the result is divided by B (number of replications) to find an estimate of the power. The result can be converted to the percentage that gives the empirical power of the test. Figure 3(a) shows the power curve for circular ANOVA, Bootstrap test, Rao's test of homogeneity and uniform scores test under H_1 and $\kappa = 2$ for each d.

Figure 3(a) and Table 8 show that when d=0, circular ANOVA, bootstrap and uniform scores tests have comparable estimated type-I error rates which are close to the nominal value of 0.05. On the other hand, Rao's test did not reach the nominal value of type-I error. Moreover, circular ANOVA is known to be powerful according to [6] when κ has 2 or higher and the data come from von Mises distribution. Bootstrap and Uniform score tests also worked as good as circular ANOVA under the data model and parameter assumptions. For larger shifts in the mean directions of the populations (for larger d values), uniform score test and Rao's tests started to lose some power as shown by Figure 3(a). In the next simulation, we assumed that all four samples are coming from von Mises populations and the common concentration parameter of $\kappa = 0.5$.



(a)



(b)

FIGURE 3. Circular ANOVA and nonparametric test alternatives are compared in terms of their power curves. All four samples are generated from von Mises with $\kappa = 2$ (figure a) and $\kappa = 0.5$ (figure b) parameters for each d.

Figure 3(b) shows that, when $\kappa = 0.5$, bootstrap test performed the best among the compared methods. Bootstrap test has an estimated type-I error rate of 0.049 which is very close to the nominal value of $\alpha = 0.05$. Uniform score and circular ANOVA tests are comparable at d=0 but circular ANOVA loses power at the larger shift values under H_1. As pointed out by [6], circular ANOVA requires κ parameter to be larger than 2 in order to maintain type-I error rate and its power. On the other hand, Rao's test of homogeneity did not perform well against the other three methods and did not reach the desired level of α or power. One reason could be that Rao's test requires large sample sizes to reach nominal value of type-I error. Table 8 has the numerical values of the simulations for $\kappa = 2$ and $\kappa = 0.5$ assumptions.

Table 8. Power simulation results for Circular ANOVA, Rao's test, Uniform Score test. All four samples are from von Mises with $\kappa = 2$ (left table) and $\kappa = 0.5$ (right table) parameters

d	CirANOVA	Boot	Uniform	Rao	CirANOVA	Boot	Uniform	Rao
d=0	0.046	0.047	0.056	0.034	0.045	0.049	0.045	0.000
d=0.1	0.120	0.140	0.083	0.086	0.054	0.051	0.057	0.003
d=0.2	0.451	0.447	0.216	0.338	0.057	0.070	0.072	0.000
d=0.3	0.831	0.816	0.486	0.575	0.091	0.105	0.093	0.000
d=0.5	0.999	0.999	0.945	0.526	0.151	0.192	0.156	0.002
d=0.7	1.000	1.000	1.000	0.826	0.252	0.330	0.324	0.002
d=0.9	1.000	1.000	1.000	0.822	0.272	0.442	0.425	0.003

In the next simulation, we considered small and large sample simulations to compare the performance of all four tests under wrapped Cauchy distribution assumption. First, four random samples of size 10 generated from wrapped Cauchy distribution with $\mu = \pi + d$ and $\rho = 0.9$ parameters. The reason that we considered the wrapped Cauchy distribution is to see the performance of circular ANOVA and alternative tests when the data come from non-von Mises models and also compare the tests under a small sample case. We repeated the same experiment for a large sample size ($n_1 = n_2 = n_3 = n_4 = 100$) using the same wrapped Cauchy distribution and parameters.

Figure 4(a) shows uniform score test performed better than bootstrap and circular ANOVA tests under a small sample case and wrapped Cauchy assumption. At d=0, uniform score test estimates the nominal type-I error rate with 0.049 which almost equals to the true rate of $\alpha = 0.05$. On the other hand, bootstrap, circular ANOVA and Rao's test did not maintain the nominal type-I error rate of $\alpha = 0.05$. Overall, Rao's test of homogeneity did not perform well again due to small sample sizes. Figure 4(b) shows the power curves under the large sample case where the random samples of size 100 created from the wrapped Cauchy distribution with $\mu = \pi + d$ and $\rho = 0.9$. Figure 4(b) shows all methods except circular ANOVA have maintained the nominal rate of type-I error as seen in Table 9. Rao's test homogeneity has an estimated type-I error rate of 0.044 for $\alpha = 0.05$ and it has shown its best performance when large samples sizes are considered. So, circular ANOVA did not perform very well under the assumption of wrapped Cauchy populations.

TABLE 9. Power simulation results for circular ANOVA, Rao's test, uniform score test. Four random samples of size 10 are from wrapped Cauchy with $\mu = \pi$ and $\rho = 0.9$ parameters (right table) and large sample case where $n_1 = n_2 = n_3 = n_4 = 100$ are again generated from wrapped Cauchy distribution with the same parameters (left table).

d	CirANOVA	Boot	Uniform	Rao	CirANOVA	Boot	Uniform	Rao
d=0	0.008	0.027	0.049	0.026	0.004	0.051	0.048	0.044
d=0.1	0.148	0.288	0.774	0.468	0.998	1.000	1.000	1.000
d=0.2	0.678	0.826	0.998	0.880	1.000	1.000	1.000	1.000
d=0.3	0.938	0.980	1.000	0.930	1.000	1.000	1.000	1.000
d=0.5	1.000	1.000	1.000	0.972	1.000	1.000	1.000	1.000
d=0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
d=0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

In the next simulation, four independent random samples ($n_1 = n_2 = n_3 = n_4 = 25$) are generated from "mixed" von Mises distribution with proportion of the mixture is defined as $p * VonM(\mu_1 = \pi/2 + d, \kappa_1 = 3) + (1 - p)VonM(\mu_2 = 2\pi, \kappa_2 = 0.5)$ where p shows the proportion of the mixture. We will consider p=0.90 (90%-10% mixture) and p=0.70 (70%-30% mixture) proportions respectively. These model assumptions can also be considered an equivalent of contaminated normal distribution in the traditional sense. The goal is to see the performance of circular ANOVA and nonparametric tests under these assumptions that random samples come from mixture of von Mises distributions. This approach is clearly a violation of the assumption for circular ANOVA since the test requires all r populations should come from von Mises distributions with equal parameters. For each d, Monte Carlo simulation is performed and estimated power curve of each method is presented in Figure 5.

Simulation results are also shown by Table 10 below. As it can be seen from Figure 5(a) and also from Table 10, uniform scores test performed the best overall when $p = 0.90$ (90%-10% mixture). At d=0, the nominal type-I error rate (α) should be reached if a test works as expected but only uniform scores comes close to the nominal value of $\alpha = 0.05$ with estimates of 0.044. Circular ANOVA and Rao's test estimates for $\alpha = 0.05$ were 0.028 and 0.015, respectively. It could be an indication that these two tests are very conservative in rejecting H_0. Bootstrap test is also under performing since its estimated type-I error rate is 0.034 but it is slightly better than circular ANOVA and Rao's test. If we assume $p = 0.70$ ($70\% - 30\%$ mixture of von Mises distributions) and generate four random samples from this mixed von Mises distribution, simulation results show uniform scores tests have an estimate of 0.047 for $\alpha = 0.05$. It is considerably close to the nominal value of type-I error rate and indication that the test works as expected even if the data come from mixture of von Mises distribution. On the other hand, circular ANOVA and bootstrap have estimates of 0.033 and 0.039 which are much smaller then the nominal value of $\alpha = 0.05$. Again, circular ANOVA and bootstrap test look very conservative when we assume mixture of von Mises distributions with $p = 0.70$. Similarly, Rao's test did not perform well for the mixture of von Mises distributions when $p = 0.70$. Thus, uniform scores tests should be considered a better performer under contaminations and violation of having von Mises distribution assumption.

Table 10. Power simulation results for circular ANOVA, bootstrap test, Rao's test of homogeneity, and uniform score test from the mixture of von Mises populations with proportion of the mixture is 90% (right table) and 70% (left table) respectively.

d	CirANOVA	Boot	Uniform	Rao
d=0	0.028	0.034	0.044	0.015
d=0.1	0.185	0.192	0.133	0.021
d=0.2	0.635	0.6466	0.458	0.044
d=0.3	0.941	0.947	0.829	0.438
d=0.5	0.998	0.998	0.989	0.902
d=0.7	1.000	1.000	0.998	0.993
d=0.9	1.000	1.000	1.000	0.999

CirANOVA	Boot	Uniform	Rao
0.033	0.039	0.047	0.003
0.090	0.089	0.103	0.028
0.332	0.341	0.334	0.007
0.725	0.729	0.720	0.061
0.942	0.945	0.935	0.307
0.991	0.992	0.990	0.645
1.000	0.999	1.000	0.803

7. Conclusion

The main motivation of this paper was about investigating circular ANOVA (one way analysis of variance in circular data analysis) against nonparametric alternatives such as bootstrap test, uniform scores test (Mardia-Watson-Wheeler test) and Rao's test of homogeneity in the analysis of multi-sample circular data problems. Circular ANOVA requires certain assumptions as we discussed in Section 2. On the other hand, bootstrap, uniform scores, and Rao's tests are considered nonparametric tests, and they do not depend on any population distributions (see Section 3) or equality of parameters. There is also a lack of study in the literature about the comparison of circular ANOVA with alternative methods if the assumptions of circular ANOVA do not meet. So, real life examples and power analysis are performed on circular ANOVA, bootstrap, uniform scores test and Rao's test of homogeneity to observe their comparative performance under von Mises, mixed von Mises and wrapped Cauchy distribution assumptions.

Section 6 presents power simulation study which is performed to see the performance of nonparametric tests against circular ANOVA under von Mises distribution. As seen in Figure 3(a) that it is an ideal case for circular ANOVA since the test gives its best performance if $\kappa = 2$ or higher but circular ANOVA starts under performing compare to the uniform score test if $\kappa < 1$ as shown by Figure 3(b) and Table 8. Moreover, Figure 4 shows power curves of all four tests under a small and large sample cases. As we see in Figure 4(a) that uniform score test performs better than bootstrap and circular ANOVA when sample sizes are small and come from wrapped Cauchy populations. Rao's test can not compete with them if the sample sizes are too small. Next, we considered a large sample case where all four random samples have a size of 100 and the results are presented by Figure 4(b) and Table 9. As we see that all four tests have converged power curves but only bootstrap and uniform score tests have maintained the nominal type-I error rate of 0.05 which is an indication that under a large sample case bootstrap and uniform score test works as expected and detect shifts in the mean directions better than circular ANOVA. Figure 5 and Table 10 are obtained by generating four random samples (sizes of 25) from mixed von Mises with ($\mu_1 = \pi/2$, $\kappa_1 = 3$) and ($\mu_2 = 2\pi, \kappa_2 = 0.5$) with a mixture rate of $p = 0.90$ and 0.70 respectively. Figure 5(a) (also Table 10) shows that only uniform scores test is almost equal to the nominal type-I error rate of 0.05. Therefore, uniform scores test could be used without sacrificing the power of the test compare to the circular ANOVA, bootstrap and Rao's test under the mixture of von Mises distributions with $p = 0.90$. Figure 5(b) also shows uniform scores test is almost equal to the nominal type-I error rate when we assume mixed von Mises with a mixture rate of $p = 0.70$. In both cases of mixed von Mises distributions, circular ANOVA and bootstrap tests are less likely to reject H_0 when it is false since their estimates of nominal type-I rate are much smaller than $\alpha = 0.05$. Similarly, Rao's test is also under performing when we assume mixture of von Mises distributions.

We can conclude that circular ANOVA shows superiority if the data come from von Mises distribution with a common concentration parameter of $\kappa = 2$ or higher which is considered an ideal case for circular ANOVA. If $\kappa < 1$, bootstrap and uniform scores tests performs slightly better overall. If we assume mixed von Mises and wrapped Cauchy distributions, uniform scores tests performs better than circular ANOVA, bootstrap and Rao's test of homogeneity in which Rao's test requires large sample sizes in order to reach the performance of the alternative tests.

8. Acknowledgement

We greatly appreciate the comments and suggestions made by the referees and the editor of the journal in preparation and publication of this manuscript.

References

[1] Agostinelli, C. and Ulric, L. (2009). Circular Package in R, http://cran.r-project.org/web/packages/circular/circular.pdf.

[2] Batschelet, E. (1981). *Circular Statistics in Biology*, Academic Press, London.

[3] Efron, B. (1979). Bootstrap methods: Another look at the Jackknife, *Ann. Statist.*, 7,pp 1-26.

[4] Efron, B. and Tibshirani, R.J. (1993).*An introduction to the Bootstrap*, Chapman and Hall, New York.

[5] Fisher, N.I. and Hall, P. (1989). Bootstrap Confidence Regions for directional data. *Journal of the American Statistical Association*, 84, 408, pp. 996-1002.

[6] Fisher, N. I. (1993). *Statistical Analysis of Circular Data*, Cambridge University Press, New York, USA.

[7] Hall, P. (1988). On the Bootstrap and Symmetric Confidence Intervals. *Journal of the Royal Statistical Socieity. Ser. B*, 50, pp. 35-45.

[8] Harrison, D., Kanji, G.K. and Gadsden, R.J.(1986). Analysis of variance for circular data, *Journal of Applied Statistics*, vol 13,issue 2.

[9] Jammalamadaka, S. R. and Sen Gupta, A. (2001). *Topics in Circular Statistics*, World Scientific Publishing Co. Pte. Ltd, London, England.

[10] Mardia, K. V. (1967). A nonparametric test for the bivariate two-sample location problem,*Jour. of Roy. Statis. Soc. Ser. B*, B29, 320-342.

[11] Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press, London and New York.

[12] Okamura, H. and Takasuka, A. (2012). A bootstrap method for testing equality of peak months. *Population Ecology* 54,1, 169-176.

[13] Rao, J.S. (1967). Large sample tests for the homogeneity of angular data, *Sankhya*, Ser, B., 28.

[14] Stephens, M. (1969). Tests for the von Mises Distribution, *Biometrika*, 56,149-160.

[15] Tasdan, F. (2013). Technical Report: R programs for Circular ANOVA and Nonparametric Alternatives, http://www.wiu.edu/users/ft100/WindRcode.pdf

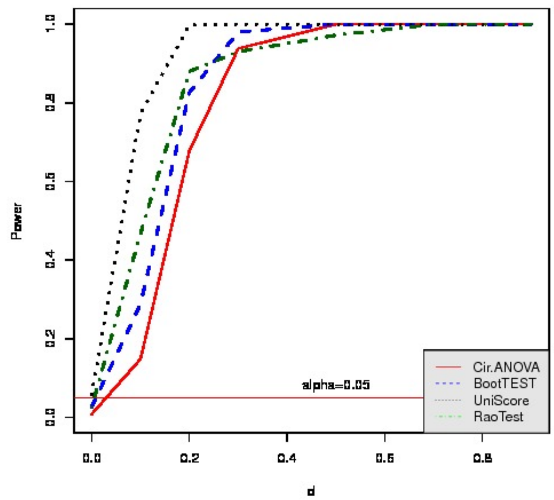
[16] Watson, G. S. (1983). *Statistics on Spheres*, Wiley, New York, USA.

[17] Watson, G. S. and Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere, *Biometrika*, 43,344-352.

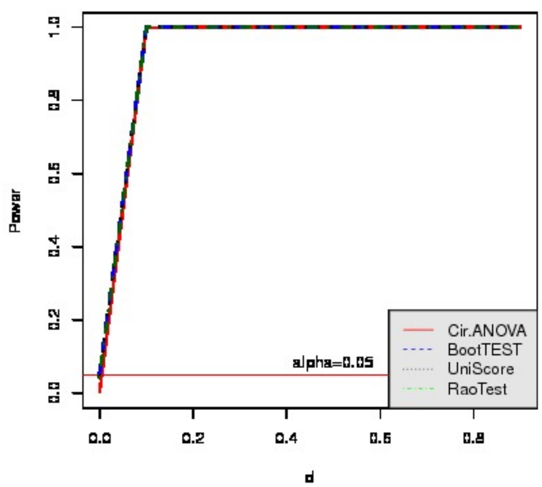
[18] Wheeler, S. and Watson, G. S. (1964). A distribution free two sample test on the circle, *Biometrika* 51, 256-7.

Appendix A

R functions that are used in this paper can be found in Tasdan ([15]). These functions require "circular" package to be installed first in order to run the functions.

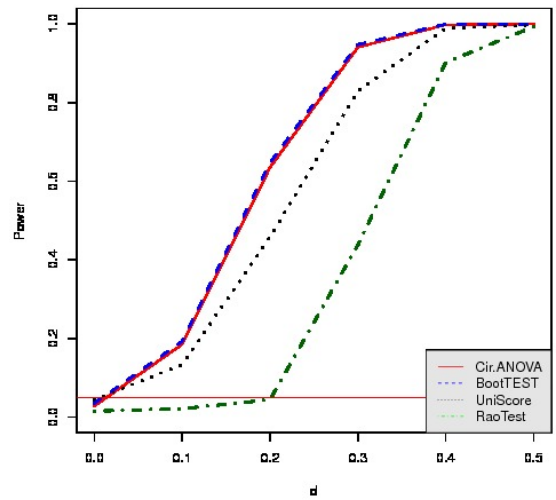


(a)

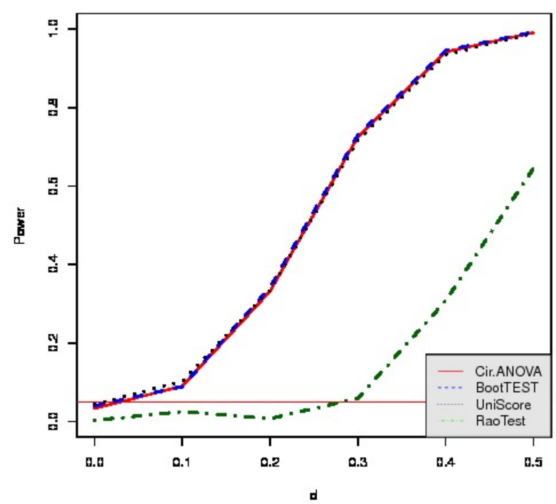


(b)

FIGURE 4. Circular ANOVA and nonparametric alternatives are compared in terms of their power curves. Figure (a) shows all four samples of size 10 (small sample case) are generated from wrapped Cauchy distribution with $\mu = \pi$ and $\rho = 0.9$ parameters and figure (b) shows large sample case where $n_1 = n_2 = n_3 = n_4 = 100$.



(a)



(b)

FIGURE 5. Circular ANOVA and nonparametric tests are compared in terms of their power curves. Figure (a) shows all four samples are generated from mixed von Mises with $\mu_1 = \pi/2 + d$, $\kappa_1 = 3$ and $\mu_2 = 2\pi$, $\kappa_2 = 0.5$ with proportion of the mixture is 90% and figure (b) shows the repeat of the simulation with proportion of 70% mixture.

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 117 – 130

SIMPLE COMPUTATIONAL FORMULAS FOR INCLUSION PROBABILITIES IN RANKED SET SAMPLING

Fikri Gökpınar * † and Yaprak Arzu Özdemir * ‡

Received 21 : 10 : 2011 : Accepted 24 : 01 : 2013

Abstract

In this study, we derived new formulations for the first and second order inclusion probabilities of a ranked set sample in a finite population setting. Gökpınar and Özdemir (2010) developed a formula to calculate the first order inclusion probabilities. However, the formula given in this study is much easier than the one given by Gökpınar and Özdemir (2010). Second order inclusion probabilities are computed based on the formulas which are used for the calculation of first order inclusion probabilities. Also, we give a numerical example to show the calculation of the formulas and Matlab codes which give first and second inclusion probabilities for any set and population sizes.

Keywords: Ranked Set Sampling, First Order Inclusion Probability, Second Order Inclusion Probability, Finite Population Setting.

2000 AMS Classification: 62D05, 65C60

1. Introduction

Ranked Set Sampling (RSS) is an efficient sampling technique than the simple random sampling (SRS) for improving the accuracy of the estimation of means. RSS was first introduced by McIntyre (1952) for estimate the mean of pasture yields. In recent years, RSS is used in many fields such as the environment, ecology and agriculture. Some applications in these fields can be found in the studies of Johnson et.al. (1993) and Al-Saleh et al(2000). Also, some recent ideas about RSS can be found in Bouza(2005).

In RSS, the inclusion probabilities of the population units are different from each other, and it is difficult to determine the inclusion probabilities for all sample sizes. Al-Saleh and Samawi (2007) obtained the inclusion probabilities in RSS for the set size 2 and 3. Özdemir and Gökpınar (2007) obtained the inclusion probabilities in RSS for all set sizes when the cycle size is one, and Özdemir and Gökpınar (2008) have adapted

*Faculty of Sciences, Department of Statistics Gazi University Ankara Turkey 06500
†Email: (F. Gökpınar) fikri@gazi.edu.tr
‡Email: (Y. A. Özdemir) yaprak@gazi.edu.tr

this procedure to Median Ranked Set Sampling (MRSS) with any set and cycle sizes. Gökpınar and Özdemir (2010) generalized the formula of inclusion probabilities in RSS for all cycle and set sizes.

Jafari et. al. (2010) derived the first and second order inclusion probabilities for Level 0 RSS procedure (sampling with replacement) of Deshpande et. al. (2006) and developed several designs based estimators of the population mean. Recently, Gökpınar and Özdemir (2011) defined the Horvitz-Thompson (HT) estimator of the population mean using the inclusion probabilities of a ranked set sample in a finite population setting. Furthermore, they give a calculation formula of the second order inclusion probabilities which is required to calculate the variance of the HT estimator.

In this study, we give a simple formula to calculate the first and second order inclusion probabilities in RSS. In the second section of this study, we give the selection procedure, required definitions, and the formulas of these inclusion probabilities in RSS. In the third section, a numerical example is given to show the calculation of the formula. Concluding remarks are given in section 4. Also in the appendix, we give Matlab codes to calculate the first and second inclusion probabilities for any set and population sizes.

2. Inclusion Probabilities in RSS

Let the population units be $X_1<X_2<...<X_N$ and let a ranked set sample from this population be $Y_1, Y_2,..., Y_m$ based on the level 1 sampling procedure. Level 1 sampling procedure is given as follows (Deshphande et al. 2006, Al-Saleh and Samawi, 2007):

In the g^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.

2. The sampled units are ranked with respect to the variable of interest and the g^{th} order statistic is selected for measurement.

3. All other m-1 units are returned to the population.

4. The steps 1-3 are repeated for $g=1,2,...,m$ to obtain a ranked set sample of size m.

The entire cycle may be repeated, if necessary, r times to produce a ranked set sample of size mr=n. In this study, we only considered the case of r=1. A generalization for r>1 can be easily derived.

To calculate the first and second order inclusion probabilities, some basic definitions are required.

A_i is the event of selecting the ith population unit in the sample (i=1,2,...,N).

A_j is the event of selecting the jth population unit in the sample (j=1,2,...,N).

$$l_g(i,j) = \begin{cases} 1 & t < i \\ 2 & t > j \\ 3 & i < t < j \end{cases}$$

where i<j and t is the rank of the population unit which is selected in the gth selection. If i=j, then $l_g(i,i) = l_g(i)$ can be defined as;

$$l_g(i) = \begin{cases} 1 & t < i \\ 2 & t > i \end{cases}$$

$B_g^1(i,j)$is the event of selecting smaller population unit than the ith population unit in the gth selection ($l_g(i,j)$=1). If i=j, then $B_g^1(i,i) = B_g^1(i)$.

$B_g^2(i,j)$is the event of selecting greater population unit than the j^{th} population unit in the g^{th} selection ($l_g(i,j)$=2). If i=j, then $B_g^2(i,i) = B_g^2(i)$.

$B_g^3(i,j)$is the event of selecting greater population unit than i^{th} and smaller population unit than the jth population unit in the gth selection ($l_g(i,j)$=3).

$a_g(i)$ is the number of smaller population units than the i^{th} population unit selected before the g^{th} selection.

$a_g(j)$is the number of smaller population units than the j^{th} population unit selected before the g^{th} selection.

So there is a relationship between $a_g(i)$ and $\{l_1(i), l_2(i), ..., l_{g-1}(i)\}$as given below

$$a_g(i) = 2(g-1) - \sum_{u=1}^{g-1} l_u(i)$$

By using these definitions, the probability of selecting the i^{th} population unit in the sample can be obtained as

$$\pi_N(A_i) = 1 - \pi_N(A_i^c) \qquad i = 1, 2, \ldots, N \tag{2.1}$$

where

$$\pi_N(A_i^c) = \sum_{l_1(i), l_2(i), \ldots, l_m(i)=1}^{2} P\left(B_1^{l_1(i)}(i) \cap B_2^{l_2(i)}(i) \cap ... \cap B_m^{l_m(i)}(i)\right)$$

$$= \sum_{l_1(i), l_2(i), \ldots, l_m(i)=1}^{2} \prod_{g=1}^{m} P_{a_g(i)}\left(B_g^{l_g(i)}(i) | B_1^{l_1(i)}(i) \cap B_2^{l_2(i)}(i) \cap ... \cap B_{g-1}^{l_{g-1}(i)}(i)\right) \tag{2.2}$$

We derive $P_{a_g(i)}\left(B_g^{l_g(i)}(i) | B_1^{l_1(i)}(i) \cap B_2^{l_2(i)}(i) \cap ... \cap B_{g-1}^{l_{g-1}(i)}(i)\right)$ in the following theorems.

2.1. Theorem. *The probability,* $P_{a_g(i)}\left(B_g^{1}(i) | B_1^{l_1(i)}(i) \cap B_2^{l_2(i)}(i) \cap ... \cap B_{g-1}^{l_{g-1}(i)}(i)\right)$ *in Eq. (2.2), can be written as follows when* $a_g(i)$*=0;*

$$P_0\left(B_g^1(i) | B_1^2(i) \cap B_2^2(i) ... \cap B_{g-1}^2(i)\right) \tag{2.3}$$

$$= \begin{cases} 0 & i = 1, 2, ..., g \\ \sum_{u=g}^{m} \dfrac{\binom{i-1}{u}\binom{N-i-g+2}{m-u}}{\binom{N-g+1}{m}} & i = g+1, ...N-m+1 \\ 1 & i = N-m+2, ..., N. \end{cases}$$

Proof. $P_0\left(B_g^1(i) | B_1^2(i) \cap B_2^2(i) \cap ... \cap B_{g-1}^2(i)\right)$ means that the probability of selection of a smaller unit than the i-th population unit in the g-th selection under the condition that there is no a smaller population unit selected before the g-th selection. So, there are i-1 smaller population units and N-i+1-(g-1) =N-i-g+2 greater population units from the i-th population unit in the g-th selection. Also, we should choose at least g population units smaller than i-th population unit to choose a population unit smaller than the i-th population unit. So, smaller population units than any of the first g population units ($i = 1, 2, ..., g$) have no chance to be selected in the g-th selection. On the other hand, greater population units than any of the last m-1 population units($i = N-m+2, ..., N$) have no chance to be selected in the g-th selection. Therefore, smaller population units than any of the last m-1 population units($i = N-m+2, ..., N$) have a %100 probability to be selected in the g-th selection. So,

$$\begin{aligned}
&P_0\left(B_g^1(i)|B_1^2(i)\cap B_2^2(i)...\cap B_{g-1}^2(i)\right)\\
&=\frac{\begin{pmatrix} i-1\\ g\end{pmatrix}\begin{pmatrix} N-i-g+2\\ m-g\end{pmatrix}}{\begin{pmatrix} N-g+1\\ m\end{pmatrix}}+...+\frac{\begin{pmatrix} i-1\\ m\end{pmatrix}\begin{pmatrix} N-i-g+2\\ 0\end{pmatrix}}{\begin{pmatrix} N-g+1\\ m\end{pmatrix}}\\
&=\sum_{u=g}^{m}\frac{\begin{pmatrix} i-1\\ u\end{pmatrix}\begin{pmatrix} N-i-g+2\\ m-u\end{pmatrix}}{\begin{pmatrix} N-g+1\\ m\end{pmatrix}},\quad i=g+1,....,N-m+1.
\end{aligned}$$

This completes the proof. □

The other probabilities required to calculate the inclusion probabilities can be obtained by using Theorem 2.1. The selection probability of the population unit smaller than $i'=i+a_g(i')$ $(a_g(i')=1,2,...,g-1)$ in the g-th selection when $a_g(i')>0$, is equal to the selection probability of the population unit smaller than the i-th population unit in the g-th selection when $a_g(i)=0$. This probability is stated at Theorem 2.2.

2.2. Theorem. $P_{a_g(i')}\left(B_g^1(i')|B_1^{l_1(i')}(i')\cap B_2^{l_2(i')}(i')...\cap B_{g-1}^{l_{g-1}(i')}(i')\right)$ *can be written as follows when* $i'=i+a_g(i')$ $(a_g(i')=1,2,...,g-1)$.

$$\begin{aligned}
&P_{a_g(i')}\left(B_g^1(i')|B_1^{l_1(i')}(i')\cap ...\cap B_{g-1}^{l_{g-1}(i')}(i')\right)=\\
&P_{a_g(i)=0}\left(B_g^1(i)|B_1^2(i)\cap ...\cap B_{g-1}^2(i)\right). \qquad (2.4)
\end{aligned}$$

Proof. In the g-th selection, the number of population units smaller than i' are

$$i'-a_g(i')-1=i+a_g(i')-a_g(i')-1=i-1.$$

By the same way, the number of population units equal or greater than i' are

$$N-i'+1-(g-1-a_g(i'))=N-(i+a_g(i'))+1-(g-1-a_g(i'))=N-i-g+2.$$

So,

$$\begin{aligned}
&P_{a_g(i')}\left(B_g^1(i')|B_1^{l_1(i')}(i')\cap B_2^{l_2(i')}(i')\cap ...\cap B_{g-1}^{l_{g-1}(i')}(i')\right)\\
&=\sum_{u=g}^{m}\frac{\begin{pmatrix} i-1\\ u\end{pmatrix}\begin{pmatrix} N-i-g+2\\ m-u\end{pmatrix}}{\begin{pmatrix} N-g+1\\ m\end{pmatrix}}.
\end{aligned}$$

This probability is equal to $P_{a_g(i)=0}\left(B_g^1(i)|B_1^2(i)\cap B_2^2(i)...\cap B_{g-1}^2(i)\right)$.
This completes the proof. □

We also required the probability of selecting of a greater unit from the i-th population unit. This probability is stated at Theorem 2.3.

2.3. Theorem. $P_{a_g(i)}\left(B_g^2(i)|B_1^{l_1(i)}(i)\cap B_2^{l_2(i)}(i)\cap ...\cap B_{g-1}^{l_{g-1}(i)}(i)\right)$ *can be written as follows:*

$$\begin{aligned}
&P_{a_g(i)}\left(B_g^2(i)|B_1^{l_1(i)}(i)\cap B_2^{l_2(i)}(i)\cap ...\cap B_{g-1}^{l_{g-1}(i)}(i)\right)\\
&=1-P_{a_g(i+1)=a_g(i)}\left(B_g^1(i+1)|B_1^{l_1(i+1)}(i+1)\cap ...\cap B_{g-1}^{l_{g-1}(i+1)}(i+1)\right). \qquad (2.5)
\end{aligned}$$

Proof. From the basic complement rule of probability, $P(A^c) = 1 - P(A)$, we know that $P_{a_g(i)}\left(\{B_g^2(i) \ |B_1^{l_1(i)}(i) \cap B_2^{l_2(i)}(i) ... \cap B_{g-1}^{l_{g-1}(i)}(i)\}\right)$ is the selection probability of a greater unit from i-th population unit (i+1,i+2,…,N) when $a_g(i)$ is known and $P_{a_g(i+1)=a_g(i)}\left(B_g^1(i+1)|B_1^{l_1(i+1)}(i+1) \cap ... \cap B_{g-1}^{l_{g-1}(i+1)}(i+1)\right)$ is the selection probability of a smaller unit from (i+1)-*th* population unit (1,2,… i) when $a_g(i+1) = a_g(i)$. So, these probabilities are complement to each other. This completes the proof. □

By using these definitions, the probability of selecting both the ith and jth population units in the sample can be obtained as

$$\begin{aligned} \pi_N(A_i \cap A_j) &= 1 - \pi_N\left((A_i \cap A_j)^c\right) \qquad i,\ j = 1, 2, \ldots, N\ (i < j) \\ &= 1 - \left[\pi_N\left(A_i^c\right) + \pi_N\left(A_j^c\right) - \pi_N\left(A_i^c \cap A_j^c\right)\right] \end{aligned} \tag{2.6}$$

where $\pi_N(A_i^c)$ and $\pi_N\left(A_j^c\right)$ probabilities can be calculated from the Theorems 2.1, 2.2, 2.3. The probability $\pi_N(A_i^c \cap A_j^c)$ can be defined as follows;

$$\begin{aligned} \pi_N(A_i^c \cap A_j^c) &= \sum_{l_1(i,j),l_2(i,j),\ldots,l_m(i,j)=1}^{3} P\left(B_1^{l_1(i,j)}(i,j) \cap ... \cap B_m^{l_m(i,j)}(i,j)\right) \\ &= \sum_{l_1(i,j),l_2(i,j),\ldots,l_m(i,j)=1}^{3} \prod_{g=1}^{m} P_{a_g(i),a_g(j)} \\ &\qquad \left(B_g^{l_g(i,j)}(i,j)|B_1^{l_1(i,j)}(i,j) \cap ... \cap B_{g-1}^{l_{g-1}(i,j)}(i,j)\right) \end{aligned} \tag{2.7}$$

The conditional probability of $B_g^{l_g(i,j)}(i,j)$ can be calculated from Theorems 2.1, 2.2, 2.3. when $l_g(i,j)$=1 and $l_g(i,j)$=2. When $l_g(i,j)$=3, the conditional probability of $B_g^{l_g(i,j)}(i,j)$ is given as following Theorem 2.4.

2.4. Theorem. $P_{a_g(i),a_g(j)}\left(B_g^3(i,j)|B_1^{l_1(i,j)}(i,j) \cap B_2^{l_2(i,j)}(i,j) ... \cap B_{g-1}^{l_{g-1}(i,j)}(i,j)\right)$ *can be written as follows:*

$$\begin{aligned} & P_{a_g(i),a_g(j)}\left(B_g^3(i,j)|B_1^{l_1(i,j)}(i,j) \cap ... \cap B_{g-1}^{l_{g-1}(i,j)}(i,j)\right) \\ &= P_{a_g(j)}\left(B_g^1(j)|B_1^{l_1(j)}(j) \cap ... \cap B_{g-1}^{l_{g-1}(j)}(j)\right) \\ &\qquad - P_{a_g(i+1)=a_g(i)}\left(B_g^1(i+1)|B_1^{l_1(i+1)}(i+1) \cap ... \cap B_{g-1}^{l_{g-1}(i+1)}(i+1)\right) \end{aligned} \tag{2.8}$$

Proof. $P_{a_g(j)}\left(B_g^1(j)|B_1^{l_1(j)}(j) \cap ... \cap B_{g-1}^{l_{g-1}(j)}(j)\right)$ is the probability of selecting smaller population unit than the j^{th} population unit in the g^{th} selection when there are a$_g(j)$ smaller unit then jth population unit. Also,

$$P_{a_g(i+1)=a_g(i)}\left(B_g^1(i+1)|B_1^{l_1(i+1)}(i+1) \cap ... \cap B_{g-1}^{l_{g-1}(i+1)}(i+1)\right)$$

is the probability of selecting a smaller population unit than the $(i{+}1)^{th}$ population unit in the g^{th} selection when there are $a_g(i{+}1)$= $a_g(i)$ smaller units then $(i{+}1)^{th}$ population unit. So, from the basic rules of probability, the probability of a population unit between i^{th} and j^{th} unit including in a ranked set sample can be obtained by using the difference of these two probabilities. This completes the proof. □

By using Theorem 2.1, 2.2, 2.3 and 2.4 we can obtain the inclusion probabilities given in Eq. (2.1) and (2.6). A simple example for calculation is given in the following section.

3. Computation of the Formula

By using the formulas in previous section, the inclusion probabilities for the all units in the population can be derived easily. For example, when N=5 and m=3, the population consists of $X_1<X_2<X_3<X_4<X_5$ elements. The inclusion probability of X_i (i=1, 2, 3, 4, 5) can be written using Eq. (2.1) as follows:

$\pi_N(A_i) = 1 - \pi_N(A_i^c)$ i=1,2,3,4,5

where

$$\pi_N(A_i^c) = \sum_{l_1,l_2,l_3=1}^{2} P\left(B_1^{l_1}(i)\cap B_2^{l_2}(i)\cap B_3^{l_3}(i)\right)$$

$$= \sum_{l_1,l_2,l_3=1}^{2} P_{a_3(i)}\left(B_3^{l_3}(i)|B_1^{l_1}(i)\cap B_2^{l_2}(i)\right) P_{a_2(i)}\left(B_2^{l_2}(i)|B_1^{l_1}(i)\right) P_{a_1(i)}\left(B_1^{l_1}(i)\right).$$

here $a_1(i) = 0$, $a_2(i) = 0, 1$ and $a_3(i) = 0, 1, 2$.

By using Theorem 2.1, the probability of selecting a smaller unit than the i-th population unit when g=1, can be written as follows;

$$P_0\left(B_1^1(i)\right) = \begin{cases} 0 & i=1 \\ \sum_{u=1}^{3} \dfrac{\binom{i-1}{u}\binom{6-i}{3-u}}{\binom{5}{3}} & i=2,3 \\ 1 & i=4,5. \end{cases}$$

$P_0\left(B_1^1(1)\right) = 0;\ P_0\left(B_1^1(2)\right) = 6/10;$

$P_0\left(B_1^1(3)\right) = 9/10;\ P_0\left(B_1^1(4)\right) = 1;\ P_0\left(B_1^1(5)\right) = 1.$

From Theorem 2.3, it can be written as follows;

$P_0\left(B_1^2(1)\right) = 4/10;\ P_0\left(B_1^2(2)\right) = 1/10;$

$P_0\left(B_1^2(3)\right) = 0;\ P_0\left(B_1^2(4)\right) = 0;\ P_0\left(B_1^2(5)\right) = 0.$

We can write the other inclusion probabilities by the same way. In Table 1, the inclusion probabilities of the all population units are given in the g^{th} selection for all possible combinations of (l_1, l_2, l_3). In Table 1, P_{l_1,l_2,l_3} is defined as follows;

$$\begin{aligned} P_{l_1,l_2,l_3} &= P\left(B_1^{l_1}(i)\cap B_2^{l_2}(i)\cap B_3^{l_3}(i)\right) \\ &= P_{a_3(i)}\left(B_3^{l_3}(i)|B_1^{l_1}(i)\cap B_2^{l_2}(i)\right) P_{a_2(i)}\left(B_2^{l_2}(i)|B_1^{l_1}(i)\right) P_{a_1(i)}\left(B_1^{l_1}(i)\right). \end{aligned}$$

The obtained inclusion probabilities in Table 1 are the same as the inclusion probabilities which are given in the study of Gökpınar and Özdemir (2010). But this formula is much easier and simpler than the formula of the inclusion probabilities given in Gökpınar and Özdemir (2010).

By the same way, the second order inclusion probabilities can be obtained as given in Table 2.

As seen from Table 1, the extreme units have greater inclusion probabilities than the others. The following figures are constructed for different population and set sizes.

As seen from Figures 1-6, units from both extremes (e.g. X_1, X_N) have greater second order inclusion probabilities than the others for all set and population sizes. Also units in the mid section of the population have smaller second order inclusion probabilities. The effects of first and second order inclusion probabilities on HT estimator under populations with different coefficient of variation and skewness values are investigated at Gökpınar

TABLE 1. The first order inclusion probabilities of the population units with N=5, m=3

X_i	(l_1,l_2,l_3)	(1,1,1)	(1,1,2)	(1,2,1)	(1,2,2)	(2,1,1)	(2,1,2)	(2,2,1)	(2,2,2)	$\pi_N(A_i^c)$	$\pi_N(A_i)$
X_1	g=1	0	0	0	0	0.40	0.40	0.40	0.40		
	g=2	0	0	1	1	0	0	1	1	0.40	0.60
	g=3	0	1	0	1	0	1	0	1		
P_{l_1,l_2,l_3}		0	0	0	0	0	0	0	0.40		
X_2	g=1	0.60	0.60	0.60	0.60	0.10	0.10	0.10	0.10		
	g=2	0	0	1	1	0	0	0.50	0.50	0.65	0.35
	g=3	0	1	0	1	0	1	0	1		
P_{l_1,l_2,l_3}		0	0	0	0.60	0	0	0	0.05		
X_3	g=1	0.90	0.90	0.90	0.90	0	0	0	0		
	g=2	0	0	0.50	0.50	0.50	0.50	0	0	0.45	0.55
	g=3	0	1	0	1	0	1	0	0		
P_{l_1,l_2,l_3}		0	0	0	0.45	0	0	0	0		
X_4	g=1	1	1	1	1	0	0	0	0		
	g=2	0.50	0.50	0	0	1	1	0	0	0.50	0.50
	g=3	0	1	0	0	0	0	1	1		
P_{l_1,l_2,l_3}		0	0.50	0	0	0	0	0	0		
X_5	g=1	1	1	1	1	0	0	0	0		
	g=2	1	1	0	0	1	1	0	0	0	1
	g=3	0	0	1	0	1	0	0	0		
P_{l_1,l_2,l_3}		0	0	0	0	0	0	0	0		

TABLE 2. The second order inclusion probabilities of the population units with N=5, m=3

$\pi_N(A_i \cap A_j)$	X_1	X_2	X_3	X_4	X_5
X_1	-	0	0.30	0.30	0.60
X_2	0	-	0.20	0.15	0.35
X_3	0.30	0.20	-	0.05	0.55
X_4	0.30	0.15	0.05	-	0.50
X_5	0.60	0.35	0.55	0.50	-

and Özdemir(2012). The results of assigning larger probabilities to the extremes are also discussed at Gökpınar and Özdemir(2012).

4. Concluding Remarks

In this study, we give a new formula for the first and the second order inclusion probabilities in RSS which is simpler and easier than the previous ones. This formula can be adapted to other modifications of RSS and can be generalized for any cycle sizes. Furthermore, a MATLAB code is given for calculate the inclusion probabilities in the Appendix.

Figure 1. The second order inclusion probabilities of the population units with N=20, m=3

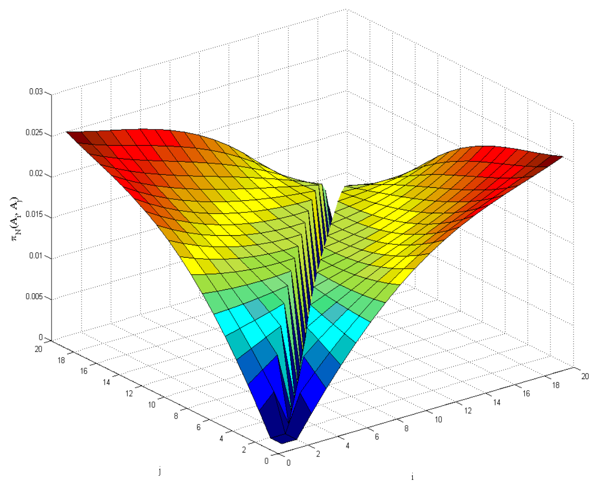


Figure 2. The second order inclusion probabilities of the population units with N=20, m=5

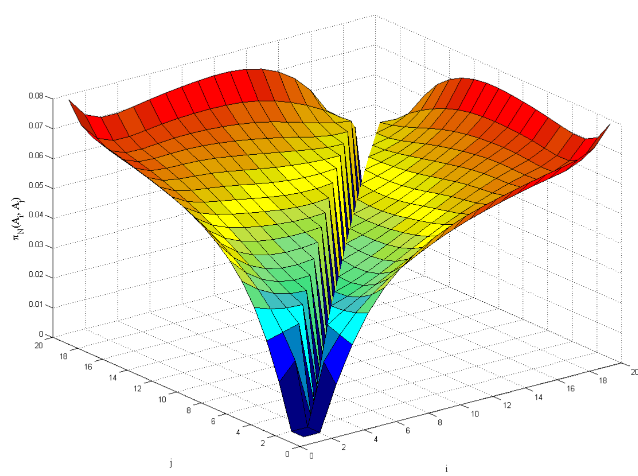


FIGURE 3. The second order inclusion probabilities of the population units with N=20, m=7

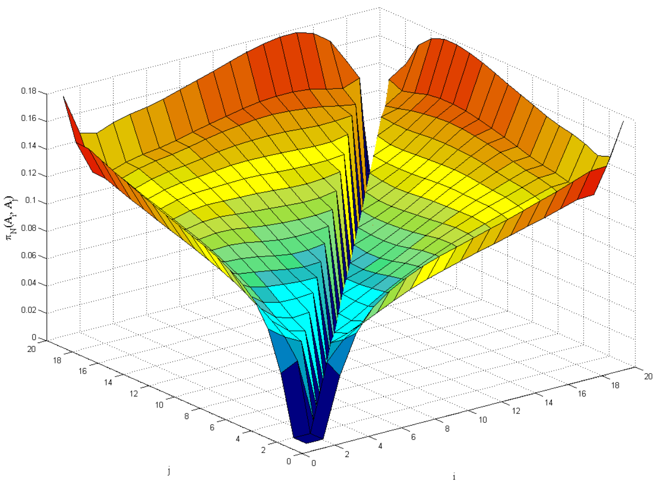


FIGURE 4. The second order inclusion probabilities of the population units with N=50, m=3

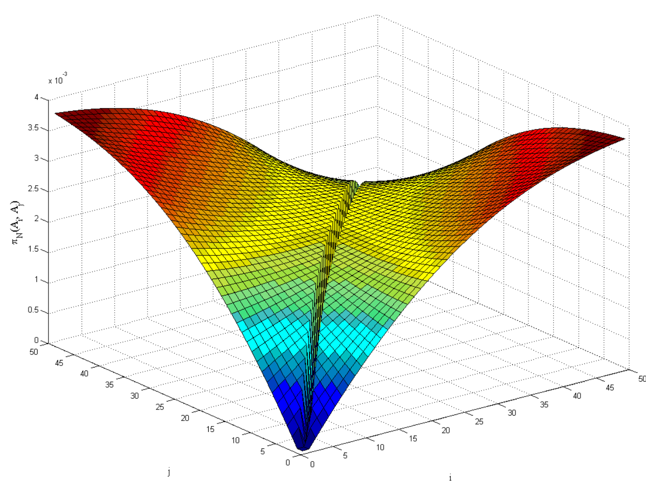


FIGURE 5. The second order inclusion probabilities of the population units with N=50, m=5

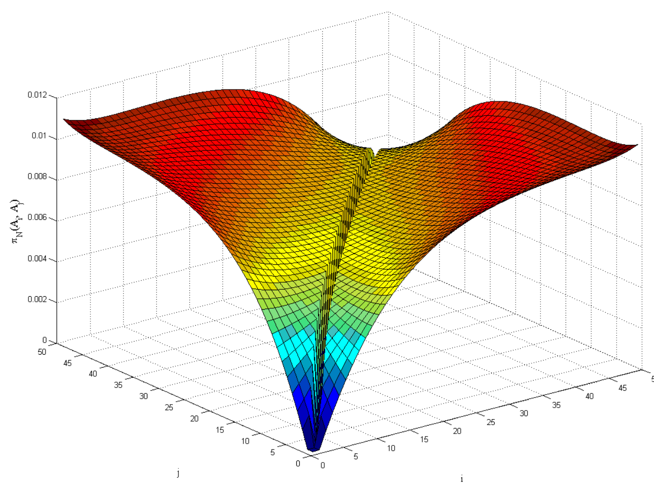
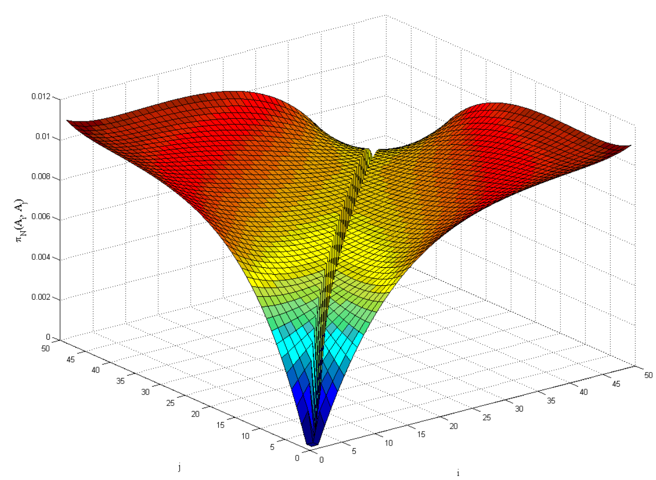


FIGURE 6. The second order inclusion probabilities of the population units with N=50, m=7



References

[1] Al Saleh, M.F. and Samawi H.M. *A note on inclusion probability in ranked set sampling and some of its variations*, Test **16**, 198-209, 2007.

[2] Al-Saleh, M.F. and Al-Shrafat K. *Estimation of average milk yield using ranked set sampling*, Environmetrics **12**(4), 395-399, 2000.

[3] Bouza, C. N. *Sampling using ranked sets: Concepts, results and perspectives*, Revista Investigación Operacional **26** (3), 275-292, 2005.

[4] Deshpande, J.V., Frey, J., Oztürk, O. *Nonparametric ranked-set sampling confidence intervals for quantiles of a finite population*, Environmental and Ecological Statistics **13**, 25-40, 2006.

[5] Johnson, G.D., Myers, W.L. *Potential of ranked-set sampling for disaster assessment.* IUFRO S4.02 Conference on Inventory and Management Techniques in the Context of Catastrophic Events, 1993.

[6] Gökpınar, F., Özdemir, Y. A. *Generalization of the inclusion probabilities in ranked set sampling*, Hacettepe Journal of Mathematics and Statistics **39**(1), 89-95, 2010.

[7] Gökpınar, F., Özdemir, Y. A. *Horvitz-Thompson Estimator Of The Population Mean Using Inclusion Probabilities Of Ranked Set Sampling*, Communications In Statistics: Theory and Methods **41**(6), 1029-1039, 2011.

[8] Jafari Jozani, M. and Johnson, B.C. *Design based estimation for ranked set sampling in finite populations*, Environ. Ecol. Stat. **18**, 663-685, 2011.

[9] McIntyre, G.A. *A method of unbiased selective sampling, using ranked sets*, Australian Journal of Agricaltural Research **3**, 385-390, 1952.

[10] Özdemir, Y.A. and Gökpınar, F. *A Generalized formula for inclusion probabilities in ranked set sampling*, Hacettepe University Bulletin of Natural Sciences & Engineering Series B: Mathematics & Statistics **36**, 89-99, 2007.

[11] Özdemir, Y.A. and Gökpınar, F. *A new Formula for inclusion probabilities in median ranked set Sampling*, Communications In Statistics:Theory and Methods. **37**, 2022-2033, 2008.

Appendix A. Matlab Code for First Order Inclusion Probabilities

```
function P =firstinc(N,m)
B(1:m,1:m,1:N,1:2)=0;
for i=1:N
for g=1:m
for u=g:m
B(1,g,i,1)=B(1,g,i,1)+nck(i-1,u)*nck(N-i-g+2,m-u)/nck(N-g+1,m);
if i>1
B(1,g,i-1,2)=1-B(1,g,i,1);
end
end
end
end
for ag=2:m
for i=1:N
for g=ag:m
for u=g:m
B(ag,g,i+ag-1,1)=B(1,g,i,1);
if i>1
B(ag,g,i-1,2)=1-B(ag,g,i,1);
end
end
end
```

```
end
end
A=allperm([1 2],m);
for i=1:N
AT(:,1,i)=B(1,1,i,A(:,1));
end
for i=1:N
for j=2:m
for t=1:2^m
AT(t,j,i)=B(2*j-1-sum(A(t,1:j-1)),j,i,A(t,j));
end
end
end
for i=1:N
for t=1:2^m
c(i,t)=1;
for j=1:m
c(i,t)=c(i,t)*AT(t,j,i);
end
end
end
P=1-sum(c');
```

B. Matlab Code for Second Order Inclusion Probabilities

```
function P2=secondinc(N,m)
B(1:m,1:m,1:N,1:2)=0;
B3(1:m,1:m,1:m,1:N,1:N)=0;
for i=1:N
for g=1:m
for u=g:m
B(1,g,i,1)=B(1,g,i,1)+nck(i-1,u)*nck(N-i-g+2,m-u)/nck(N-g+1,m);
if i>1
B(1,g,i-1,2)=1-B(1,g,i,1);
end
end
end
end
for ag=2:m
for i=1:N
for g=ag:m
for u=g:m
B(ag,g,i+ag-1,1)=B(1,g,i,1);
if i>1
B(ag,g,i-1,2)=1-B(ag,g,i,1);
end
end
end
end
end
for aig=1:m
for ajg=aig:m
for i=1:N
```

```
for j=i+1:N
for g=1:m
B3(aig,ajg,g,i,j)=B(ajg,g,j,1)-B(aig,g,i+1,1);
if B3(aig,ajg,g,i,j)<0
B3(aig,ajg,g,i,j)=0;
end
end
end
end
end
end
A=allperm([1 2 3],m);
for k=1:size(A,1)
for l=1:size(A,2)
if A(k,l)==1;
AA{k,l}={1 1};
elseif A(k,l)==3;
AA{k,l}={2 1};
elseif A(k,l)==2;
AA{k,l}={2 2};
end
end
end
for i=1:N-1
for j=i+1:N
for k=1:3^m
if A(k,1)==1;
AT(i,j,k,1)=B(1,1,i,1);
elseif A(k,1)==2;
AT(i,j,k,1)=B(1,1,j,2);
elseif A(k,1)==3;
AT(i,j,k,1)=B3(1,1,1,i,j);
end
end
end
end
for i=1:N-1
for j=i+1:N
for l=2:size(A,2)
for k=1:3^m
if A(k,l)==1;
aa=2*l-1;
for t=1:l-1
aa=aa-AA{k,t}{1};
end
AT(i,j,k,l)=B(aa,l,i,AA{k,l}{1});
elseif A(k,l)==2
aa=2*l-1;
for t=1:l-1
aa=aa-AA{k,t}{2};
end
```

```
AT(i,j,k,l)=B(aa,l,j,AA{k,l}{2});
elseif A(k,l)==3
aai=2*l-1;
for t=1:l-1
aai=aai-AA{k,t}{1};
end
aaj=2*l-1;
for t=1:l-1
aaj=aaj-AA{k,t}{2};
end
AT(i,j,k,l)=B3(aai,aaj,l,i,j);
end
end
end
end
end
for i=1:N-1
for j=i+1:N
for k=1:3^m
c(i,j,k)=1;
for l=1:m
c(i,j,k)=c(i,j,k)*AT(i,j,k,l);
end
end
P(i,j)=sum(c(i,j,:));
end
end
P(N,1:N)=0;
P1=firstinc(N,m);
for i=1:N-1
for j=i+1:N
P2(i,j)=1-((1-P1(i))+(1-P1(j))-P(i,j));
end
end
```

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 131 – 140

A NEW CLASS OF EXPONENTIAL REGRESSION CUM RATIO ESTIMATOR IN TWO PHASE SAMPLING

Nilgun Ozgul * † and Hulya Cingi * ‡

Received 07 : 03 : 2013 : Accepted 15 : 05 : 2013

Abstract

In this paper, we propose a new class of exponential regression cum ratio estimator using the auxiliary variable for the estimation of the finite population mean under two phase sampling scheme. The Bias and Mean Square Error (MSE) equations of the proposed estimator are obtained and compared with the MSE equations of some existing estimators in two phase sampling. We find theoretically the proposed estimator is always more efficient than classical ratio and regression estimators, Singh and Vishwakarma [17] ratio type exponential estimator in two phase sampling. In addition, theoric results are supported by a numerical example using original data sets.

Keywords: Two phase sampling, Auxiliary variable, Exponential estimation, Efficiency.

2000 AMS Classification:

1. Introduction

In the sampling theory, the use of auxiliary information results in considerable improvement in the precision of estimators of population mean. The ratio and regression methods have been widely used when auxiliary information is available. In literature, number of authors introduced many ratio and regression type estimators by using general linear transformation of the auxiliary variable. For recent development, exponential estimators have been widely studied by several authors such as Bahl and Tuteja [2], Singh et al. [19] and Grover and Kaur [6].

Under various sampling schemes, many exponential estimators, using the population information of the auxiliary variable, have been proposed. However, the knowledge on the population mean of the auxiliary variable is not always available. In this situation, two phase sampling method is the most popular sampling scheme in literature. Two

*Hacettepe University, Department of Statistics, Beytepe, 06800, Ankara, Turkey.
†Email: (N. Ozgul) nozgul@hacettepe.edu.tr
‡Email: (H. Cingi) hcingi@hacettepe.edu.tr

phase sampling, first introduced by Neyman [13], is a cost effective technique in survey sampling. It is typically used when it is very expensive to collect data on the variables of interest, but it is relatively inexpensive to collect data on variables that are correlated with the variables of interest. By these aspects two phase sampling is a powerful and cost economical procedure for finding the reliable estimate in first phase sample for the unknown parameters of the auxiliary variable x. Simply, a field survey is to be undertaken to determine the average value of some characters of a population. For example, the amount of money families spend on food. As the collection of data requires long interviews by specially trained enumerators, the cost per family is quite high. The cost of survey is constrained within a specified amount but the sample does not appear to yield an estimate of desired precision because of the great variability of the character. Nevertheless, the character is correlated with a second character that can be determined at a lower cost per family so that a precise estimate of the distribution of this second character is readily obtained. Hence, a more precise estimate of the original character can be found by first estimating the distribution of the second character alone from a large random sample [10]. In literature, many authors improved ratio and regression estimators using at least one auxiliary variable under two phase sampling scheme. Singh and Espejo [16] suggested a class of ratio-product estimators in two phase sampling with its properties and identified asymptotically optimum estimators from proposed class of estimators. Samiuddin and Hanif [14] proposed ratio and regression estimation procedures to estimate the population mean in two-phase sampling using idea of partial and no information cases. Ahmad [1] has proposed various estimators for two phase and multiphase sampling using information on several auxiliary variables. Hanif et al. [7] proposed regression estimator using several auxiliary variables. In recent years, exponential estimators have not been studied sufficiently in two phase sampling. Singh and Vishwakarma [17] adapted Bahl and Tuteja [2] exponential ratio type estimator into two phase sampling. We, here, give the notations about two phase sampling and various estimators of the population mean in two phase sampling method in Section 2. We propose a class of exponential regression cum ratio estimator in Section 3. In Section 4, the proposed estimator is compared with other existing estimators in two phase sampling and we obtain certain conditions that proposed estimator is found to be more efficient than other estimators. In Section 5, the theoretical results are supported by a numerical example. In Section 6, we give conclusion.

2. Notations and Various Existing Estimators

Consider a finite population $U = U_1, U_2, \ldots, U_N$, of size N units. Let y denote the study variable taking the values y_i on the unit U_i, $(i = 1, 2, \ldots, N)$ and $\overline{Y}$ is its unknown population mean. Let x denotes the auxiliary variable taking the values x_i on the unit U_i, $(i = 1, 2, \ldots, N)$ positively correlated with y and $\overline{X}$ is its unknown population mean.

It is well known that when the population mean of the auxiliary variable is not known, two phase sampling is used. Two phase sampling consists of two phase. In first phase, a sample of fixed size is drawn by Simple Random Sampling Without Replacement (SR-SWOR) from the finite population to estimate the mean of the auxiliary variable. The sample is drawn in first phase is named as primary sample and expressed by s'. The usual practice is to estimate the mean of the auxiliary variable by sample mean. In second phase, a sample s $(s \subset s')$ of fixed size n is drawn SRSWOR from the primary sample (s') to estimate the mean of the study variable. The sample is drawn in second phase is named as sub sample and expressed by s [14].

When information is not available on the auxiliary variable, x, that is positively correlated with the study variable, y, the classical ratio estimator is a widely used estimator to estimate the population mean, $\overline{Y}$, in two phase sampling as follows:

$$\overline{y}_R = \frac{\overline{y}}{\overline{x}}\overline{x}' \tag{2.1}$$

where $\overline{x}'$ is the primary sample mean of the auxiliary variable, $\overline{y}$ and $\overline{x}$ are the sub sample means of the study and auxiliary variables, respectively. It is well known that the MSE equation of the classical ratio estimator is given by

$$MSE\left(\overline{y}_R\right) \cong \overline{Y}^2\left[\lambda C_y^2 + \lambda^* C_x^2\left(1 - 2K_{yx}\right)\right] \tag{2.2}$$

where $K_{yx} = \rho_{yx}\dfrac{C_y}{C_x}$; $\lambda = \dfrac{1}{n} - \dfrac{1}{N}$; $\lambda^* = \dfrac{1}{n} - \dfrac{1}{n'}$; n' is the primary sample size; n is the sub sample size; N is the number of units in the population; ρ_{yx} is the population correlation coefficient between the auxiliary and the study variables, C_x and C_y are the population coefficients of variation of the auxiliary and study variables, respectively.

When auxiliary variable is correlated with the study variable, the classical unbiased regression estimator is used to estimate the population mean, in two phase sampling as follows:

$$\overline{y}_{lr} = \overline{y} + \beta_{yx}\left(\overline{x}' - \overline{x}\right) \tag{2.3}$$

where β_{yx} is the regression coefficient between the auxiliary and the study variables. It is well known that the variance of the classical regression estimator is given by

$$Var\left(\overline{y}_{lr}\right) = \overline{Y}^2 C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right) \tag{2.4}$$

Singh and Vishwakarma [17] suggested the following modified exponential ratio estimator in two phase sampling

$$\overline{y}_{svr} = \overline{y}\exp\left(\frac{\overline{x}' - \overline{x}}{\overline{x}' + \overline{x}}\right) \tag{2.5}$$

The MSE equation of the estimator can be given by

$$MSE\left(\overline{y}_{svr}\right) \cong \overline{Y}^2\left[\lambda C_y^2 + \lambda^*\left(\frac{C_x^2}{4} - \rho_{yx}C_yC_x\right)\right] \tag{2.6}$$

In sampling literature, the authors rarely consider the exponential estimators in two phase sampling scheme. For this reason, we improved a class of exponential regression cum ratio estimator in two phase sampling using the ratio and regression methods and their linear transformation in this study.

3. Suggested Exponential Estimator in Two Phase Sampling

Replacing regression estimator instead of sample mean and using linear transformation in exponential term in Singh and Vishwakarma [17] exponential ratio estimator given in (2.5), we improve a class of exponential regression cum ratio estimator as follows:

$$\overline{y}_{NH} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right]\exp\left(\frac{\overline{z}' - \overline{z}}{\overline{z} + \overline{z}'}\right) \tag{3.1}$$

where k_1 and k_2 are some known constants, $\overline{z}^{'}$ is a transformation of the auxiliary variable at first phase as $\overline{z}^{'} = a\overline{x}^{'} + b$, and $\overline{z}$ is a transformation of the auxiliary variable at second phase as $\overline{z} = a\overline{x} + b$.

Then, we have

$$\left.\begin{array}{l}\overline{z}^{'} = a\overline{x}^{'} + b \\ \overline{z} = a\overline{x} + b\end{array}\right\} \tag{3.2}$$

where $a\,(\neq 0)$ and b are either any known constants or functions of any known population parameters of the auxiliary variable, such as standard deviation (σ_x), coefficient of variation (C_x), coefficient of skewness $\{\beta_1(x)\}$, coefficient of kurtosis $\{\beta_2(x)\}$, coefficient of correlation (ρ_{yx}) [9]. The list of new exponential estimator generated from (3.1) is given in Table 1.

To obtain the Bias and MSE equations for the proposed estimator, we define following notations:

$$e_0 = \frac{\left(\overline{y} - \overline{Y}\right)}{\overline{Y}}, e_1 = \frac{\left(\overline{x} - \overline{X}\right)}{\overline{X}}, {e_1}' = \frac{\left(\overline{x}^{'} - \overline{X}\right)}{\overline{X}} \tag{3.3}$$

such that

$$\begin{array}{l} E\left(e_0\right) = E\left(e_1\right) = E\left(e_1^{'}\right) = 0; E\left(e_0^2\right) = \lambda C_y^2; E\left(e_1^2\right) = \lambda C_x^2; \\ E\left(e_1^{'2}\right) = \lambda^{'} C_x^2; E\left(e_0 e_1\right) = \lambda \rho_{yx} C_y C_x; E\left(e_0 e_1^{'}\right) = \lambda^{'} \rho_{yx} C_y C_x; \\ E\left(e_1 e_1^{'}\right) = \lambda^{'} C_x^2 \end{array} \tag{3.4}$$

where

$$\lambda = \frac{1}{n} - \frac{1}{N},\ \ \lambda^{'} = \frac{1}{n^{'}} - \frac{1}{N},\ \ C_y^2 = \frac{S_y^2}{\overline{Y}^2},\ \ C_x^2 = \frac{S_x^2}{\overline{X}^2},\ \ S_y^2 = \frac{\sum_{i=1}^{N}\left(y_i - \overline{Y}\right)^2}{N-1},$$

$$S_x^2 = \frac{\sum_{i=1}^{N}\left(x_i - \overline{X}\right)^2}{N-1},\ \ \rho_{yx} = \frac{\sum_{i=1}^{N}\left(y_i - \overline{Y}\right)\left(x_i - \overline{X}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i - \overline{X}\right)^2 \sum_{i=1}^{N}\left(y_i - \overline{Y}\right)^2}}$$

and we use Taylor series method [4] for two variables to solve the exponential term as

$$\begin{aligned} f\left(e_1, e_1^{'}\right) &= f\left(e_1, e_1^{'}\right)\Big|_{e_1 = e_1^{'} = 0} + \frac{1}{1!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1}\Big|_{e_1 = e_1^{'} = 0} \\ &+ \frac{1}{1!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1^{'}}\Big|_{e_1 = e_1^{'} = 0} + \frac{1}{2!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1^2}\Big|_{e_1 = e_1^{'} = 0} \\ &+ \frac{1}{2!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1^{'2}}\Big|_{e_1 = e_1^{'} = 0} + \frac{1}{2!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1 e_1^{'}}\Big|_{e_1 = e_1^{'} = 0} \\ &+ \frac{1}{2!}\frac{\partial f\left(e_1, e_1^{'}\right)}{\partial e_1^{'} e_1}\Big|_{e_1 = e_1^{'} = 0} + \ldots \end{aligned} \tag{3.5}$$

Expressing (3.1) in terms of e's and using (3.5) for the exponential term, we have

$$\overline{y}_{NH} = \left[k_1\overline{Y}\left(1 + e_0\right) + k_2\overline{X}\left(e_1^{'} - e_1\right)\right]\exp\left\{\frac{a\overline{X}\left(e_1^{'} - e_1\right)}{a\overline{X}\left(e_1 + e_1^{'} + 2\right) + 2b}\right\} \tag{3.6}$$

where $f(e_1, e_1^{'}) = \exp\left\{\dfrac{a\overline{X}\left(e_1^{'} - e_1\right)}{a\overline{X}\left(e_1 + e_1^{'} + 2\right) + 2b}\right\}$ and we solve the exponential term from (3.5) as

$$\begin{aligned}\overline{y}_{NH} &= \left[k_1\overline{Y}\left(1+e_0\right) + k_2\overline{X}\left(e_1^{'} - e_1\right)\right]\\ &\left\{1 - \theta\left(e_1 - e_1^{'}\right) + \frac{3\theta^2}{2}e_1^2 - \frac{\theta^2}{2}e_1^{'\,2} - \theta^2 e_1 e_1^{'} + ...\right\}\end{aligned}$$

where $\theta = \dfrac{a\overline{X}}{2\left(a\overline{X}+b\right)}$.

Assuming $|e_1| < 1$, expanding the right hand side of (3.6), and retaining terms up to the second degree of e's, we have

$$\begin{aligned}\overline{y}_{NH} - \overline{Y} \cong \overline{Y}&\left[(k_1 - 1) - k_1\theta\left(e_1 - e_1^{'}\right) - \frac{3\theta^2}{2}\left(e_1^2 - e_1^{'\,2}\right) + k_1 e_0 - \right.\\ &\left. k_1\theta\left(e_0e_1 - e_0e_1^{'}\right)\right] + k_2\overline{X}\left[e_1^{'} - e_1 + \theta\left(e_1^2 - e_1^{'\,2}\right)\right]\end{aligned} \tag{3.7}$$

Squaring both sides of (3.7), retaining terms of e's up to the second degree and taking expectation, we get the Bias and MSE Equations of $\overline{y}_{NH}$ to the second degree of approximation as

$$Bias\left(\overline{y}_{NH}\right) \cong E\left(\overline{y}_{NH} - \overline{Y}\right) \cong \overline{Y}\left[(k_1 - 1) + k_1\lambda^*\theta C_x^2\right]\left(\frac{3\theta}{2} - K_{yx}\right) \tag{3.8}$$

$$\begin{aligned}MSE\left(\overline{y}_{NH}\right) &\cong E\left(\overline{y}_{NH} - \overline{Y}\right)^2\\ &\cong \bar{Y}^2\left[(k_1-1)^2 + k_1^2\left\{\lambda C_y^2 + 4\lambda^*\theta C_x^2\left(\theta - K_{yx}\right)\right\} + \right.\\ &\left. k_1\lambda^*\theta C_x^2\left(2K_{yx} - 3\theta\right) + k_2^2\lambda^*\overline{X}^2C_x^2 + 2k_2\overline{X}\,\overline{Y}\lambda^*C_x^2\left\{k_1\left(2\theta - K_{yx}\right)\right\}\right.\end{aligned} \tag{3.9}$$

To obtain the minimum $MSE\left(\overline{y}_{NH}\right)$, we get

$$\frac{\partial}{\partial k_i}\left\{MSE\left(\overline{y}_{NH}\right)\right\} = 0; i = 1,\,2. \tag{3.10}$$

Solving two equations simultaneously, the optimum values of k_1 and k_2 are respectively,

$$k_1 = 1 - \frac{2 - \lambda^*\theta^2C_x^2}{1 + \left(\lambda - \lambda^*\rho_{yx}^2\right)} \tag{3.11}$$

$$k_2 = \frac{\overline{Y}}{\overline{X}}\left\{(\theta - 1) + \frac{2 - \lambda^*\theta^2C_x^2}{1 + \left(\lambda - \lambda^*\rho_{yx}^2\right)}\left(2\theta - K_{yx}\right)\right\} \tag{3.12}$$

k_1 and k_2 quantities can be guessed quite accurately through a pilot sample survey or sample data or experience gathered in due course of time, see Das and Tripathi [5], Singh and Ruiz-Espejo [16], Singh, H.P. et al. [18] and Koyuncu and Kadilar [11].

When k_1 and k_2 are replaced in (3.9), the minimum MSE of the proposed estimator can be written as

$$(3.13)\quad MSE_{\min}\left(\overline{y}_{NH}\right) \cong \overline{Y}^2 \frac{C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)\left(1-\lambda^*\theta^2C_x^2\right)-\dfrac{\lambda^{*2}\theta^4C_x^4}{4}}{\left\{1+C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)\right\}}$$

$$\cong \overline{Y}^2 \frac{Var\left(\overline{y}_{lr}\right)\left(1-\lambda^*\theta^2C_x^2\right)-\dfrac{\lambda^{*2}\overline{Y}^2\theta^4C_x^4}{4}}{\left\{\overline{Y}^2+Var\left(\overline{y}_{lr}\right)\right\}}$$

Table 1. Some Members of the Suggested Estimator $\overline{y}_{NH}$

A subset of $\overline{y}_{NH}$	**a**	**b**
$\overline{y}_{NH1}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left(\frac{\overline{x}'-\overline{x}}{\overline{x}+\overline{x}'}\right)$	1	0
$\overline{y}_{NH2}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left(\frac{\overline{x}'-\overline{x}}{\overline{x}+\overline{x}'+2}\right)$	1	1
$\overline{y}_{NH3}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left(\frac{\overline{x}'-\overline{x}}{\overline{x}+\overline{x}'+2\beta_2(x)}\right)$	1	$\beta_2(x)$
$\overline{y}_{NH4}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{\beta_2(x)\left(\overline{x}'-\overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2}\right\}$	$\beta_2(x)$	1
$\overline{y}_{NH5}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{C_x\left(\overline{x}'-\overline{x}\right)}{C_x\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	C_x	$\beta_2(x)$
$\overline{y}_{NH6}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{\beta_2(x)\left(\overline{x}'-\overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2C_x}\right\}$	$\beta_2(x)$	C_x
$\overline{y}_{NH7}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{\rho_{yx}\left(\overline{x}'-\overline{x}\right)}{\rho_{yx}\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	ρ_{yx}	$\beta_2(x)$
$\overline{y}_{NH8}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{\beta_2(x)\left(\overline{x}'-\overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2\rho_{yx}}\right\}$	$\beta_2(x)$	ρ_{yx}
$\overline{y}_{NH9}=\left[k_1\overline{y}+k_2\left(\overline{x}'-\overline{x}\right)\right]\exp\left\{\frac{C_x\left(\overline{x}'-\overline{x}\right)}{C_x\left(\overline{x}+\overline{x}'\right)+2\rho_{yx}}\right\}$	C_x	ρ_{yx}

4. Efficiency Comparisons in Two Phase Sampling

In this section, we obtain the efficiency conditions for the proposed estimator by comparing the MSE of the proposed estimators with the MSE of classical ratio and regression estimators and the exponential ratio estimator suggested by Singh and Vishwakarma [17].

We compare the MSE of the proposed estimator, $\overline{y}_{NH}$, given in (3.13), with the MSE of the existing estimators, $\overline{y}_R$, $\overline{y}_{lr}$, $\overline{y}_{svr}$.

From (2.2) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < MSE\left(\overline{y}_R\right)$$

$$\overline{Y}^2 \frac{C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)\left(1-\lambda^*\theta^2C_x^2\right)-\dfrac{\lambda^{*2}\theta^4C_x^4}{4}}{1+C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)} < \overline{Y}^2\left[\lambda C_y^2+\lambda^*C_x^2\left(1-2K_{yx}\right)\right]$$

Table 1 Continued: Some Members of the Suggested Estimator $\overline{y}_{NH}$

A subset of $\overline{y}_{NH}$	**a**	**b**
$\overline{y}_{NH10} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\rho_{yx}\left(\overline{x}' - \overline{x}\right)}{\rho_{yx}\left(\overline{x}+\overline{x}'\right)+2C_x}\right\}$	ρ_{yx}	C_x
$\overline{y}_{NH11} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\sigma_x\left(\overline{x}' - \overline{x}\right)}{\sigma_x\left(\overline{x}+\overline{x}'\right)+2\rho_{yx}}\right\}$	σ_x	ρ_{yx}
$\overline{y}_{NH12} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\rho_{yx}\left(\overline{x}' - \overline{x}\right)}{\rho_{yx}\left(\overline{x}+\overline{x}'\right)+2\sigma_x}\right\}$	ρ_{yx}	σ_x
$\overline{y}_{NH13} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2\sigma_x}\right\}$	$\beta_2\left(x\right)$	σ_x
$\overline{y}_{NH14} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\sigma_x\left(\overline{x}' - \overline{x}\right)}{\sigma_x\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	σ_x	$\beta_2\left(x\right)$
$\overline{y}_{NH15} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\beta_1(x)\left(\overline{x}' - \overline{x}\right)}{\beta_1(x)\left(\overline{x}+\overline{x}'\right)+2\beta_2(x)}\right\}$	$\beta_1\left(x\right)$	$\beta_2\left(x\right)$
$\overline{y}_{NH16} = \left[k_1\overline{y} + k_2\left(\overline{x}' - \overline{x}\right)\right] \exp\left\{\frac{\beta_2(x)\left(\overline{x}' - \overline{x}\right)}{\beta_2(x)\left(\overline{x}+\overline{x}'\right)+2\beta_1(x)}\right\}$	$\beta_2\left(x\right)$	$\beta_1\left(x\right)$

Note: In addition to estimators listed in Table 1, a large number of estimators can also be generated from (3.1) by putting 1, C_x, $\beta_2\left(x\right)$, ρ_{yx}, σ_x, $\beta_1\left(x\right)$ values for a and b.

$$\frac{C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1 + C_y^2\left(\lambda - \lambda^*\rho_{yx}^2\right)} < \lambda C_y^2 + \lambda^*\left(C_x - \rho_{yx}C_y\right)^2 - \lambda^*\rho_{yx}C_y$$

$$\left\{\frac{\lambda^*\theta^2 C_x^2}{2} + \frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\}^2 + \lambda^*\left(C_x - \rho_{yx}C_y\right)^2\left\{1 + \frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\} > 0 \tag{4.1}$$

The condition (4.1) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than the classical ratio estimator, $\overline{y}_R$.

From (2.4) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < Var\left(\overline{y}_{lr}\right)$$

$$\overline{Y}^2\frac{Var\left(\overline{y}_{lr}\right)\left(1 - \lambda^*\theta^2 C_x^2\right) - \dfrac{\lambda^{*2}\overline{Y}^2\theta^4 C_x^4}{4}}{\overline{Y}^2 + Var\left(\overline{y}_{lr}\right)} < Var\left(\overline{y}_{lr}\right)$$

$$\left\{\frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2} + \frac{\lambda^*\theta^2 C_x^2}{2}\right\}^2 > 0 \tag{4.2}$$

The condition (4.2) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than the classical regression estimator, $\overline{y}_{lr}$.

From (2.6) and (3.13), we have the condition

$$MSE\left(\overline{y}_{NH}\right) < MSE\left(\overline{y}_{svr}\right)$$

$$\overline{Y}^2 \frac{C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)\left(1-\lambda^*\theta^2 C_x^2\right)-\dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1+C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)} < \overline{Y}^2\left[\lambda C_y^2+\lambda^*\left(\frac{C_x^2}{4}-\rho_{yx}C_yC_x\right)\right]$$

$$\frac{C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)\left(1-\lambda^*\theta^2 C_x^2\right)-\dfrac{\lambda^{*2}\theta^4 C_x^4}{4}}{1+C_y^2\left(\lambda-\lambda^*\rho_{yx}^2\right)} < \lambda C_y^2+\lambda^*\left(\frac{C_x}{2}-\rho_{yx}C_y\right)-\lambda^*\rho_{yx}^2C_y^2$$

$$(4.3)\qquad \left(\frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}+\lambda^*\theta^2C_x^2\right)^2+\lambda^*\left(\frac{C_x}{2}-\rho_{yx}C_y\right)^2\left\{1+\frac{Var\left(\overline{y}_{lr}\right)}{\overline{Y}^2}\right\}>0$$

The condition (4.3) is always satisfied, the proposed estimator, $\overline{y}_{NH}$, is always more efficient than Singh and Vishwakarma [17] exponential ratio estimator, $\overline{y}_{svr}$.

Thus, finally, we conclude from the efficiency comparisons that the class of exponential regression cum ratio estimator, $\overline{y}_{NH}$, is always more efficient than the estimators, $\overline{y}_R$, $\overline{y}_{lr}$ and $\overline{y}_{svr}$.

5. Numerical Example

To show the performance of the proposed estimator in comparison to other estimators in two phase sampling, four original data sets used by other authors in literature has been considered. The descriptions of the populations are given below.

Population I : Cingi et. al. [3],

y : the number of teachers

x : the number of student in both primary and secondary school for 923 districts

$N=923$, $n^{'}=400$, $n=200$, $\overline{Y}=436,3$, $\overline{X}=11440,50$, $C_y=1,72$, $C_x=1,86$, $\rho_{yx}=0,955$.

Population II : Sukhatme and Sukhatme [20],

y: No. of villages in the circle.

x: A circle consisting more than five villages.

$N=89$, $n^{'}=30$, $n=20$, $\overline{Y}=3,360$, $\overline{X}=0,124$, $C_y=0,604$, $C_x=2,190$, $\rho_{yx}=0,766$.

Population III : Kadilar and Cingi [9],

y: Level of apple production.

x: No. of apple trees.

$N=104$, $n^{'}=40$, $n=20$, $\overline{Y}=625,37$, $\overline{X}=13,930$, $C_y=1,866$, $C_x=1,653$, $\rho_{yx}=0,865$.

Population IV : Murthy [12],

y: Output

x: fixed capital

$N=80$, $n^{'}=40$, $n=20$, $\overline{Y}=51,826$, $\overline{X}=11,265$, $C_y=0,354$, $C_x=0,751$, $\rho_{yx}=0,9413$.

We compute the MSE values of classical ratio and regression estimators, Singh and Vishwakarma [17] estimator and proposed estimator using the equations, (2.2), (2.4), (2.6), and (3.13), respectively. We have taken $a = b = 1$, that is, $\theta = \dfrac{\overline{X}}{2\left(\overline{X}+1\right)}$, just for the sake of simplicity.

These MSE values are shown in Table 2. We observe that the most efficient estimator is the proposed exponential regression cum ratio estimator as compared to those existing ones.

Table 2. MSE Values of Estimators in Two Phase Sampling

	Population			
Estimators	**I**	**II**	**III**	**IV**
Classical Ratio($\bar{y}_R$)	807,59	0,30	54993,75	12,64
Classical Regression($\bar{y}_{lr}$)	780,89	1,86	29536,17	16,87
Singh and Vishwakarma ($\bar{y}_{svr}$)	1045,59	0,40	35586,14	5,29
Proposed Est.($\bar{y}_{NH}$)	**774,71**	**0,12**	**26960,89**	**5,12**

6. Conclusion

We propose a class of regression cum estimator using the exponential function for the population mean in two phase sampling improving the exponential ratio estimator suggested in Singh and Vishwakarma [17]. Theoretically, we demonstrate that the proposed estimator is always the most efficient estimator in two phase sampling and numerically, for various specific data sets, we show that the proposed estimator has small MSE value according to other estimators. In future work, we will improve the proposed estimator, presented here, with using several auxiliary variables and adding more parameters for other sampling schemes.

References

[1] Ahmad, Z. *Generalized Multivariate Ratio and Regression Estimators for Multi-phase Sampling* (PhD thesis, National College of Business Administration and Economics, Lahore, Pakistan, 2008).

[2] Bahl, S. and Tuteja, R.K. *Ratio and product exponential estimators*, Journal of Information and Optimization Sciences, **12** (1), 159-164, 1991.

[3] Cingi, H., Kadilar, C. and Kocberber, G. *Determination of the Opportunities of the Primary and the Secondary Schools in Turkey and Suggestions to Solve the Determined Problems* (TUBITAK, SOBAG106K077.(in Turkish), 2007).

[4] Cingi, H. and Kadilar, C. *Advances in Sampling Theory-Ratio Method of Estimation* (Bentham Science Publishers, 2009).

[5] Das, A.K. and Tripathi, T.P. *Use of auxiliary information in estimating the finite population variance*, Sankhya C **40** (2): 139-148, 1978.

[6] Grover, L.K. and Kaur, P. *An improved estimator of the finite population mean in simple random sampling*, Model Assisted Statistics and Applications **6** (1), 47-55, 2011.

[7] Hanif, M., Shahbaz, M.Q. and Ahmad, Z. *Some Improved Estimators in Multiphase Sampling*, Pak. J. Statist. **26** (1), 195-202, 2010.

[8] Kadilar, C. and Cingi, H. *Ratio Estimators in Stratified Random Sampling*, Biometrical Journal **45** (2), 218-225, 2003.

[9] Kadilar, C. and Cingi, H. *An Improvement in Estimating The Population Mean By Using The Correlation Coefficient*, Hacettepe Journal of Mathematics and Statistics **35** (1), 103-109, 2006.
[10] Keen, K. J. *Two-Phase Sampling*, Encyclopedia of Biostatistics, 2005.
[11] Koyuncu, N. and Kadilar, C. *On the family of estimators of population mean in stratified random sampling*, Pak. Jour. Statist. **26** (2), 427-443, 2010.
[12] Murthy, M.N. *Sampling Theory and methods* (Statistical Publishing Society, Calcutta, India, 1967).
[13] Neyman, J. *Contribution to the theory of sampling human populations*, Journal of the American Statistical Association **33**, 101-116, 1938.
[14] Ozgul, N. *Mean Estimators in Two Phase Sampling* (MSc.Thesis (in Turkish), Department of Statistics, Hacettepe University, Ankara, Turkey, 2007).
[15] Samiuddin, M. and Hanif, M. *Estimation of population mean in single and two phase sampling with or without additional information*, Pak. J. Stat. **23** (2), 99-118, 2007.
[16] Singh, H.P. and Ruiz-Espejo, M. *On linear regression and ratio-product estimation of a finite population mean*, The Statistician **52** (1), 59-67, 2003.
[17] Singh, H.P. and Vishwakarma, G.K. *Modified exponential ratio and product estimators for finite population mean in two phase sampling*, Austrian Journal of Statistics **36** (3), 217-225, 2007.
[18] Singh, H.P., Tailor, R. and Tailor, R. O*n ratio and product methods with certain known population parameters of auxiliary variable in sample surveys*, SORT **34** (2), 157-180, 2010.
[19] Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. *Improvement in estimating the population mean using exponential estimator in simple random sampling*, Bulletin of Statistics and Economics **3** (9), 13-18, 2009.
[20] Sukhatme, P.V. and Sukhatme, B.V. *Sampling theory of surveys with applications (Second Edition)* (Asia Publishing House, Bombay, India, 1970).

Hacettepe Journal of Mathematics and Statistics
Volume 43 (1) (2014), 141 – 152

TYPE I ERROR RATE FOR TWO-SAMPLE TESTS IN STATISTICAL SHAPE ANALYSIS

Gokhan Ocakoglu * † and Ilker Ercan * ‡

Received 19 : 02 : 2013 : Accepted 25 : 06 : 2013

Abstract

Nowadays, with the help of advanced imaging techniques the image or shape of an organ or organism can be used as input data. Therefore, the statistical analysis of shape has recently become more important in the medical and biological sciences. Methods related to two-sample tests have been developed for statistical shape analysis, giving rise to considerable interest in research that evaluates the performance of these tests. In this study, two sample procedures are used to compare the mean shapes from the statistical shape analysis literature according to type I error rate.

Keywords: Statistical shape analysis, two-sample tests, type I error rate.

2000 AMS Classification:

1. Introduction

In the biological and medical sciences, morphometric methods are frequently preferred for examining the morphologic structures of organs or organisms with regard to diseases or environmental factors. Therefore, the statistical analysis of shapes has recently become more important in the medical and biological sciences. Data sets include qualitative and quantitative measurements for use in the statistical analyses associated with medical research. Nowadays, with the help of advanced imaging techniques the image or shape of an organ or organism can be used as input data [1].

Shape is defined as all the geometrical information that remains when location, scale and rotational effects are filtered from an object [2], [3], [4], [5]. Statistical shape analysis is a geometrical analysis of the statistics measured from sets of shapes that determines the features of similar shapes or of different groups comprising similar shapes. Distance between shapes, mean shape and shape variation can be predicted and obtained using statistical shape analysis [3]. A comparison of shapes between groups can also be done at a particular significance level.

*Department of Biostatistics, Faculty of Medicine, Uludag University, Bursa, Turkey.
†Corresponding author. Email: (G. Ocakoglu) ocakoglu@uludag.edu.tr
‡Email: (I. Ercan) ercan@uludag.edu.tr

Inferential methods described in the shape analysis literature make use of landmark configurations that are optimally superimposed via either a least-squares procedure or an analysis of interlandmark distance matrices [6].

Methods concerning two-sample tests have been developed for statistical shape analysis, giving rise to considerable interest in research that evaluates the performance of these tests. In this study, the Hotelling T^2, Goodall's F and James F_j tests as well as the $\lambda_{\min}$ test statistic are used to compare the mean shapes of two samples from the statistical shape analysis literature according to type I error rates derived from various variance values in different sample sizes. This simulation study considers both isotropic and anisotropic cases for which tangent space is used as shape space and considers methods that use complex arithmetic and exploit the geometry of the shape space.

2. Materials and Methods

2.1. Shape Space. The shape space is the set of all possible shapes [3]. For any set of landmarks $\{X_i\}$ in the original Euclidean plane, we can imagine the set of shapes derived by holding all but one of the X's at fixed position and varying that one in a circle about its original position. We would like the metric assigned to shape space (the set of "shapes" of all such sets of X's, correcting for centroid, orientation, and scale, all of which usually change whenever one of the X's moves) to be such that the shapes generated by circles in the original landmark plane are all at the same distance from the original shape $\{X_i\}$ in the shape space. That is, to a circle around one landmark in data space should correspond something very nearly a circle in shape space [7]. Although shape spaces defined by superimposition methods have less dimensions than raw data or non-redundant measurements, they are non-Euclidean and correspond to a curved surface. Nobody will recommend applying traditional statistics directly in this space because traditional statistics relies on the Euclidean metric, which is not the same as the Procrustes one [8]. Special statistical methods (rather than the usual linear multivariate methods) are required to take into account the non-Euclidean geometry of Kendall's shape space for both two and three-dimensional landmarks [4]. To perform usual statistical methods, one must first project the surface of the hyperhemisphere onto a "flat" tangent space where the Euclidean metrics allows us to use Euclidean statistics. The data are projected on a tangent shape space (also called Kendall tangent space or Kent tangent space). The contact between spaces is chosen as the mean shape. Working on variation in the tangent space is a rather perilous estimation since the projection can introduce distortion for the largest distances. However, provided that variation is small, one can assume that the portion of the shape hyperhemisphere and tangent space are nearly flat and nearly confused [8].

The projection onto a Euclidean space can be orthogonal or stereographic. Note that both projections will introduce biases for shapes being very different from the mean shape: the orthogonal projection minimizes large differences while stereographic projection accentuates them. The stereographic projection is produced by adjusting the size scale factor for the configuration to be projected onto the tangent space. To perform this projection, we use simple trigonometric relationships and divide the coordinates of the aligned configurations by the cosine of the Procrustes distance ρ between shapes and the mean shape [8].

In this study the performances of two-sample test procedures that examine differences in mean shape between two independent populations were evaluated in case of using tangent shape space as a shape space. For these test procedures the case in terms of using complex arithmetic and exploiting the geometry of the shape space which is an alternative computational method was also considered for examining tests performances.

2.2. Two-Sample Hotelling T^2 Test. The two-sample Hotelling T^2 test is used to test an alternative hypothesis related to the differences of the mean shapes of two groups and is accordingly applied to shape coordinates [9]. The Hotelling T^2 test assumes that the samples have multivariate normal distributions and equal variance-covariance matrices [10].

Consider two independent random samples $X_1, \ldots, X_{n1}$ and $Y_1, \ldots, Y_{n2}$ from two independent populations with mean shapes $[\mu_1]$ and $[\mu_2]$. To test the hypothesis $H_0 : [\mu_1] = [\mu_2]$, a two-sample Hotelling T^2 test can be performed in the Procrustes tangent space where the pole corresponds to overall pooled full Procrustes mean shape $\widehat{\mu}$. Let $v_1, \ldots, v_{n1}$ and $w_1, \ldots, w_{n2}$ be the partial Procrustes tangent coordinates (with pole $\widehat{\mu}$) [3].

A multivariate normal model is proposed in the tangent space, where $v_i \sim N(\xi_1, \sum_1)$ for $i = 1, \ldots, n_1$, $w_j \sim N(\xi_2, \sum_2)$ for $j = 1, \ldots, n_2$, and the v_i and w_j values are all mutually independent. $\bar{v}$ and $\bar{w}$ and S_v, S_w represent the sample means and sample covariance matrices respectively (with divisors n_1 and n_2) in each group. If the covariance matrices are assumed to be equal ($\sum_1 = \sum_2$), then the squared Mahalanobis distance between $\bar{v}$ and $\bar{w}$ is given by Equation-2.1.

$$D^2 = (\bar{v} - \bar{w})^T S_U^{+} (\bar{v} - \bar{w}) \tag{2.1}$$

where $S_U = (n_1 S_1 + n_2 S_2)/(n_1 + n_2 - 2)$ and S_U^{+} is the Moore-Penrose generalized inverse of S_U. Under the null hypothesis, we have $\xi_1 = \xi_2$ and the two-sample Hotelling statistic, which is given by Equation 2.2

$$F_H = \frac{n_1 n_2 (n_1 + n_2 - M - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)M} D^2 \tag{2.2}$$

where $M = 2d - 2$ is the dimension of the planar shape space. The test statistic has an $F_{M, n_1 + n_2 - M - 1}$ distribution under the null hypothesis [2], [3].

2.3. James F_j Test. When covariances are not assumed to be equal, an alternative method is to use the statistic proposed by James, which represents an effort to solve the multivariate Behrens-Fisher problem [2], [6].

$$F_j = (\bar{v} - \bar{w})^T \left(\frac{1}{n_1} S_v + \frac{1}{n_2} S_w \right)^{+} (\bar{v} - \bar{w}) \tag{2.3}$$

The J-statistic has an asymptotic χ^2_M distribution under the null hypothesis regardless of whether $\sum_1$ and $\sum_2$ are equal, and we reject the null hypothesis for large values of this statistic [2].

2.4. Two-Sample Goodall's F Test. Goodall presented a statistical framework for analyzing Procrustes shape data and developed a possible F test. This test is based on the Procrustes chord distance and should work under the assumption that variation is isotropic and is equal for each landmark [8]. This assumption implies that the variances of all landmarks (that is, the amount of dispersion) are expected to be the same. The assumption also implies that the patterns of dispersion across landmarks are expected to be uncorrelated [11].

If $\sum_1 = \sum_2 = \sum$ and we have isotropic covariance structure ($\sum = \sigma^2 I$) [2].

In an isotropic variance structure, the diagonal elements and the variance values of the covariance matrix are equal for each landmark, and all elements except the diagonal elements are equal to zero. Perhaps the simplest type of covariance structure for the perturbation distribution is one in which all landmarks are perturbed with the same variance irrespective of direction. This isotropic variance structure is easy to visualize, but may not be biologically realistic in the study of certain biological structures or certain populations [12]. An isotropic normal model with mean μ and transformed by an additional location, rotation and scale effects are given by Equation-2.4

$$x_i = \beta_i(\mu + E_i)\Gamma_i + 1_k\gamma_i^T \quad vec(E_i) \sim N(0, \sigma^2 I_{km}) \tag{2.4}$$

where $\beta_i > 0$ (scale), $\Gamma_i \in SO(m)$ (rotation) and $\gamma_i \in \mathbb{R}^m$ (translation), and σ is small.

Consider independent random samples $x_1, x_2, \ldots, x_n$ from a population modeled by Equation-2.4 with μ_1 and $y_1, y_2, \ldots, y_n$ from Equation-2.4 with mean μ_2. Both populations are assumed to have a common σ^2 variance for each coordinate [3].

We wish to test $H_0 : [\mu_1] = [\mu_2](= [\mu_0])$ against $H_1 : [\mu_1] \neq [\mu_2]$. $[\widehat{\mu}_1]$ and $[\widehat{\mu}_2]$ are the full Procrustes means of each sample. Under the H_0 hypothesis, with a small σ the Procrustes distances are approximately distributed as

$$\sum_{i=1}^{n_1} d_F^2(X_i, \widehat{\mu}_1) \sim \tau_0^2 \chi^2_{(n_1-1)M} \tag{2.5}$$

$$\sum_{i=1}^{n_2} d_F^2(Y_i, \widehat{\mu}_2) \sim \tau_0^2 \chi^2_{(n_2-1)M} \tag{2.6}$$

$$d_F^2(\widehat{\mu}_1, \widehat{\mu}_2) \sim \tau_0^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \chi^2_M \tag{2.7}$$

where $\tau = \sigma/\delta$, $\delta_0 = S(\mu_0)$ and d_F^2 represents the squared full Procrustes distance between two configurations. In addition, these statistics are approximately mutually independent [3]. Hence, under the null hypothesis, we have the approximate distribution as given in equation-2.8.

$$F_G = \frac{n_1 + n_2 - 2}{n_1^{-1} + n_2^{-1}} \frac{d_F^2(\widehat{\mu}_1, \widehat{\mu}_2)}{\sum_{i=1}^{n_1} d_F^2(X_i, \widehat{\mu}_1) + \sum_{i=1}^{n_2} d_F^2(Y_i, \widehat{\mu}_2)} \sim F_{M,(n_1+n_2-2)M} \tag{2.8}$$

We reject the null hypothesis for large values of this test statistic. The Hotelling T^2 procedure is less powerful than Goodall's F test, for which the isotropic normal model holds [3], [13].

2.5. $\lambda_{\min}$ Test Statistic. Amaral et al. [2] proposed a novel bootstrap approach to k-sample testing problems in which each sample consists of a set of real or complex unit vectors. The basic assumption is that the distribution of the sample mean shape (or direction or axis) is highly concentrated [6]. Consider k samples of unit vectors in $\mathbb{C}^d$ (in most traditional applications, $d = 2; 3$, but sometimes the case $d \geq 4$ is also relevant), and let $\widehat{\mu}_i$ be the estimator of μ_0 (i.e., the mean shape under the hypothesis) based on sample i, for $i = 1, \ldots, k$. Assume that $n^{\frac{1}{2}} \widehat{M}_i \mu_0 \xrightarrow{D} \mathbb{C}N_{d-1}(0, G_i)$ for $i = 1, \ldots, k$ where G_i denotes asymptotic covariance matrix has full rank and $\widehat{M}_i$ represents a projection onto the tangent space at $\widehat{\mu}_i$ [6].

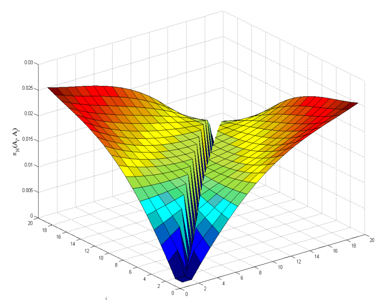
Define $\widehat{A}_0 = n \sum_{i=1}^{k} \widehat{M}_i^T \widehat{G}_i^{-1} \widehat{M}_i$ and $T_0(\mu) = 2\mu^T \widehat{A}_0 \mu$, where T denotes the conjugate transpose, μ is a complex unit vector and $\widehat{G}_i$ is a consistent estimator of G_i. We thus obtain

$$\lambda_{\min} \equiv \min_{\mu:\|\mu\|=1} T_0(\mu) = T_0(\widehat{\mu}) \tag{2.9}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\widehat{A}_0$ and $\widehat{\mu}_0$ is the corresponding unit eigenvector [2], [6]. It is proven that $\lambda_{\min} \xrightarrow{D} \chi^2_{2(k-1)(d-1)}$ as $n \to \infty$ under the null hypothesis of equality of means across populations [6].

2.6. A Simulation Study. In this study we aim to compare type I error rates of the tabular, bootstrap and permutation adaptations of Hotelling T^2, Goodall's F and James F_j tests as well as the $\lambda_{\min}$ test statistic. A mean vector and a variance-covariance matrix are computed from a data set obtained from the landmark markings of the nose in the anterior views of the faces of 50 subjects. Eleven landmarks (Figure 1) are applied to the images in the manner described by Ercan et al. [14]. In the present study, the data are simulated from a multivariate normal distribution under isotropic and anisotropic models.

FIGURE 1. Landmark markings for the source data set used in the simulation study.



The samples for which type I error rates are examined in the simulation study are $n_1 = n_2 = 20, 50, 100$ and 500.

A mean vector that computed from a data set obtained from the landmark markings as mentioned above is $(\bar{x}_1, \ldots, \bar{x}_{11}, \bar{y}_1, \ldots, \bar{y}_{11})$ = (501, 590, 546, 522, 568, 546, 521, 570, 532, 563, 547, 399, 398, 384, 398, 397, 409, 425, 426, 469, 469, 500).

Variance values are determined to be $0.001, 0.01, 0.05, 0.1, 0.5, 1, 5, 737, 1703$ and 2949 in the isotropic case. The values $737, 1703$ and 2949 values are the minimum, maximum and mean variance values of the variance-covariance matrix, which contains real values from the sample data set.

Isotropic structures are used in studies and when comparing the methods; however it is not the case as in real-world applications; therefore, in our study we also compare methods by simulating with anisotropic structures. The real variance-covariance matrix computed from the sample data set is used as input for the simulation of the anisotropic case.

In the examination of type I error rate in the simulation study, it is assumed that related tests use tangent space as shape space, that they use complex arithmetic and that they exploit the geometry of shape space.

The simulation study has been conducted with 1000 replications, and the number of bootstrap and permutation resamples is set to 100.

We used TPSDIG 2.04 software to mark the landmarks on the images. The simulation study and analyses were performed using R 2.12.0 software [15].

3. Results and Discussion

In Table 1, we give type I error rates as determined for both cases according to the exploitation of shape space, according to various variance values for the isotropic model and according to the variance-covariance matrix computed from the real data set for the anisotropic model in different sample sizes.

It has been observed that applications of statistical shape analysis have recently been used more than ever before in medical and biological sciences to compare the structures of shapes [14], [16], [17], [18]. For example, forensics analyses [19], computer-assisted neurosurgery methods [20] anthropological studies [14], [17], [18], [21] and MRI-based morphological analyses of the brain [22], [23], [24] make use of statistical shape analysis. Therefore, it is of great importance that shape objects be recognized, measured and compared.

Newly developed methods utilize two-sample tests in statistical shape analysis, which is a geometric morphometric concept. However, more emphasis has been placed on studies of the comparative performance of related tests. In this study, we aim to compare the type I error rates of the Hotelling T^2, Goodall's F and James F_j tests as well as the $\lambda_{\min}$ test statistic, which are all used in the shape analysis literature to compare mean shapes. In this simulation study, the performance of tabular, bootstrap and permutation adaptations of the related procedures are examined in terms of type I error rate. We also consider isotropic and anisotropic cases for different variance values and sample sizes using the tangent space as the shape space. Finally, we consider related procedures that use complex arithmetic and exploit the geometry of the shape space.

We examined the procedures of bootstrap adaptations through simulation results, considered isotropic covariance structure, exploited tangent space and used complex arithmetic with the geometry of the shape space, thus evaluating small samples. In light of these findings, the application of the Hotelling T^2, James F_j and Goodall's F tests in tangent space put the type I error rate under the determined nominal level. Additionally, we observe that the type I error rates remained under the nominal level following the application of $\lambda_{\min}$ test statistic with the Hotelling T^2, Goodall's F and James F_j tests when complex arithmetic was applied and the geometry of the shape space was exploited. In a similar study of small samples, Brombin and Salmaso [6] conducted the Hotelling T^2, Goodall's F and James F_j tests and generally found that the type I error rate was under the nominal level in the isotropic covariance structure when using complex arithmetic with the geometry of the shape space. Brombin and Salmaso [6] also observed a value close to the determined nominal level when using the $\lambda_{\min}$ test statistic. Amaral et al. [2] carried out a similar study with small samples and observed a value close to

Table-1: Type I error rates for $n_1=n_2=20, 50, 100, 500$ and $\sigma^2= 0.001, 0.01, 0.05, 0.1$ in the case of using shape space as tangent space and exploiting complex arithmetic with geometry of shape space.

		$\sigma^2 = 0.001$				$\sigma^2 = 0.01$				$\sigma^2 = 0.05$				$\sigma^2 = 0.1$			
		n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500
Using Tangent space as shape space	H_bootstrap	0.000	0.024	0.051	0.049	0.000	0.033	0.037	0.055	0.000	0.021	0.036	0.066	0.000	0.027	0.040	0.048
	H_permutation	0.047	0.042	0.057	0.048	0.037	0.045	0.054	0.056	0.048	0.040	0.046	0.057	0.062	0.046	0.042	0.050
	H_tabular	0.018	0.024	0.044	0.044	0.032	0.041	0.037	0.056	0.047	0.042	0.044	0.061	0.052	0.052	0.043	0.043
	G_ bootstrap	0.017	0.020	0.047	0.041	0.019	0.036	0.033	0.058	0.019	0.038	0.036	0.053	0.028	0.041	0.038	0.041
	G_ permutation	0.046	0.032	0.050	0.041	0.051	0.050	0.043	0.057	0.045	0.047	0.045	0.063	0.060	0.053	0.057	0.051
	G _ tabular	0.044	0.030	0.044	0.035	0.053	0.048	0.037	0.059	0.054	0.049	0.041	0.056	0.059	0.052	0.052	0.042
	J_ bootstrap	0.000	0.024	0.051	0.049	0.000	0.033	0.037	0.055	0.000	0.021	0.036	0.066	0.000	0.027	0.040	0.048
	J_ permutation	0.047	0.032	0.057	0.048	0.037	0.045	0.054	0.056	0.048	0.040	0.046	0.057	0.062	0.046	0.042	0.050
	J_ tabular	0.120	0.035	0.053	0.044	0.167	0.051	0.042	0.056	0.191	0.062	0.046	0.062	0.228	0.066	0.052	0.045
Using complex arithmetic and exploit geometry of shape space	H_ bootstrap	0.000	0.012	0.047	0.049	0.000	0.032	0.038	0.059	0.000	0.022	0.039	0.054	0.000	0.031	0.041	0.046
	H_ permutation	0.050	0.034	0.043	0.042	0.043	0.051	0.044	0.056	0.047	0.048	0.038	0.061	0.051	0.054	0.046	0.046
	H_ tabular	0.044	0.033	0.059	0.045	0.041	0.048	0.042	0.060	0.047	0.042	0.044	0.061	0.052	0.052	0.043	0.043
	G_ bootstrap	0.020	0.017	0.040	0.039	0.030	0.042	0.036	0.056	0.019	0.036	0.042	0.050	0.026	0.042	0.042	0.043
	G_ permutation	0.050	0.037	0.053	0.039	0.050	0.050	0.043	0.060	0.047	0.044	0.044	0.058	0.048	0.050	0.055	0.047
	G _ tabular	0.054	0.032	0.051	0.039	0.055	0.049	0.042	0.060	0.053	0.049	0.041	0.056	0.059	0.052	0.052	0.042
	J_ bootstrap	0.000	0.012	0.047	0.049	0.000	0.032	0.038	0.059	0.000	0.022	0.039	0.054	0.000	0.031	0.041	0.046
	J_ permutation	0.050	0.034	0.043	0.042	0.043	0.051	0.044	0.056	0.047	0.048	0.038	0.061	0.051	0.054	0.046	0.046
	J_ tabular	0.197	0.066	0.063	0.046	0.198	0.067	0.049	0.060	0.191	0.062	0.046	0.062	0.228	0.066	0.052	0.045
	λ_{min}_ bootstrap	0.045	0.029	0.044	0.041	0.056	0.058	0.043	0.059	0.049	0.046	0.038	0.053	0.072	0.038	0.045	0.043
	λ_{min}_ permutation	0.052	0.033	0.051	0.037	0.050	0.057	0.038	0.058	0.044	0.051	0.052	0.060	0.056	0.050	0.051	0.049
	λ_{min}_ tabular	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.000	0.108	0.112	0.076	0.066	0.286	0.106	0.076	0.047

Table-1 (continued): Type I error rates for $n_1=n_2=20, 50, 100, 500$ and $\sigma^2= 0.5, 1, 5, 737$ in the case of using shape space as tangent space and exploiting complex arithmetic with geometry of shape space.

		$\sigma^2=0.5$				$\sigma^2=1$				$\sigma^2=5$				$\sigma^2=737$			
		n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500
Using Tangent space as shape space	H_bootstrap	0.000	0.023	0.041	0.043	0.000	0.028	0.045	0.052	0.000	0.024	0.039	0.054	0.000	0.023	0.040	0.047
	H_permutation	0.045	0.054	0.047	0.048	0.049	0.045	0.055	0.045	0.054	0.044	0.045	0.048	0.047	0.049	0.043	0.046
	H_tabular	0.049	0.046	0.055	0.051	0.047	0.049	0.058	0.048	0.052	0.041	0.045	0.054	0.046	0.049	0.047	0.048
	G_ bootstrap	0.022	0.028	0.042	0.046	0.019	0.040	0.044	0.047	0.020	0.034	0.044	0.054	0.021	0.035	0.038	0.048
	G_ permutation	0.054	0.047	0.046	0.046	0.060	0.050	0.055	0.042	0.049	0.047	0.049	0.050	0.047	0.045	0.037	0.052
	G _ tabular	0.046	0.045	0.048	0.049	0.057	0.050	0.058	0.044	0.046	0.046	0.049	0.053	0.046	0.051	0.041	0.049
	J_ bootstrap	0.000	0.023	0.041	0.043	0.000	0.028	0.045	0.052	0.000	0.024	0.039	0.054	0.000	0.023	0.040	0.047
	J_ permutation	0.045	0.054	0.047	0.048	0.049	0.045	0.055	0.045	0.054	0.044	0.045	0.048	0.047	0.049	0.043	0.046
	J_ tabular	0.222	0.066	0.056	0.053	0.207	0.066	0.065	0.050	0.232	0.048	0.051	0.055	0.188	0.070	0.053	0.048
Using complex arithmetic and exploit geometry of shape space	H_ bootstrap	0.000	0.023	0.051	0.047	0.000	0.022	0.045	0.047	0.000	0.028	0.042	0.059	0.000	0.031	0.045	0.044
	H_ permutation	0.052	0.045	0.055	0.053	0.051	0.046	0.059	0.050	0.056	0.041	0.047	0.054	0.050	0.054	0.048	0.053
	H_ tabular	0.049	0.046	0.055	0.051	0.047	0.049	0.058	0.048	0.052	0.041	0.045	0.054	0.047	0.052	0.048	0.047
	G_ bootstrap	0.022	0.029	0.046	0.044	0.024	0.029	0.048	0.045	0.023	0.038	0.038	0.048	0.008	0.039	0.035	0.044
	G_ permutation	0.049	0.042	0.058	0.050	0.053	0.044	0.054	0.045	0.042	0.051	0.048	0.053	0.053	0.052	0.046	0.051
	G _ tabular	0.046	0.045	0.048	0.049	0.057	0.050	0.059	0.044	0.048	0.048	0.049	0.053	0.486	0.524	0.548	0.568
	J_ bootstrap	0.000	0.023	0.051	0.047	0.000	0.022	0.045	0.047	0.000	0.028	0.042	0.059	0.000	0.031	0.045	0.044
	J_ permutation	0.052	0.045	0.055	0.053	0.051	0.046	0.059	0.050	0.056	0.041	0.047	0.054	0.050	0.054	0.048	0.053
	J_ tabular	0.222	0.066	0.056	0.053	0.207	0.066	0.065	0.050	0.232	0.058	0.051	0.055	0.185	0.071	0.053	0.047
	λ_{min}_ bootstrap	0.009	0.026	0.047	0.048	0.013	0.036	0.046	0.046	0.010	0.029	0.038	0.054	0.004	0.039	0.042	0.041
	λ_{min}_ permutation	0.042	0.038	0.052	0.059	0.057	0.045	0.044	0.048	0.049	0.048	0.043	0.053	0.047	0.050	0.043	0.049
	λ_{min}_ tabular	0.281	0.117	0.078	0.055	0.264	0.110	0.087	0.054	0.294	0.109	0.075	0.057	0.240	0.109	0.073	0.048

Table-1 (continued): Type I error rates for n_1=n_2=20, 50, 100, 500 and σ^2= 1703, 2949 and anisotropic covariance structure in the case of using shape space as tangent space and exploiting complex arithmetic with geometry of shape space.

		$\sigma^2 = 1703$				$\sigma^2 = 2949$				Anisotropic covariance structure			
		n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500	n=20	n=50	n=100	n=500
Using Tangent space as shape space	H_bootstrap	0.000	0.034	0.037	0.050	0.000	0.022	0.034	0.046	0.000	0.016	0.038	0.045
	H_permutation	0.053	0.052	0.043	0.050	0.049	0.042	0.044	0.052	0.060	0.048	0.050	0.047
	H_tabular	0.054	0.054	0.047	0.054	0.053	0.042	0.041	0.053	0.046	0.049	0.050	0.045
	G_ bootstrap	0.030	0.040	0.033	0.047	0.024	0.041	0.027	0.043	0.052	0.041	0.051	0.051
	G_ permutation	0.057	0.055	0.041	0.053	0.051	0.052	0.038	0.055	0.061	0.046	0.051	0.050
	G _ tabular	0.057	0.054	0.044	0.054	0.053	0.047	0.033	0.048	0.169	0.140	0.122	0.154
	J_ bootstrap	0.000	0.034	0.037	0.050	0.000	0.022	0.034	0.046	0.000	0.016	0.038	0.045
	J_ permutation	0.053	0.052	0.043	0.050	0.049	0.042	0.044	0.052	0.060	0.048	0.050	0.047
	J_ tabular	0.214	0.073	0.055	0.055	0.212	0.066	0.046	0.054	0.207	0.064	0.055	0.050
Using complex arithmetic and exploit geometry of shape space	H_ bootstrap	0.000	0.033	0.031	0.036	0.000	0.016	0.018	0.028	0.000	0.022	0.039	0.005
	H_ permutation	0.054	0.059	0.045	0.053	0.051	0.046	0.037	0.046	0.051	0.056	0.050	0.042
	H_ tabular	0.053	0.055	0.050	0.051	0.056	0.043	0.041	0.053	0.046	0.049	0.050	0.045
	G_ bootstrap	0.011	0.038	0.033	0.054	0.004	0.021	0.027	0.047	0.047	0.042	0.050	0.049
	G_ permutation	0.093	0.065	0.059	0.050	0.146	0.077	0.046	0.051	0.069	0.047	0.053	0.054
	G _ tabular	0.914	0.931	0.927	0.936	0.988	0.992	0.995	0.997	0.168	0.145	0.122	0.151
	J_ bootstrap	0.000	0.033	0.031	0.036	0.000	0.016	0.018	0.028	0.000	0.022	0.039	0.045
	J_ permutation	0.054	0.059	0.045	0.053	0.051	0.046	0.037	0.046	0.051	0.056	0.050	0.042
	J_ tabular	0.218	0.075	0.053	0.054	0.216	0.067	0.044	0.055	0.207	0.064	0.055	0.047
	λ_{min}_ bootstrap	0.006	0.040	0.036	0.051	0.003	0.023	0.028	0.048	0.016	0.040	0.048	0.040
	λ_{min}_ permutation	0.059	0.055	0.037	0.044	0.050	0.051	0.035	0.045	0.061	0.055	0.051	0.051
	λ_{min}_ tabular	0.243	0.110	0.062	0.058	0.239	0.116	0.064	0.056	0.315	0.146	0.117	0.093

In Table-1, H indicates Hotelling T^2 test, G indicates Goodall's F test and J indicates James F_J test respectively.

the determined level in terms of type I error rates in related procedures. As for large samples, while the type I error rates converged to the nominal level in both usages of shape space, we found results under the nominal level in the simulation study of high variance values.

In the simulation study in which we exploited the variance-covariance matrix of real landmark values, the anisotropic covariance structure and the procedures of bootstrap adaptations, we found that type I error rates stayed under the nominal level according to the Hotelling T^2, Goodall's F and James F_j tests as well as the $\lambda_{\min}$ test statistic for both usages of shape space in small samples. When large samples were evaluated, we found that the type I error rates remained under the determined nominal level only when the Hotelling T^2 test was applied in the case of exploiting complex arithmetic with the geometry of the shape space.

Following the examination of the permutation adaptation of procedures through the simulation results and considering the isotropic covariance structure, the tests showed an overall performance in all sample sizes in both usages of shape space. However, the Goodall's F test tends to overestimate the nominal level in small samples in the case of exploiting complex arithmetic with the geometry of the shape space. In a similar study of small samples, Amaral et al. [2] found an overall results that were close the nominal level for the type I error rates; however, Amaral et al. [2] reported that as the variance values in the Goodall's F test increased, the related procedure tended to overestimate the nominal level of the type I error rate. Compared to the variance values in Amaral et al. [2], the variance values of the Goodall's F test are close to the values of the nominal level of the type I error rate. Brombin and Salmaso [6] stated that the Hotelling T^2 and James F_j tests showed similar values but that the Goodall's F test and the $\lambda_{\min}$ test statistic tended to underestimate the nominal level. In the anisotropic covariance structure, the examined procedures showed similar results to the nominal type I error rate in small and large sample sizes.

When tabular versions of procedures were analyzed through simulation results, the James F_j test tended to overestimate the nominal level in small samples in both usages of shape space in the case of isotropic covariance structure. The Hotelling T^2 test underestimated the nominal level in small samples in tangent space with reference to type I error rate in low variance values, but the Goodall's F test overestimated the nominal level in the case of exploiting complex arithmetic with the geometry of the shape space in high variance values. We found that comparison with the $\lambda_{\min}$ test statistic generally underestimated and overestimated the nominal level. We found that the James F_j and Goodall's F tests as well as the $\lambda_{\min}$ test statistic underestimated and overestimated the nominal level; on the other hand, the Hotelling T^2 test revealed values close to the nominal level, which Brombin and Salmaso [6] also observed in a similar study of small samples in the case of exploiting isotropic covariance structure and in the cases of related procedures that use complex arithmetic and exploit the geometry of the shape space. Amaral et al. [2] also found that the Goodall's F test and the $\lambda_{\min}$ test statistic overestimated the nominal level; however, the Hotelling T^2 and James F_j tests resulted in values close to the nominal level in a similar study of small samples. The Goodall's F test overestimated the nominal level in large samples when exploiting complex arithmetic with the geometry of the shape space and in the case of high variance values. It was observed that the Goodall's F test and $\lambda_{\min}$ test statistic overestimated the nominal level in both usages of shape space in anisotropic covariance structure.

When the present study is compared with the similar studies [2], [6] in the literature, performances of two-sample test procedures used in this study were examined in terms of both using tangent space as a shape space and using complex arithmetic with exploiting the geometry of shape space. This study also differs from other literatures in terms of using variance-covariance matrix of real-life data set to examine the performances of related procedures in anisotropic case. In addition, it has been observed that the variance values given in simulation scenarios in similar studies are smaller than the variance values of real-life data sets. For this reason, in this study two-sample test procedures' performances were also examined for large variance values computed from a real-life data set. Present study also differs in terms of including large sample size values.

4. Conclusions

As predicted, the results of the present study indicate that tests perform better with large samples than with small samples. For small samples, permutation test adaptations gave the most favorable results in all isotropic and anisotropic covariance structures. For large samples, permutation test adaptations gave the most favorable results with regard to type I error rate in all low and high variance values and in all isotropic and anisotropic covariance structures. It was concluded that bootstrap adaptations of tests gave the most unfavorable results in all isotropic and anisotropic covariance structures in small samples.

Conflict of Interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by Uludag University Scientific Research Project. Project number: (UAP(T)-2010/34). The numerical calculations reported in this paper were performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TR-Grid e-Infrastructure).

References

[1] Ercan, I., Ocakoglu, G., Sigirli, D. and Ozkaya, G. *Statistical shape analysis and usage in medical sciences*, Turkiye Klinikleri J Biostat. 4(1), 27-35 (2012).
[2] Amaral, G.J.A, Dryden, I. and Wood, A.T.A. *Pivotal bootstrap methods for k-sample problems in directional statistics and shape analysis*, Journal of the American Statistical Association. 102(478), 695-707 (2007).
[3] Dryden, I.L. and Mardia, K.V. Statistical shape analysis (Vol. 4). Wiley, New York, 1998.
[4] Rohlf, F.J. *Shape statistics. Procrustes superimpositions and tangent spaces*, Journal of Classification. 16(2), 197-223, 1999.
[5] Rohlf, F.J. *On the use of shape spaces to compare morphometric methods*, Hystrix-the Italian Journal of Mammalogy, 11(1), 12-13, 2000.
[6] Brombin, C. and Salmaso, L. *Multi-aspect permutation tests in shape analysis with small sample size*, Computational Statistics & Data Analysis, 53(12), 3921-3931, 2009.
[7] Bookstein, F. L. Morphometric tools for landmark data: Geometry and Biology. Cambridge Univ. Press, New York, 1991.
[8] Claude, J. Morphometrics with R. Springer Verlag, 2008.
[9] Zelditch, M., Swiderski, D.L., Sheets, H.D. and Fink, W.L. Geometric morphometrics for biologists. A primer. Academic Press, 2004.
[10] Hammer, Ø. and Harper, D.A.T. Paleontological data analysis. Malden (Massachusetts), Blackwell Publishing Ltd Oxford, 2005.
[11] Richtsmeier, J.T., Lele, S. and Cole, T. *Landmark morphometrics and the analysis of variation*, In. Hallgrimsson, B, Hall, BK(eds.)Variation. A Central Concept in Biology. Amsterdam. Elsevier, 2005.
[12] Lele, S. and Richtsmeier, J.T. *An invariant approach to statistical analysis of shapes*, Florida. Chapman & Hall, 2001.
[13] Goodall, C. *Procrustes methods in the statistical analysis of shape*, Journal of the Royal Statistical Society. Series B (Methodological), 53(2), 285-339, 1991.

[14] Ercan, I., Etoz, A., Guney, I., Ocakoglu, G., Ozdemir, S.T., Kan, I. and Kahveci, R. *Statistical shape analysis of nose in Turkish young adults*, Journal of Craniofacial Surgery, 18(1), 219, 2007.

[15] Dryden, I. L. Shapes package. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[16] Ercan, I., Ocakoglu, G., Guney, I. and Yazici, B. *Adaptation of generalizability theory for inter-rater reliability for landmark localization*, International Journal of Tomography & Statistics, 9(S08), 51-58, 2008.

[17] Ercan, I., Ozdemir, S.T, Etoz, A., Sigirli, D., Tubbs, R.S., Loukas, M. and Guney, I. *Facial asymmetry in young healthy subjects evaluated by statistical shape analysis*, Journal of Anatomy, 213(6), 663-669, 2008.

[18] Ozdemir, S.T., Ercan, I., Ozkaya, G., Cankur, N.S. and Erdal Y.S. *Geometric Morphometric Study and Cluster Analysis of Late Byzantine and Modern Human Crania*, Collegium Antropologicum, 34(2), 493-499, 2010.

[19] De Greef, S., Claes, P., Vandermeulen, D., Mollemans, W., Suetens, P. and Willems, G. *Large-scale in-vivo Caucasian facial soft tissue thickness database for craniofacial reconstruction*, Forensic science international, 159, S126-S146, 2006.

[20] Aldridge, K., Marsh, J.L., Govier, D. and Richtsmeier, J.T. *Central nervous system phenotypes in craniosynostosis*, Journal of Anatomy, 201(1), 31-39, 2002.

[21] Jonke, E., Schaefer, K., Freudenthaler, J.W., Prossinger, H. and Bookstein, F.L. *A cephalometric comparison of skulls from different time periods-the bronze age, the 19th century and the present*, Collegium antropologicum, 27(2), 789-802, 2003.

[22] Bookstein, F.L., Sampson, P.D., Streissguth, A.P. and Connor, P.D. *Geometric morphometrics of corpus callosum and subcortical structures in the fetal-alcohol-affected brain*, Teratology. 64(1).4-32, 2001.

[23] Colak, C., Ercan, I., Dogan, M., Ozdemir, S.T., Sener, S. and Alkan, A. *Detecting the shape differences of the corpus callosum in behcet's disease by statistical shape analysis*, Anatomical Record-Advances in Integrative Anatomy and Evolutionary Biology, 294(5), 870-874, 2011.

[24] Ozdemir, S.T., Ercan, I., Sevinc, O., Guney, I., Ocakoglu, G., Aslan, E. and Barut, C. *Statistical shape analysis of differences in the shape of the corpus callosum between genders The Anatomical Record*, Advances in Integrative Anatomy and Evolutionary Biology, 290(7), 825-830, 2007.

INSTRUCTIONS FOR 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: Editor: `+ 90 312 297 7898`
Associate Editor: `+ 90 312 297 7880`
Fax : `+ 90 312 299 2017`
Web Page : `http//www.mat.hacettepe.edu.tr/hjms/`

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. 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 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).

Submission of Manuscripts : Manuscript will be typeset using the LaTeX typesetting system. Authors are invited to submit a LaTeX file of their paper by e-mail, or if this is not possible to send three copies of their article by post. In the latter case the article should be printed on one side only of white A4 size paper. Files should be prepared using the LaTeX article style or the AMS-TeX preprint style. Authors are requested to use only standard LaTeX or AMS-TeX commands and environments. Figures and other artwork should be submitted in electronic form wherever possible. Postscript (eps) or Windows metafile (wmf) are the preferred formats, but a separate Word file of each figure may be suitable. More detailed information, and a style-file for preparing the final form of accepted articles, will soon be available from the HJMS Web Page.

Authors who are unable to produce a suitable LaTeX or AMS-TeX file may submit three wordprocessor or typewritten copies of their article by post, printed with wide margins on one side only of A4 size white paper. If accepted for publication the article will be typeset in LaTeX by the Journal. Authors should be aware, however, that submitting a suitable LaTeX file will substantially shorten the time taken to process their paper.

While great care will be taken of manuscripts and files, no responsibility can be accepted for any loss. In their own interest, authors are advised to keep a copy of everything they send.

Deadline for Submissions : Manuscripts may be submitted at any time.

Page Charges : There are **no** page charges for articles published in HJMS.

Offprints : Twenty four offprints of each article will be supplied free of charge. Additional offprints may be ordered prior to publication and will be charged at cost.

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 or 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.