Online
ISSN 1346-0447

ISSN 1346-0862

Scientiae Mathematicae Japonicae

(Scientiae Mathematicae / Mathematica Japonica New Series)

Vol. 78, No. 1 Whole Number 278

June 2015

International Society for Mathematical Sciences

INTERNATIONAL SOCIETY FOR MATHEMATICAL SCIENCES
Scientiae Mathematicae Japonicae, Notices from the ISMS

The International Society for Mathematical Sciences (ISMS) is an international society consisting of mathematical scientists throughout the world.

The main activities of the ISMS are to publish (1) the (**print and online**) journal **Scientiae Mathematicae Japonicae (SCMJ)** and (2) Notices from the ISMS and to hold **assembly meeetings** in Japan and **international internet meetings** (distance symposium) of mathematical sciences (IVMS) accessible from all over the world.

SCMJ is the 21st Century New Unified Series of **Mathematica Japonica (MJ)** and **Scientiae Mathematicae (SCM). MJ** was first published in 1948 and was one of the oldest mathematical journals in Japan. **SCM** was an online and print journal started in 1998 in celebration of the semi-centurial anniversary and received 26000 visits per month from 50 countries in the world. **SCMJ** contains original papers in mathematical sciences submitted from all over the world and receives **38000** visits per month now. Not only papers in pure and applied mathematics but those devoted to mathematical statistics, operations research, informatics, computer science, biomathematics, mathematical economics and other mathematical sciences are also welcome. The journal is published in January, March, May, July, September, and November in each calendar year.

The ISMS has enhanced the journal, begining from July 1995, by including excellent Research-Expository papers in the section "**International Plaza** for Mathematical Sciences " as well as original research papers. The section provides papers dealing with broad overviews of contemporary mathmatical sciences, written by experts mainly at our invitation. Papers shedding lights on open problems or new directions or new breakthroughs for future research are especially welcome.

As is shown in the **Editorial Board of SCMJ**, we have invited **many distinguished professors** of 20 countries as editors, who will receive and referee the papers of their special fields with their **high standard**.

Beginning from 2007, we make the online version of SCMJ more readable and convenient to the readers by adding the specialized contents. By this, the readers can access to the online version, in which the papers appear in the order of acceptance, from (i) the contents of the printed version, and (ii) the specialized contents of a volume. From 2007, the subscription fee of the printed version plus the online version of SCMJ becomes lower and the same of the printed version only. Therefore, the subscribers of the printed version can read the online version without no additional cost.

For benefit of the ISMS members, we publish "Notices from the ISMS" 6 times a year. We are enhancing it by adding interesting articles, including book reviewing, written by eminent professors.

The ISMS has set up a **videoconferencing system** (**IVMS**) which can connect up to twenty sites of a reserch group in the same or **different countries in the world**. Using this system, speakers of the session can write on a white board or an OHP sheet or use PowerPoint. On the other hand participants can ask questions or make comments from any connected site in the world. All these are performed similarly to the traditional meetings.

To connect with our system, you can use your own videoconferencing system only if it satisfies the International Telecommunication Union-Technical Committee Standards (ITU-T Standard).

Copyright Transfer Agreement

A copyright transfer agreement is required before a paper is published in this journal. By submitting a paper to this journal, authors are regarded to certify that the manuscript has not been submitted to nor is it under consideration for publication by another journal, conference proceedings or similar publication.

For more information, please visit http://www.jams.or.jp.

Scientiae Mathematicae Japonicae, **78**, No. 1 (2015), 1–6

ON 0-MINIMAL IDEALS IN A DUAL ORDERED SEMIGROUP WITH ZERO

Thawhat Changphas, Wichayaporn Jantanan

Received March 5, 2013 ; revised December 18, 2013

Abstract. An ordered semigroup S is called a *dual ordered semigroup* if $l(r(L)) = L$ for every left ideal L of S and $r(l(R)) = R$ for every right ideal R of S where $r(A)$ and $l(A)$ denoted the *right annihilator* and the *left annihilator* of a nonempty subset A of S, respectively. The main result of this paper is to show the existence of 0-minimal ideals of a dual ordered semigroup.

1 Preliminaries Dual ring credited to Baer [1] and Kaplansky [8] have been widely studied (see [3], [5], [4], [9]). Using only the multiplication properties of the elements of a ring, Schwarz ([10], [11]) introduced and studied dual semigroups. Let S be a semigroup with zero 0 and let A be a nonempty subset of S. The *left annihilator* of A, denoted by $l(A)$, is defined by $l(A) = \{x \in S \mid xA = \{0\}\}$. Dually, the *right annihilator* of A, denoted by $r(A)$, is defined by $r(A) = \{x \in S \mid Ax = \{0\}\}$. The semigroup S is said to be *dual* if $l(r(L)) = L$ for all left ideals L of S and $r(l(R)) = R$ for all right ideals R of S. In [11], the author proved the existence of 0-minimal ideals of a dual semigroup. The purpose of this paper is to extend the results to ordered semigroups.

A semigroup $(S, \cdot)$ together with a partial order $\leq$ on S that is *compatible* with the semigroup operation, meaning that for $x, y, z \in S$,

$$x \leq y \Rightarrow zx \leq zy,\ xz \leq yz,$$

is called an *ordered semigroup* ([2], [4]). If A, B are nonempty subsets of S, we let

$$\begin{aligned} AB &= \{xy \in S \mid x \in A, y \in B\}, \\ (A] &= \{x \in S \mid x \leq a \text{ for some } a \in A\}. \end{aligned}$$

If $x \in S$, then we write Ax and xA instead of $A\{x\}$ and $\{x\}A$, respectively.

If A, B are non-empty subsets of an ordered semigroup $(S, \cdot, \leq)$, then it was proved in [6] that the following conditions hold:

(1) $A \subseteq (A]$;

(2) $A \subseteq B \Rightarrow (A] \subseteq (B]$;

(3) $((A]] = (A]$;

(4) $(A](B] \subseteq (AB]$;

(5) $(A \cup B] = (A] \cup (B]$;

(6) $((A](B]] = (AB]$.

2010 *Mathematics Subject Classification.* 06F05 .

Key words and phrases. semigroup, dual semigroup, ordered semigroup, dual ordered semigroup, left (right) ideal, 0-minimal ideal, left (right) annihilator .

The concepts of left ideals, right ideals and (two-sided) ideals in an ordered semigroup have been introduced in [6] as follows: let $(S, \cdot, \leq)$ be an ordered semigroup. A nonempty subset A of S is called a *left ideal* of S if

(i) $SA \subseteq A$;

(ii) if $x \in A$ and $y \in S$ such that $y \leq x$, then $y \in A$.

A nonempty subset A of S is called a *right ideal* of S if $AS \subseteq A$ and (ii) holds. If A is both a left and a right ideal of S, then A is called a (two-sided) *ideal* of S. It is known that, for $x \in S$, $(Sx]$ is a left ideal of S, $(xS]$ is a right ideal of S and $(SxS]$ is an ideal of S.

An element 0 of an ordered semigroup $(S, \cdot, \leq)$ is called a *zero* [2] if

(i) $0x = x0 = 0$ for all $x \in S$;

(ii) $0 \leq x$ for all $x \in S$.

Clearly, $\{0\}$ is an ideal of S which will be denoted by 0. To exclude the trivial case, if an ordered semigroup $(S, \cdot, \leq)$ has a zero 0 then we assume that $S \neq \{0\}$.

Let $(S, \cdot, \leq)$ be an ordered semigroup with zero 0. A left ideal A of S is said to be 0-*minimal* if $\{0\} \neq A$ and $\{0\}$ is the only left ideal of S properly contained in A. Similarly, we define 0-minimal right ideals and 0-minimal two-sided ideals.

Let $(S, \cdot, \leq)$ be an ordered semigroup with zero 0. Analogously to [11], if A is a nonempty subset of S, then the *left annihilator* of A, denoted by $l(A)$, is defined by

$$l(A) = \{x \in S \mid xA = 0\}.$$

Dually, the *right annihilator* of A, denoted by $r(A)$, is defined by

$$r(A) = \{x \in S \mid Ax = 0\}.$$

It is easy to see that $l(A)A = 0$ and $Ar(A) = 0$.

Lemma 1.1 *Let $(S, \cdot, \leq)$ be an ordered semigroup with zero 0 and A, B nonempty subsets of S. Then the following statements hold:*

(1) *$l(A)$ is a left ideal of S and $r(A)$ is a right ideal of S;*

(2) *$A \subseteq r(l(A))$, $A \subseteq l(r(A))$;*

(3) *if $A \subseteq B$, then $l(B) \subseteq l(A)$ and $r(B) \subseteq r(A)$;*

(4) *if $A_\alpha \subseteq S$, $\alpha \in \Lambda$, then*

$$l(\bigcup_\alpha A_\alpha) = \bigcap_\alpha l(A_\alpha),\ r(\bigcup_\alpha A_\alpha) = \bigcap_\alpha r(A_\alpha).$$

Proof. (1) We will show that $l(A)$ is a left ideal of S. Dually, we have $r(A)$ is a right ideal of S. Clearly, $l(A) \neq \emptyset$. If $x \in S, y \in l(A)$, then $(xy)A = x(yA) = 0$, and so $xy \in l(A)$. Let $x \in l(A)$ and $y \in S$ such that $y \leq x$. Then $yA \subseteq (yA] \subseteq (xA] = 0$, and hence $y \in l(A)$.

(2) Since $l(A)A = 0$, so $A \subseteq r(l(A))$. Similarly, $A \subseteq l(r(A))$.

(3) Assume that $A \subseteq B$. Let $x \in l(B)$. Since $A \subseteq B$, we get $xA \subseteq xB = 0$, and so $x \in l(A)$. Thus $l(B) \subseteq l(A)$. Similarly, $r(B) \subseteq r(A)$.

(4) The proof is straightforward.

2 Main Results

Analogously to [11], we define a dual ordered semigroup as follows:

Definition 2.1 *Let $(S, \cdot, \leq)$ be an ordered semigroup. Then S is called a dual ordered semigroup if*

(i) $l(r(L)) = L$ *for all left ideals* L *of* S*;*

(ii) $r(l(R)) = R$ *for all right ideals* R *of* S*.*

Lemma 2.2 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero* 0*.*

(1) *If $\{R_\alpha \mid \alpha \in \Lambda\}$ is a family of right ideals of S, then*

$$l(\textstyle\bigcap_\alpha R_\alpha) = \bigcup_\alpha l(R_\alpha).$$

(2) *If $\{L_\alpha \mid \alpha \in \Lambda\}$ is a family of left ideals of S, then*

$$r(\textstyle\bigcap_\alpha L_\alpha) = \bigcup_\alpha r(L_\alpha).$$

(3) $l(S) = r(S) = 0$.

(4) *If L is a 0-minimal left ideal of S, then $r(L)$ is a maximal right ideal of S.*

(5) *If A is a 0-minimal ideal of S, then $r(A)$ and $l(A)$ are maximal ideals of S.*

Proof. For (1) and (2), the proofs are straightforward.
(3) We have

$$r(S) = r(S \cup l(0)) = r(S) \cap r(l(0)) = r(S) \cap 0 = 0.$$

Similarly, $l(S) = 0$.

(4) Assume that L is a 0-minimal left ideal of S. Since $L \neq 0$, $r(L) \neq S$. Let R be a proper right ideal of S such that $r(L) \subseteq R$. Then $0 \neq l(R) \subseteq l(r(L)) = L$, and thus $l(R) = L$. Hence $R = r(l(R)) = r(L)$.

(5) Assume that A is a 0-minimal ideal of S. We will show that $r(A)$ is a maximal ideal of S. It is easy to see that $r(A)$ is an ideal of S. Let M be a proper ideal of S such that $r(A) \subseteq M$. Then $0 \neq l(M) \subseteq l(r(A)) = A$, and thus $l(M) = A$. Hence $M = r(l(M)) = r(A)$. Therefore, $r(A)$ is a maximal ideal of S. Similar arguments show that $l(A)$ is a maximal ideal of S.

Lemma 2.3 *If $(S, \cdot, \leq)$ is a dual ordered semigroup with zero 0, then $a \in (Sa]$ and $a \in (aS]$ for every $a \in S$. In particular, $(S^2] = S$.*

Proof. Let $a \in S$. Since $(Sa]$ is a left ideal of S, by assumption, we have $l(r((Sa])) = (Sa]$. If $x \in r((Sa])$, then $(Sa]x = 0$, and hence $(Sax] = 0$. By Lemma 2.2, $ax \in r(S)$, and so $ax = 0$. This proves that $a \in l(r((Sa]))$. Hence $a \in (Sa]$. Dually, $a \in (aS]$.

Lemma 2.4 *Let $(S, \cdot, \leq)$ is a dual ordered semigroup with zero 0 and $a \in S$. If $(aS] = 0$ or $(Sa] = 0$, then $a = 0$.*

Proof. This follows by Lemma 2.3.

Lemma 2.5 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero 0. If $S = (aS]$ for every $a \in S \setminus \{0\}$, then S is itself a 0-minimal right ideal of S.*

Proof. Assume that $S = (aS]$ for every $a \in S \setminus \{0\}$. Let A be a right ideal of S such that $A \neq \{0\}$. Then there exists $a \in A \setminus \{0\}$. By assumption, $S = (aS]$, and thus $S = A$. This shows that S contains only the right ideals S and $\{0\}$. Therefore, the assertion follows.

We now prove the main result analogue to ([11], Theorem 4).

Theorem 2.6 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero 0. Every nonzero right ideal of S contains a 0-minimal right ideal of S.*

Proof. Let R be a non-zero right ideal of S. There are two cases to consider:

Case 1: $S = (aS]$ for every $a \in S \setminus \{0\}$. By Lemma 2.5, we have S is itself a 0-minimal right ideal of S. Therefore, R contains a 0-minimal right ideal of S.

Case 2: $(aS] \neq S$ for some $a \in S \setminus \{0\}$. We have $a \in (aS] \subseteq S$. Since $a \in (Sa]$, there exists $y \in S$ such that $a \leq ya$. If $y \in l(aS)$, then $yaS = 0$, and so $(yaS] = 0$. Hence $ya = 0$. This is a contradiction. This shows that $y \notin l(aS)$ which implies $y \notin l((aS])$. If $l((aS]) = 0$, then $(aS] = r(l((aS])) = r(0) = S$. This is a contradiction. We have $l((aS]) \neq 0$.

Let L_0 be the union of all left ideals of S which does not contain y. Since

$$l((aS]) \subseteq L_0 \neq S,$$

it follows that

$$r(L_0) \subseteq r(l((aS])) = (aS] \subseteq S$$

and $r(L_0) \neq 0$.

We will show that $r(L_0)$ is a 0-minimal right ideal of S. Let R_1 be a right ideal of S such that $0 \neq R_1 \subset r(L_0)$. Then $L_0 \subset l(R_1) \subset S$, and thus $y \in l(R_1)$. Since $l(R_1)R_1 = 0$, $yR_1 = 0$. Since $l((aS]) \subseteq L_0$, $R_1 \subseteq r(L_0) \subseteq (aS]$. If $x \in R_1 \subseteq (aS]$, then there is $z \in S$ such that $x \leq az \leq yaz = 0$, and thus $R_1 = 0$. This is a contradiction. Hence the proof is completed.

Theorem 2.7 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero 0. Every nonzero left ideal of S contains a 0-minimal left ideal of S.*

Proof. This can be proved similarly to Theorem 2.6.

Corollary 2.8 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero 0. Every right ideal R of S such that $R \neq S$ is contained in a maximal right ideal of S.*

Proof. Let R be a right ideal of S such that $R \neq S$. Since $l(R)$ is a left ideal of S, by Theorem 2.6(6), $l(R)$ contains a 0-minimal left ideal L_0 of S. Since $0 \neq L_0 \subseteq l(R)$, we have $R \subseteq r(L_0) \subset S$, By Lemma 2.2, $r(L_0)$ is a maximal right ideal of S.

Theorem 2.9 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero 0. Every 0-minimal left ideal of S is contained in a 0-minimal ideal of S.*

Proof. Let L_0 be a 0-minimal left ideal of S. By Lemma 2.3, $L_0 \subseteq (L_0S]$. We have $(L_0S]$ is a 0-minimal ideal of S. This proves the assertion.

We will show that $M_0 := (L_0S]$ is a 0-minimal ideal of S. It is easy to see that M_0 is an ideal of S. Setting

$$Z = S \setminus r(L_0) := \{z_\alpha \mid \alpha \in \Lambda\},$$

we have

$$M_0 = (L_0(r(L_0) \cup Z)] = (L_0 Z] = \bigcup_{\alpha\in\Lambda}(L_0 z_\alpha].$$

Note that for $a \in S$, $(L_0 a] = 0$ or $(L_0 a]$ is a 0-minimal left ideal of S. In fact: we assume that $(L_0 a] \neq 0$. Let L be a left ideal of S such that $0 \neq L \subseteq (L_0 a]$. Setting $L_1 = \{x \in L_0 \mid xa \in L\}$. It is easy to see that L is a left ideal of S. By the minimality of L_0, we obtain $L = L_0$. Hence, $L = (L_0 a]$.

Now, since $L_0 \subseteq M_0$, there exists $z_0 \in Z$ such that $L_0 = (L_0 z_0]$.

Let M be an ideal of S such that $0 \neq M \subseteq M_0$. We claim that $L_0 \subseteq M$. Suppose not, then

$$M = \bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha]$$

for some $\Lambda_1 \subseteq \Lambda$ such that $z_0 \notin \{z_\alpha \mid \alpha \in \Lambda_1\}$. Since $MS \subseteq M$, we obtain

$$\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha]S \subseteq \bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha],$$

thus

$$(\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha](S]] = (\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha S] \subseteq (\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha]] \subseteq \bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha].$$

Since

$$(\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha](S]] = \bigcup_{\alpha\in\Lambda_1}((L_0 z_\alpha](S]] = \bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha S],$$

we get $\bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha S] \subseteq \bigcup_{\alpha\in\Lambda_1}(L_0 z_\alpha]$.

Let $\alpha \in \Lambda_1$. Since

$$(L_0 z_\alpha S] = ((L_0](z_\alpha S]] = (L_0(z_\alpha S]]$$

and $(L_0 z_0]$ is not contained in M, we have $z_0 \notin (z_\alpha S]$. Since $r(L_0)$ is a maximal right ideal of S, it follows that $S = (z_\alpha S] \cup r(L_0)$. This is a contradiction sine $z_0 \notin r(L_0)$. So we have the claim.

Now, we get $L_0 \subseteq M \subseteq (L_0 S]$, and thus $(L_0 S] \subseteq (MS] \subseteq (L_0 S]$. Since $M = (MS]$, we have $M = (L_0 S] = M_0$. This completes the proof.

Corollary 2.10 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero* 0. *Every ideal of S contains (at least one) 0-minimal ideal of S.*

Proof. This follows by Theorem 2.9.

Corollary 2.11 *Let $(S, \cdot, \leq)$ be a dual ordered semigroup with zero* 0. *Every maximal left ideal of S contains a maximal ideal of S.*

Proof. Let L be a maximal left ideal of S. By Theorem 2.9, the 0-minimal right ideal $r(L)$ is contained in the 0-minimal ideal $(Sr(L)]$. Since $r(L) \subseteq (Sr(L)] \subseteq S$, we have $0 \subseteq l((Sr(L)]) \subseteq L$. By Lemma 2.2, $l((Sr(L)])$ is a maximal ideal of S.

Acknowledgment. The authors would like to thank Professor Klaus Denecke for his suggestions.

References

[1] R. Baer, *Rings with duals*, Amer. J. Math., **65** (1943), 569-584.

[2] G. Birkhoff, *Lattice Theory*, **25**, Rhode Island, American Mathematical Society Colloquium Publications, Am. Math. Soc., Providence, 1984.

[3] F. F. Bonsall, A. W. Goldie, *Annihilator algebra*, Proc. London Math. Soc., **4** (1954), 154-167.

[4] L. Fuchs, *Partially Ordered Algebraic Systems*. Great Britain: Addison-Wesley Publ. Comp., 1963.

[5] M. Hall, *A type of algebraic closure*, Ann. of Math., **40** (1939), 360-369.

[6] N. Kehayopulu, M. Tsingelis, *On left regular ordered semigroups*, Southeast Asian Bulletin of Mathematics, **25** (2002), 609-615.

[7] N. Kehayopulu, M. Tsingelis, *Ideal extensions of ordered semigroups*, Comm. Algebra, **31** (2003), 4939-4969.

[8] I. Kaplansky, *Dual rings*, Ann. of Math., **49** (1948), 689-701.

[9] M. A. Najmark, *Normirovannije koljca (Russian)*, Gostechizdat, Moskva, 1943.

[10] S. Schwarz, *On dual semigroups*, Czechoslovak Mathematical Journal, **10**(2) (1960), 201-230.

[11] S. Štefan, *On the structures of dual semigroups*, Czechslovak Mathematical Journal, **21** (1971), 461-483.

Communicated by *Klaus Denecke*

Department of Mathematics
Faculty of Science
Khon Kaen University
Khon Kaen, 40002, Thailand
e-mail: thacha@kku.ac.th, jantanan-2903@hotmail.com

SELBERG TYPE INEQUALITIES IN A HILBERT C^*-MODULE AND ITS APPLICATIONS

Kyoko Kubo*, Fumio Kubo** and Yuki Seo***

Received December 10, 2013

Abstract. In this paper, we present a Selberg type inequality in a Hilbert C^*-module, which ia simultaneous extensions of the Cauchy-Schwarz inequality and the Bessel inequality in a Hibert C^*-module. As an application, we give a generalization of the Selberg inequality in a Hilbert C^*-module.

1 Introduction The theory of Hilbert C^*-modules over non-commutative C^*-algebras firstly appeared in Paschke [18] and Rieffel [19], and it has contributed greatly to the developments of operator algebras. Recently, many researchers have studied geometric properties of Hilbert C^*-modules from a viewpoint of the operator theory. For example, Dragomir, Khosravi and Moslehian [4], and Bounader and Chahbi [3] showed several variants of the Bessel inequality, the Selberg inequality and these generalizations in the framework of a Hilbert C^*-module. We showed in [6] the new Cauchy-Schwarz inequality in a Hilberet C^*-module by means of the operator geometric mean. From the viewpoint, we show a Hilbert C^*-module version of the Selberg inequality which is simultaneous extensions of the Cauchy-Schwarz inequality and the Bessel one in a Hilbert C^*-module.

We briefly review the Selberg inequality and its generalization in a Hilbert space.

Let H be a Hilbert space with the inner product $\langle \cdot, \cdot \rangle$. The Selberg inequality [2, 17] states that if $y_1, y_2, \ldots, y_n$ and x are nonzero vectors in H, then

$$\sum_{i=1}^{n} \frac{|\langle y_i, x\rangle|^2}{\sum_{j=1}^{n} |\langle y_j, y_i\rangle|} \leq \| x \|^2 . \tag{1.1}$$

Moreover, Furuta [10] posed conditions enjoying the equality: The equality in (1.1) holds if and only if $x = \sum_{i=1}^{n} a_i y_i$ for some scalars $a_1, a_2, \ldots, a_n \in \mathbb{C}$ such that for arbitrary $i \neq j$

$$\langle y_i, y_j\rangle = 0 \quad \text{or} \quad |a_i| = |a_j| \ \text{ with } \ \langle a_i y_i, a_j y_j\rangle \geq 0, \tag{1.2}$$

also see [7]. Note that the Selberg inequality is simultaneous extensions of the Bessel inequality and the Cauchy-Schwarz inequality.

Fujii and Nakamoto [9] showed a refinement of the Selberg inequality: If $\langle y, y_i\rangle = 0$ for given nonzero vectors $y_1, \ldots, y_n \in H$, then

$$|\langle x, y\rangle|^2 + \sum_{i=1}^{n} \frac{|\langle x, y_i\rangle|^2}{\sum_{j=1}^{n} |\langle y_j, y_i\rangle|} \| y \|^2 \leq \| x \|^2 \| y \|^2 \tag{1.3}$$

holds for all $x \in H$. Also, Bombieri [1] showed the following generalization of the Bessel inequality: If $x, y_1, \ldots, y_n$ are nonzero vectors in H, then

$$\sum_{i=1}^{n} |\langle x, y_i\rangle|^2 \leq \| x \|^2 \max_{1 \leq i \leq n} \sum_{j=1}^{n} |\langle y_j, y_i\rangle|. \tag{1.4}$$

2010 *Mathematics Subject Classification.* 46L08, 47A63.

Key words and phrases. Hilbert C^*-module, Selberg inequality, Bessel inequality, Cauchy-Schwarz inequality.

Moreover, Mitrinović, Pečărić and Fink [17, Theorem 5 in pp394] mentioned the following inequality equivalent to Bombieri's type: If $x, y_1, \ldots, y_n$ are nonzero vectors in H and $a_1, \ldots, a_n \in \mathbb{C}$, then

$$|\sum_{i=1}^{n} a_i \langle x, y_i \rangle|^2 \leq \| x \|^2 \sum_{i=1}^{n} |a_i|^2 \sum_{j=1}^{n} |\langle y_j, y_i \rangle|. \tag{1.5}$$

In this paper, from a viewpoint of the operator theory, we propose a Selberg type inequality in a Hilbert C^*-module, which ia simultaneous extensions of the Bessel inequality and the Cauchy-Schwarz inequality in a Hibert C^*-module. As applications, we show Hilbert C^*-module versions of Fujii-Nakamoto type (1.3), Bombieri type (1.4) and Mitrinović, Pečărić and Fink type (1.5). Moreover, we give a generalization of the Selberg inequality in a Hilbert C^*-module.

2 Preliminaries Let $\mathscr{A}$ be a unital C^*-algebra with the unit element e. An element $a \in \mathscr{A}$ is called positive if it is selfadjoint and its spectrum is contained in $[0, \infty)$. For $a \in \mathscr{A}$, we denote the absolute value of a by $|a| = (a^*a)^{\frac{1}{2}}$. For positive elements $a, b \in \mathscr{A}$, the operator geometric mean of a and b is defined by

$$a \sharp b = a^{\frac{1}{2}} \left(a^{-\frac{1}{2}} b a^{-\frac{1}{2}} \right)^{\frac{1}{2}} a^{\frac{1}{2}}$$

for invertible a. If a and b are non invertible, then $a \sharp b$ belongs to the double commutant $\mathscr{A}''$ in general. In fact, since $a \sharp b$ satisfies the upper semicontinuity, it follows that $a \sharp b = \lim_{\varepsilon \to +0} (a + \varepsilon e) \sharp (b + \varepsilon e)$ in the strong operator topology. If $\mathscr{A}$ is monotone complete in the sense that every bounded increasing net in the self-adjoint part has a supremum with respect to the usual partial order, then we have $a \sharp b \in \mathscr{A}$, see [13]. The operator geometric mean has the symmetric property: $a \sharp b = b \sharp a$. In the case that a and b commute, we have $a \sharp b = \sqrt{ab}$. For more details on the operator geometric mean, see [12, 8].

A complex linear space $\mathscr{X}$ is said to be an inner product $\mathscr{A}$-module (or a pre-Hilbert $\mathscr{A}$-module) if $\mathscr{X}$ is a right $\mathscr{A}$-module together with a C^*-valued map $(x, y) \mapsto \langle x, y \rangle : \mathscr{X} \times \mathscr{X} \to \mathscr{A}$ such that

(i) $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle \quad (x, y, x \in \mathscr{X}, \alpha, \beta \in \mathbb{C})$,

(ii) $\langle x, ya \rangle = \langle x, y \rangle a \quad (x, y \in \mathscr{X}, a \in \mathscr{A})$,

(iii) $\langle y, x \rangle = \langle x, y \rangle^* \quad (x, y \in \mathscr{X})$,

(iv) $\langle x, x \rangle \geq 0 \ (x \in \mathscr{X})$ and if $\langle x, x \rangle = 0$, then $x = 0$.

We always assume that the linear structures of $\mathscr{A}$ and $\mathscr{X}$ are compatible. Notice that (ii) and (iii) imply $\langle xa, y \rangle = a^* \langle x, y \rangle$ for all $x, y \in \mathscr{X}, a \in \mathscr{A}$. If $\mathscr{X}$ satisfies all conditions for an inner-product $\mathscr{A}$-module except for the second part of (iv), then we call $\mathscr{X}$ a semi-inner product $\mathscr{A}$-module.

In this case, we write $\| x \| := \sqrt{\| \langle x, x \rangle \|}$, where the latter norm denotes the C^*-norm of $\mathscr{A}$. If an inner-product $\mathscr{A}$-module $\mathscr{X}$ is complete with respect to its norm, then $\mathscr{X}$ is called a *Hilbert C^*-module.* In [6], from a viewpoint of operator theory, we presented the following Cauchy-Schwarz inequality in the framework of a semi-inner product C^*-module over a unital C^*-algebra: If $x, y \in \mathscr{X}$ such that the inner product $\langle x, y \rangle$ has a polar decomposition $\langle x, y \rangle = u |\langle x, y \rangle|$ with a partial isometry $u \in \mathscr{A}$, then

$$|\langle x, y \rangle| \ \leq \ u^* \langle x, x \rangle u \sharp \langle y, y \rangle. \tag{2.1}$$

An element x of a Hilbert C^*-module $\mathscr{X}$ is called nonsingular if the element $\langle x, x\rangle \in \mathscr{A}$ is invertible. The set $\{x_i\} \subset \mathscr{X}$ is called orthonormal if $\langle x_i, x_j\rangle = \delta_{ij}e$. For more details on Hilbert C^*-modules, see [16].

In [4], Dragomir, Khosravi and Moslehian showed a version of the Bessel inequality and some generalizations of this inequality in the framework of Hilbert C^*-modules. Moreover, in [3], Bounader and Chahbi showed a type and refinement of Selberg inequality in Hilbert C^*-modules. We shall show an improvement of the Selberg type inequality due to Bounader and Chahbi.

3 Main theorem Fiest of all, we show the following Selberg type inequality in a Hilbert C*-module.

Theorem 1. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebra $\mathscr{A}$. If $x, y_1, \ldots, y_n$ are nonzero vectors in $\mathscr{X}$ such that $y_1, \ldots, y_n$ are nonsingular, then*

$$\sum_{i=1}^{n} \langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle \le \langle x, x\rangle. \tag{3.1}$$

The equality in (3.1) *holds if and only if $x = \sum_{i=1}^{n} y_i a_i$ for some $a_i \in \mathscr{A}$ and $i = 1, \ldots, n$ such that for arbitrary $i \ne j$ $\langle y_i, y_j\rangle = 0$ or $|\langle y_j, y_i\rangle| a_i = \langle y_i, y_j\rangle a_j$.*

Theorem 1 is simultaneous extensions of the Bessel inequality [4] and the Cauchy-Schwarz inequality [6] in a Hilbert C^*-module. As a matter of fact, if $\{y_1, \ldots, y_n\}$ is orthonormal in Theorem 1, then we have the Bessel inequality:

$$\sum_{i=1}^{n} |\langle y_i, x\rangle|^2 \le \langle x, x\rangle$$

holds for all $x \in \mathscr{X}$. If $n = 1$ and $y = y_1$ in Theorem 1 and $\langle x, y\rangle$ has a polar decomposition $\langle x, y\rangle = u|\langle x, y\rangle|$ with a partial isometry $u \in \mathscr{A}$, then we have $u|\langle x, y\rangle| \langle y, y\rangle^{-1} |\langle y, x\rangle| u^* \le \langle x, x\rangle$ and hence

$$|\langle x, y\rangle| = |\langle x, y\rangle| \langle y, y\rangle^{-1} |\langle y, x\rangle| \ \sharp\ \langle y, y\rangle \le u^* \langle x, x\rangle u \ \sharp\ \langle y, y\rangle.$$

This implies the Cauchy-Schwarz inequality (2.1).

To prove Theorem 1, we need the following two lemmas:

Lemma 2. *If $a \in \mathscr{A}$, then the operator matrix on $\mathscr{A} \oplus \mathscr{A}$*

$$A = \begin{pmatrix} |a^*| & -a \\ -a^* & |a| \end{pmatrix}$$

is positive, and $\begin{pmatrix} \xi \\ \eta \end{pmatrix} \in \mathrm{N}(A)$ if and only if $|a^*|\xi = a\eta$, where $N(A)$ is the kernel of A.

Proof. Let $a = u|a|$ be the polar decomposition of a, where u is the partial isometry in the double commutant $\mathscr{A}''$. Since it follows that $|a^*| = u|a|u^*$, we have

$$A = \begin{pmatrix} u|a|u^* & -u|a| \\ -|a|u^* & |a| \end{pmatrix} = \begin{pmatrix} u|a|^{1/2} & 0 \\ 0 & |a|^{1/2} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u|a|^{1/2} & 0 \\ 0 & |a|^{1/2} \end{pmatrix}^* \ge 0.$$

Next, it is obvious that $\begin{pmatrix}\xi\\ \eta\end{pmatrix} \in \mathrm{Ker}(A)$ if and only if $|a|\eta = a^*\xi$ and $|a^*|\xi = a\eta$. Moreover, it follows that $|a|\eta = a^*\xi$ if and only if $|a^*|\xi = a\eta$. In fact, if $|a|\eta = a^*\xi$, then we have $a\eta = u|a|\eta = ua^*\xi = u|a|u^*\xi = |a^*|\xi$. Conversely, if $|a^*|\xi = a\eta$, then we have $a^*\xi = u^*|a^*|\xi = u^*a\eta = u^*u|a|\eta = |a|\eta$. □

Lemma 3. *For any $y_1, y_2, \ldots, y_n \in \mathscr{X}$*

$$\begin{pmatrix} \langle y_1, y_1\rangle & \cdots & \langle y_1, y_n\rangle \\ & \ddots & \\ \langle y_n, y_1\rangle & \cdots & \langle y_n, y_n\rangle \end{pmatrix} \le \begin{pmatrix} \sum_{j=1}^n |\langle y_j, y_1\rangle| & & 0 \\ & \ddots & \\ 0 & & \sum_{j=1}^n |\langle y_j, y_n\rangle| \end{pmatrix}. \tag{3.2}$$

Proof. The difference between both sides of (3.2) is the following form:

$$\sum_{i,j=1}^n \begin{pmatrix} 0 & & & 0 \\ & |\langle y_j, y_i\rangle| & -\langle y_i, y_j\rangle & \\ & -\langle y_i, y_j\rangle & |\langle y_i, y_j\rangle| & \\ 0 & & & 0 \end{pmatrix}$$

and for each pair i, j it is positive by Lemma 2. □

Proof of Theorem 1 For each $i = 1, \ldots, n$, put $c_i = \sum_{j=1}^n |\langle y_j, y_i\rangle|$. Since y_i is nonsingular, it follows that c_i is invertible in $\mathscr{A}$. It follows from Lemma 3 that

$$\begin{aligned}
&\sum_{i=1}^n \langle x, y_i\rangle c_i^{-1}\langle y_i, y_j\rangle c_j^{-1}\langle y_j, x\rangle \\
&= (\langle x, y_1\rangle c_1^{-1} \cdots \langle x, y_n\rangle c_n^{-1}) \begin{pmatrix} \langle y_1, y_1\rangle & \cdots & \langle y_1, y_n\rangle \\ & \ddots & \\ \langle y_n, y_1\rangle & \cdots & \langle y_n, y_n\rangle \end{pmatrix} \begin{pmatrix} c_1^{-1}\langle y_1, x\rangle \\ \vdots \\ c_n^{-1}\langle y_n, x\rangle \end{pmatrix} \\
&\le (\langle x, y_1\rangle c_1^{-1} \cdots \langle x, y_n\rangle c_n^{-1}) \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \begin{pmatrix} c_1^{-1}\langle y_1, x\rangle \\ \vdots \\ c_n^{-1}\langle y_n, x\rangle \end{pmatrix} \\
&= \sum_{i=1}^n \langle x, y_i\rangle c_i^{-1}\langle y_i, x\rangle
\end{aligned}$$

and this implies

$$\begin{aligned}
0 &\le \langle x - \sum_{i=1}^n y_i c_i^{-1}\langle y_i, x\rangle, x - \sum_{i=1}^n y_i c_i^{-1}\langle y_i, x\rangle\rangle \\
&= \langle x, x\rangle - 2\sum_{i=1}^n \langle x, y_i\rangle c_i^{-1}\langle y_i, x\rangle + \sum_{i=1}^n \langle x, y_i\rangle c_i^{-1}\langle y_i, y_j\rangle c_j^{-1}\langle y_j, x\rangle \\
&\le \langle x, x\rangle - \sum_{i=1}^n \langle x, y_i\rangle c_i^{-1}\langle y_i, x\rangle.
\end{aligned}$$

Hence we have the desired inequality (3.1).

The equality in (3.1) holds if and only if the following (3.3) and (3.4) are satisfied:

$$x = \sum_{i=1}^{n} y_i c_i^{-1} \langle y_i, x\rangle \tag{3.3}$$

and for arbitrary $i \neq j$

$$(\langle x, y_i\rangle c_i^{-1} \quad \langle x, y_j\rangle c_j^{-1}) \begin{pmatrix} |\langle y_j, y_i\rangle| & -\langle y_i, y_j\rangle \\ -\langle y_j, y_i\rangle & |\langle y_i, y_j\rangle| \end{pmatrix} \begin{pmatrix} c_i^{-1}\langle y_i, x\rangle \\ c_j^{-1}\langle y_j, x\rangle \end{pmatrix} = 0. \tag{3.4}$$

Put $A = \begin{pmatrix} |\langle y_j, y_i\rangle| & -\langle y_i, y_j\rangle \\ -\langle y_j, y_i\rangle & |\langle y_i, y_j\rangle| \end{pmatrix}$ and it follows that the condition (3.4) holds if and only if

$$A^{1/2} \begin{pmatrix} c_i^{-1}\langle y_i, x\rangle \\ c_j^{-1}\langle y_j, x\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad A \begin{pmatrix} c_i^{-1}\langle y_i, x\rangle \\ c_j^{-1}\langle y_j, x\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence it follows from Lemma 2 that the condition (3.4) is equivalent to the following (3.5) and (3.6): For arbitrary $i \neq j$

$$\langle y_i, y_j\rangle = 0 \tag{3.5}$$

or

$$|\langle y_j, y_i\rangle| c_i^{-1}\langle y_i, x\rangle = \langle y_i, y_j\rangle c_j^{-1}\langle y_j, x\rangle. \tag{3.6}$$

Conversely, suppose that $x = \sum_{i=1}^{n} y_i a_i$ for some $a_i \in \mathscr{A}$ and for $i \neq j$ $\langle y_i, y_j\rangle = 0$ or $|\langle y_j, y_i\rangle| a_i = \langle y_i, y_j\rangle a_j$. Then

$$\begin{aligned}
&\sum_{i=1}^{n}\langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle = \sum_{i=1}^{n}\langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \sum_{j=1}^{n}\langle y_i, y_j\rangle a_j \\
&= \sum_{i=1}^{n}\langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \sum_{j=1}^{n} |\langle y_j, y_i\rangle| a_i \\
&= \sum_{i=1}^{n}\langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right) a_i \\
&= \sum_{i=1}^{n}\langle x, y_i\rangle a_i \\
&= \langle x, x\rangle.
\end{aligned}$$

Whence the proof is complete. □

Remark 4. (1) In the case that $\mathscr{X}$ is a Hilbert space, the equality condition $|\langle y_j, y_i\rangle| a_i = \langle y_i, y_j\rangle a_j$ in Theorem 1 implies the condition (1.2). In fact, for some scalars $a_i, a_j \in \mathbb{C}$, it follows that $\langle a_i y_i, a_j y_j\rangle = a_i^*\langle y_i, y_j\rangle a_j = a_i^* |\langle y_j, y_i\rangle| a_i \geq 0$, and $|\langle y_j, y_i\rangle| = |\langle y_j, y_i\rangle^*|$ implies $|a_i| = |a_j|$.
(2) In the Hilbert space setting, K. Kubo and F. Kubo [15] showed another proof of Selberg's inequality (1.1) using Geršgorin's location of eigenvalues [14, Theorem 6.1.1] and a diagonal domination theorem of positive semidefinite matrix.

4 Applications In this section, by using Theorem 1, we consider several Hilbert C^*-module versions of the Selberg inequality and the Bessel inequality.

Bounader and Chahbi in [3, Theorem 3.1] showed that if $\mathscr{X}$ is an inner product C^*-module and $y_1, \ldots, y_n$ are nonzero vectors in $\mathscr{X}$, and $x \in \mathscr{X}$, then

$$\sum_{i=1}^{n} \frac{|\langle y_i, x\rangle|^2}{\sum_{j=1}^{n} \| \langle y_j, y_i\rangle \|} \leq \langle x, x\rangle. \tag{4.1}$$

By Theorem 1, we have the following corollary, which is an improvement of (4.1):

Corollary 5. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebera $\mathscr{A}$. If $x, y_1, \ldots, y_n$ are nonzero vectors in $\mathscr{X}$ such that $y_1, \ldots, y_n$ are nonsingular, then*

$$\sum_{i=1}^{n} \frac{|\langle y_i, x\rangle|^2}{\| \sum_{j=1}^{n} |\langle y_j, y_i\rangle| \|} \leq \langle x, x\rangle.$$

Proof. By assumption it follows that $\sum_{i=1}^{n} |\langle y_j, y_i\rangle|$ is invertible in $\mathscr{A}$ and hence

$$\left(\sum_{i=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \geq \| \sum_{i=1}^{n} |\langle y_j, y_i\rangle| \|^{-1}.$$

Therefore, Theorem 1 implies Corollary 5. □

Moreover, Bounader and Chahbi showed a Hilbert C^*-module version of Fujii-Nakamoto type (1.3), which is a refinement of (4.1): If y and $y_1, \ldots, y_n$ are nonzero vectros in $\mathscr{X}$ such that $\langle y, y_i\rangle = 0$ for $i = 1, \ldots, n$, and $x \in \mathscr{X}$, then

$$|\langle y, x\rangle|^2 + \sum_{i=1}^{n} \frac{|\langle y_i, x\rangle|^2}{\sum_{j=1}^{n} \| \langle y_i, y_j\rangle \|} \| \langle y, y\rangle \| \leq \| \langle y, y\rangle \| \langle x, x\rangle. \tag{4.2}$$

We show a Hilbert C^*-module version of a refinement of the Selberg inequality due to Fujii and Nakamoto, which is another version of (4.2):

Theorem 6. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebera $\mathscr{A}$. If $x, y, y_1, \ldots, y_n$ are nonzero vectors in $\mathscr{X}$ such that $y_1, \ldots, y_n$ are nonsingular, $\langle y, y_i\rangle = 0$ for $i = 1, \cdots, n$ and $\langle x, y\rangle = u|\langle x, y\rangle|$ is a polar decomposition in $\mathscr{A}$, i.e., $u \in \mathscr{A}$ is a partial isometry, then*

$$|\langle y, x\rangle| \leq u^*\langle y, y\rangle u \ \sharp \ \left(\langle x, x\rangle - \sum_{i=1}^{n} \langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle\right) \tag{4.3}$$
$$\left(\leq u^*\langle y, y\rangle u \ \sharp \ \langle x, x\rangle\right).$$

Proof. Put $z = x - \sum_{i=1}^{n} y_i \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle$. By the proof of Theorem 1, we have

$$\langle z, z\rangle \leq \langle x, x\rangle - \sum_{i=1}^{n} \langle x, y_i\rangle \left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle.$$

Since $\langle y, z\rangle = \langle y, x\rangle$, it follows from the monotonicity of the operator geometric mean that

$$\begin{aligned}|\langle y, x\rangle| = |\langle y, z\rangle| &\le u^*\langle y, y\rangle u \ \sharp \ \langle z, z\rangle \quad \text{by the Cauchy-Schwarz inequality (2.1)}\\ &\le u^*\langle y, y\rangle u \ \sharp \ \left(\langle x, x\rangle - \sum_{i=1}^n \langle x, y_i\rangle \left(\sum_{j=1}^n |\langle y_j, y_i\rangle|\right)^{-1} \langle y_i, x\rangle\right).\end{aligned}$$

□

In [3, Corollary 3.5], Bounader and Chahbi showed a Hilbert C^*-module version of Bombieri type (1.4): If $y_1, \dots, y_n$ are nonzero vectors in $\mathscr{X}$ and $x \in \mathscr{X}$, then

$$\sum_{i=1}^n |\langle y_i, x\rangle|^2 \le \langle x, x\rangle \max_{1\le i\le n} \sum_{j=1}^n \| \langle y_i, y_j\rangle \| . \tag{4.4}$$

We show a Hilbert C^*-module version of Bombieri type, which is an improvement of (4.4):

Theorem 7. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebra $\mathscr{A}$. If $x, y_1, \dots, y_n$ are nonzero vectors in $\mathscr{X}$ such that $y_1, \dots, y_n$ are nonsingular, then*

$$\sum_{i=1}^n |\langle y_i, x\rangle|^2 \le \langle x, x\rangle \max_{1\le i\le n} \| \sum_{j=1}^n |\langle y_j, y_i\rangle| \| .$$

Proof. Since for $i = 1, \dots, n$

$$\sum_{j=1}^n |\langle y_j, y_i\rangle| \le \| \sum_{j=1}^n |\langle y_j, y_i\rangle| \| \le \max_{1\le i\le n} \| \sum_{j=1}^n |\langle y_j, y_i\rangle| \|,$$

we have this theorem by virtue of Theorem 1. □

As a corollary, we have the following Boas-Bellman type inequality [3, Corollary 3.6]:

Corollary 8. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebra $\mathscr{A}$. If $x, y_1, \dots, y_n$ are nonzero vectors in $\mathscr{X}$ such that $y_1, \dots, y_n$ are nonsingular, then*

$$\sum_{i=1}^n |\langle y_i, x\rangle|^2 \le \langle x, x\rangle \left(\max_{1\le i\le n} \| \langle y_i, y_i\rangle \| + (n-1) \max_{j\ne i} \| \langle y_j, y_i\rangle \| \right).$$

Finally, we show a Mitrinović-Pečarić-Fink type inequality [17, Theorem 5 in pp394] in Hilbert C^*-modules, which is another version of [4, Theorem 3.8]:

Theorem 9. *Let $\mathscr{X}$ be an inner product C^*-module over a unital C^*-algebra $\mathscr{A}$. If $x, y_1, \dots, y_n$ are nonzero vectors in $\mathscr{X}$ and $a_1, \cdots, a_n \in \mathscr{A}$ such that $y_1, \dots, y_n$ are nonsingular and $\langle x, \sum_{i=1}^n y_i a_i\rangle = u|\langle x, \sum_{i=1}^n y_i a_i\rangle|$ is a polar decomposition in $\mathscr{A}$, i.e., $u \in \mathscr{A}$ is a partial isometry, then*

$$|\sum_{i=1}^n \langle x, y_i\rangle a_i| \le u^*\langle x, x\rangle u \ \sharp \ \left(\sum_{i=1}^n a_i^* \left(\sum_{j=1}^n |\langle y_j, y_i\rangle|\right) a_i\right).$$

Proof. By the Cauchy-Schwarz inequality (2.1), we have

$$\begin{aligned}
|\sum_{i=1}^{n}\langle x, y_i\rangle a_i| &= |\langle x, \sum_{i=1}^{n} y_i a_i\rangle| \\
&\le u^*\langle x,x\rangle u \ \sharp\ \left(\langle \sum_{i=1}^{n} y_i a_i, \sum_{i=1}^{n} y_i a_i\rangle\right) \\
&= u^*\langle x,x\rangle u \ \sharp\ \left(\sum_{i,j=1}^{n} a_i^*\langle y_i, y_j\rangle a_j\right) \\
&\le u^*\langle x,x\rangle u \ \sharp\ \left(\sum_{i=1}^{n} a_i^*\left(\sum_{j=1}^{n} |\langle y_j, y_i\rangle|\right) a_i\right) \qquad \text{by Lemma 3.}
\end{aligned}$$

□

5 Generalization In this section, we present a generalization of the Selberg inequality in a Hilbert C^*-module.

We review the basic concepts of adjointable operators on a Hilbert C^*-module$\mathscr{X}$ over a unital C^*-algebra $\mathscr{A}$. We define $\mathcal{L}(\mathscr{X})$ to be the set of all maps $T : \mathscr{X} \mapsto \mathscr{X}$ for which there is a map $T^* : \mathscr{X} \mapsto \mathscr{X}$ such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in \mathscr{X}$. For $T \in \mathcal{L}(\mathscr{X})$, we denote the kernel of T by $N(T)$. A closed submodule $\mathscr{M}$ of $\mathscr{X}$ is said to be complemented if $\mathscr{X} = \mathscr{M} \oplus \mathscr{M}^{\perp}$. Suppose that the closures of the ranges of T and T^* are both complemented. Then it follows from [16, Proposition 3.8] that T has a polar decomposition $T = U|T|$ with a partial isometry $U \in \mathcal{L}(\mathscr{X})$ and $N(U) = N(|T|)$, and the following hold:

(i) $N(|T|) = N(T)$.

(ii) $|T^*|^q = U|T|^qU^*$ for any positive number $q > 0$.

(iii) $N(S^q) = N(S)$ for any positive operator $S \in \mathcal{L}(\mathscr{X})$ and $q > 0$,

also see [5, 20].

Theorem 10. *Let T be an operator in $\mathcal{L}(\mathscr{X})$ such that the closures of the ranges of T and T^* are both complemented. If $y_1, \dots, y_n \notin \mathrm{N}(T^*)$ are nonsingular, then*

$$\sum_{i=1}^{n}\langle Tx, y_i\rangle \left(\sum_{j=1}^{n} |\langle |T^*|^{2\beta)} y_j, y_i\rangle|\right)^{-1} \langle y_i, Tx\rangle \le \langle |T|^{2\alpha}x, x\rangle \tag{5.1}$$

holds for every $x \notin \mathrm{N}(T)$ and for any $\alpha, \beta \in [0,1]$ with $\alpha + \beta = 1$. In particular,

$$\sum_{i=1}^{n}\langle Tx, y_i\rangle \left(\sum_{j=1}^{n} |\langle TT^* y_j, y_i\rangle|\right)^{-1} \langle y_i, Tx\rangle \le \langle U^*Ux, x\rangle \tag{5.2}$$

and

$$\sum_{i=1}^{n}\langle Tx, y_i\rangle \left(\sum_{j=1}^{n} |\langle UU^* y_j, y_i\rangle|\right)^{-1} \langle y_i, Tx\rangle \le \langle T^*Tx, x\rangle. \tag{5.3}$$

Moreover, the equality in (5.1) *holds if and only if* $Tx = \sum_{i=1}^n |T^*|^{2\beta} y_i a_i$ *for some* $a_1, \ldots, a_n \in \mathscr{A}$ *such that for arbitrary* $i \neq j$, $\langle |T^*|^{2\beta} y_i, y_j \rangle = 0$ *or* $|\langle |T^*|^{2\beta} y_j, y_i \rangle| a_i = \langle |T^*|^{2\beta} y_i, y_j \rangle a_j$.

Proof. Let $T = U|T|$ be the polar decomposition of T, where U is the partial isometry. In the case of $\alpha = 0$ or 1, it follows from Theorem 1 that replacing x by U^*Ux (resp. $|T|x$) and y_i by $|T|U^*y_i$ (resp. U^*y_i) for all $i = 1, \ldots, n$, it follows that $\langle U^*Ux, |T|U^*y_i \rangle = \langle Ux, U|T|U^*y_i \rangle = \langle x, U^*|T^*|y_i \rangle = \langle x, T^*y_i \rangle = \langle Tx, y_i \rangle$ and we have (5.2) (resp. (5.3)). In the case of $0 < \alpha < 1$, we replace x by $|T|^\alpha x$ and also replace y_i by $|T|^\beta U^* y_i$ for all $i = 1, \ldots, n$. Then we have

$$\langle |T|^\beta U^* y_i, |T|^\beta U^* y_j \rangle = \langle U|T|^{2\beta} U^* y_i, y_j \rangle = \langle |T^*|^{2\beta} y_i, y_j \rangle$$

and $y_1, \ldots, y_n \notin \mathrm{N}(T^*) = \mathrm{N}(|T^*|) = \mathrm{N}(|T^*|^\beta)$. Thus we have (5.1) by Theorem 1.

Next, we consider the equality condition in (5.1). By (iii), we have

$$|T|^\alpha x = \sum_{i=1}^n |T|^\beta U^* y_i a_i \quad \Longleftrightarrow \quad |T|^{2\alpha} x = \sum_{i=1}^n |T| U^* y_i a_i = \sum_{i=1}^n T^* y_i a_i.$$

Hence we have the following implication:

$$\begin{aligned} |T|^\alpha x = \sum_{i=1}^n |T|^\beta U^* y_i a_i \quad &\Longleftrightarrow \quad |T|x = |T|^{\alpha+\beta} x = \sum_{i=1}^n |T|^{2\beta} U^* y_i a_i \quad \text{by (iii)} \\ &\Longleftrightarrow \quad U|T|x = \sum_{i=1}^n U|T|^{2\beta} U^* y_i a_i \quad \text{by (i) and (iii)} \\ &\Longleftrightarrow \quad Tx = \sum_{i=1}^n |T^*|^{2\beta} y_i a_i. \quad \text{by (ii).} \end{aligned}$$

Whence the proof is complete. □

References

[1] E. Bombieri, *A note on the large sieve*, Acta. Arith., **18** (1971), 401–404.

[2] E. Bombieri, *Le Grand Grible dans la Theórie Analytique des Nombres*, Asterisque 18, Societe Mathematique de France, 1974.

[3] N. Bounader and A. Chahbi, *Selberg type inequalities in Hilbert C^*-modules*, Int. J. Analy., **7** (2013), 385–391.

[4] S.S. Dragomir, M. Khosravi and M.S. Moslehian, *Bessel type inequalities in Hilbert C^*-modules*, Linear Multilinear Algebra, **58** (2010), 967–975.

[5] J.I. Fujii, M. Fujii and Y. Seo, *Operator inequalities on Hilbert C^*-modules via the Cauchy-Schwarz inquality*, to appear in Math. Ineq. Appl.

[6] J.I. Fujii, M. Fujii, M.S. Moslehian and Y. Seo, *Cauchy-Schwarz inequality in semi-inner product C^*-modules via polar decomposition*, J. Math. Anal. Appl., **394** (2012), 835-840.

[7] M. Fujii, K. Kubo and S. Otani, *A graph theoretical observation on the Selberg inequality*, Math. Japon., **35** (1990), 381–385.

[8] M. Fujii, J. Mićić Hot, J. Pečarić and Y. Seo, *Recent Developments of Mond-Pečarić Method in Operator Inequalities*, Monographs in Inequalities 4, Element, Zagreb, 2012.

[9] M. Fujii and R. Nakamoto, *Simultaneous extensions of Selberg inequality and Heinz-Kato-Furuta inequality*, Nihonkai Math. J., **9** (1998), 219–225.

[10] T. Furuta, *When does the equality of a generalized Selberg inequality hold?*, Nihonkai Math. J., **2** (1991), 25–29.

[11] T. Furuta, *Invitation to Linear Operators*, Taylor & Francis, London, 2001.

[12] T. Furuta, J. Mićić Hot, J. Pečarić and Y. Seo, *Mond-Pečarić Method in Operator Inequalities*, Monographs in Inequalities 1, Element, Zagreb, 2005.

[13] M. Hamana, *Partial $*$-automorphisms, normalizers, and submodules in monotome complete C^*-algebras*, Canad. J.Math., **58** (2006), 1144–1202.

[14] R.A. Horn and C.A. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.

[15] K. Kubo and F. Kubo, *Diagonal matrix that dominates a positive semidefinite matrix*, Technical note, 1988.

[16] E.C. Lance, *Hilbert C^*-Modules*, London Math. Soc. Lecture Note Series 210, Cambridge Univ. Press, 1995.

[17] D.S. Mitrinović, J. Pecărić and A. M. Fink, *Classical and New Inequalities in Analysis*, Kluwer Academic, Dordrecht, 1993.

[18] W.L. Paschke, *Inner product modules over B^*-algebras*, Trans. Amer. Math. Soc. **182** (1973), 443-468.

[19] M.A. Rieffel, *Morita equivalence for C^*-algebras and W^*-algebras*, J. Pure Applied Algebra, **5** (1974), 51-96.

[20] Y. Seo, *Hölder type inequalities on Hilbert C^*-modules and its reverses*, Ann. Funct. Anal., **5** (2014), 1–9.

Communicated by *Masatoshi Fujii*

* 2-7-5 Nagaemachi, Toyama, Toyama 930-0076, Japan

** Department of mathematics, Toyama University, Gofuku, Toyama 930-8555, Japan
E-mail address : fkubo@sci.u-toyama.ac.jp

*** Department of Mathematics Education, Osaka Kyoiku University, 4-698-1 Asahigaoka Kashiwara Osaka 582-8582 Japan.
E-mail address : yukis@cc.osaka-kyoiku.ac.jp

BIFURCATION OF SOLUTIONS TO SEMILINEAR ELLIPTIC PROBLEMS ON S^2 WITH A SMALL HOLE

Atsushi Kosaka

Received October 29, 2013; revised January 17, 2014

Abstract. In this paper we prove the existence of non-positive or non-radial solutions to semilinear elliptic problems on $\mathbf{S}^2$ with a small hole. When the hole is sufficiently small, we prove that the multiplicity of eigenvalues to the corresponding linearized problem is 1 or 2. Thus, by using the result, we show those eigenvalues are bifurcation points, and the corresponding bifurcating solutions are not positive except for a bifurcating solution which is corresponding to the first eigenvalue. Moreover if the multiplicity of a eigenvalue is 2, then the corresponding bifurcating solution is not radially symmetric.

1 Introduction

We investigate the existence of non-trivial solutions to

$$\begin{cases} \Delta_{\mathbf{S}^N} u + \lambda u + |u|^{p-1}u = 0 & \text{in } B_{\theta_0}, \\ u = 0 & \text{on } \partial B_{\theta_0}, \end{cases} \tag{1.1}$$

where $\Delta_{\mathbf{S}^N}$ is the Laplace–Beltrami operator on the N-dimensional unit sphere $\mathbf{S}^N$ $(N \geq 2)$ and $1 < p < \infty$. Here B_{θ_0} is a geodesic ball on $\mathbf{S}^N$ with the geodesic radius θ_0. In addition the origin of B_{θ_0} is at the North Pole $(0, 0, ..., 0, 1)$ in the $(N+1)$-dimensional Euclidean space $\mathbf{R}^{N+1}$. In this paper we consider a classical solution to (1.1) (in fact we shall prove the existence of a solution $u \in C^{2,\alpha}(B_{\theta_0})$ to (1.1) with some $\alpha \in (0, 1)$).

When $(N-2)p < N+2$ and $\lambda < \lambda_1$ (λ_1 is the first eigenvalue of $\Delta_{\mathbf{S}^N}$ on B_{θ_0} with the homogeneous Dirichlet boundary condition), we can prove the existence of a solution to (1.1) by using the mountain pass lemma (e.g., see Theorem 6.2 in Chapter II of Struwe [17]). In fact, by the Rellich-Kondrachov theorem (e.g., see Theorem 2.34 in Aubin [2]), the compactness of $H_0^1(B_{\theta_0}) \hookrightarrow L^2(B_{\theta_0})$ is guaranteed, and we can apply the mountain pass lemma.

In the case of $N \geq 3$ and $p \geq (N+2)/(N-2)$, the compactness of $H_0^1(B_{\theta_0}) \hookrightarrow L^2(B_{\theta_0})$ is lost, and hence we need other approaches to prove the existence of solutions. The first result on this problem is by Bandle, Brillard and Flucher [5]. For $N \geq 3$, $p = (N+2)/(N-2)$ and $\lambda = 0$, they proved the following result: there exists some $\theta_c \in [0, \pi)$ such that (1.1) has a positive and radial solution if and only if $\theta_0 \in (\theta_c, \pi)$ (the *radial solution* means a solution depending only on the geodesic distance from the North Pole). Additionally if $N \geq 4$, then $\theta_c = 0$. On the other hand, if $N = 3$, then $\theta_c \neq 0$. Later Bandle and Peletier [8] investigated the case $N = 3$, $p = 5$ and $\lambda = 0$ in detail, and they showed that $\theta_c = \pi/2$. Moreover the author of this paper [14] also focused attention on the case $N = 3$, $p = 5$ and $\lambda = 0$. Namely, instead of the Dirichlet boundary condition, the author assumed the Robin boundary condition and clarified the structure of positive and radial solutions to (1.1) under

2010 *Mathematics Subject Classification.* 35J61, 35B32, 33C45.

Key words and phrases. a semilinear elliptic equation, the associated Legendre function, the Lyapunov-Schmidt method.

$N = 3$ and $p = 5$. In addition, for $p > (N+2)/(N-2)$ with $N \geq 3$, a solution to (1.1) seems to exist, but it seems difficult to investigate the structure of solutions to (1.1).

The case $\lambda \neq 0$ is also studied. Bandle and Benguria [6] proved that, for $N = 3$, $p = 5$ and $\lambda > -3/4$, there exists a unique, positive and radial solution to (1.1). Here $\lambda = -3/4$ is the second eigenvalue of

$$\Delta_{\mathbf{S}^3} w + 4\lambda w = 0 \qquad \text{in } \mathbf{S}^3$$

which is the linearized equation of $\Delta_{\mathbf{S}^3} u + \lambda(|u|^4 u - u) = 0$ around $u \equiv 1$. On the other hand, Brezis and Peletier [9] showed the existence of positive and radial solutions for negatively large λ (they proved that many positive and radial solutions exist). Furthermore Bandle, Kabeya and Ninomiya [7] investigated the bifurcation structure for $\lambda < -3/4$ in detail. In studies above positive solutions are only treated, and no one investigates the structure of *non-positive* or *non-radial* solutions. Thus we focus our attention on those kinds of solutions.

Linearizing (1.1) around $u \equiv 0$, we obtain

$$\begin{cases} \Delta_{\mathbf{S}^N} w + \lambda w = 0 & \text{in } B_{\theta_0}, \\ w = 0 & \text{on } \partial B_{\theta_0}. \end{cases} \tag{1.2}$$

In this paper we only consider the case $N = 2$ and shall prove that eigenvalues of (1.2) are bifurcation points of (1.1). For the purpose we shall apply the *Lyapunov–Schmidt reduction method* and construct bifurcating solutions (for the Lyapunov–Schmidt reduction method, e.g., see Section 5.3 in Ambrosetti and Prodi [1]). To apply the method, we are required to know the multiplicity of eigenvalues for (1.2). Moreover, to see the positivity and the radial symmetry of the bifurcating solutions, we need to know profiles of eigenfunctions. Hence we shall investigate eigenvalues and eigenfunctions.

Let the polar coordinates

$$\begin{cases} y_1 = \sin\varphi \sin\theta \\ y_2 = \sin\varphi \cos\theta \\ y_3 = \cos\theta \end{cases}$$

with $(y_1, y_2, y_3) \in \mathbf{S}^2 \subset \mathbf{R}^3$, $\theta \in (0, \theta_0)$ and $\varphi \in [0, 2\pi]$. Then the operator $\Delta_{\mathbf{S}^2} + \lambda$ is expressed as

$$\Delta_{\mathbf{S}^2} w + \lambda w = \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\frac{\partial w}{\partial\theta} \sin\theta \right) + \frac{1}{\sin^2\theta} \frac{\partial^2 w}{\partial\varphi^2} + \lambda w.$$

Additionally, for convenience below, we define $\nu \geq 0$ satisfying

$$\lambda := \nu(\nu+1). \tag{1.3}$$

Solutions to (1.2) are expressed by using the separation of variables. Namely let

$$w(\theta, \varphi) = P(x)\Phi(\varphi)$$

with

$$x = \cos\theta.$$

Here, by considering the regularity of solutions, $|P(1)| < \infty$, $\Phi(0) = \Phi(2\pi)$ and $\Phi'(0) = \Phi'(2\pi)$ must be satisfied. Functions $P(x)$ and $\Phi(\varphi)$ satisfy

$$(1-x^2)\frac{d^2P}{dx^2} - 2x\frac{dP}{dx} + \left\{ \nu(\nu+1) - \frac{m^2}{1-x^2} \right\} P = 0, \tag{1.4}$$

and

$$\frac{d^2\Phi}{d\varphi^2} + m^2\Phi = 0. \tag{1.5}$$

From the periodicity of $\Phi(\varphi)$, m is a non-negative integer, and any solutions to (1.5) are expressed as $\Phi(\varphi) = C_1 \cos m\varphi + C_2 \sin m\varphi$.

On the other hand, (1.4) is known as the *associated Legendre equation.* The equation (1.4) has two kinds of solutions $P = P_\nu^m(x)$ and $Q_\nu^m(x)$ such that $|P_\nu^m(1)| < \infty$ and $|Q_\nu^m(x)| \to \infty$ as $x \to 1$, respectively. In addition these solutions are linearly independent. From the condition $|P(1)| < \infty$, we have only to treat $P = P_\nu^m(x)$. Each $P_\nu^m(x)$ satisfies

$$P_\nu^m(1) = \begin{cases} 1 & (m = 0), \\ 0 & (m = 1, 2, ...), \end{cases}$$

and

$$\mathrm{sgn}(P_\nu^m(x)) = \quad (-1)^m \quad \text{near } x = 1, \tag{1.6}$$

where $\mathrm{sgn}(a)$ denotes the signature of a (see (4.1) and (4.3) in Appendix).

We assume $\nu = j \geq 0$ which is an integer. Then, from (1.6) and

$$P_j^m(-x) = (-1)^{j+m} P_j^m(x)$$

(e.g., p.131 in Moriguchi, Udagawa and Hitotsumatsu [15]), it follows that

$$P_j^m(-1) = \begin{cases} (-1)^j & (m = 0), \\ 0 & (m = 1, 2, ...). \end{cases}$$

Hence $\lambda = j(j+1)$ and $C_1 P_j^m(\cos\theta)\cos m\varphi + C_2 P_j^m(\cos\theta)\sin m\varphi$ are an eigenvalue and an eigenfunction of $\Delta_{\mathbf{S}^2} w + \lambda w = 0$ on $\mathbf{S}^2$, respectively.

On the other hand, to solve the eigenvalue problem (1.2), we are required to find solutions to (1.4) satisfying the boundary condition

$$P(\cos\theta_0) = 0. \tag{1.7}$$

For any fixed $m = 0, 1, 2...$, there exist infinitely many $\lambda = \nu(\nu+1)$ satisfying (1.4), (1.7) and $P(1) = 1$ or 0 (e.g., see Chapter 10.6 in Ince [13], which is a general result on the Sturm–Liouville equations). But, in general, it seems difficult to investigate the multiplicity of eigenvalues for (1.2). In fact, for any m and n $(m \neq n)$, it is not known whether $P_\nu^m(\cos\theta_0) = P_\nu^n(\cos\theta_0) = 0$ holds or not (partial results are obtained by Baginski [3], [4]). Hence, for any $\theta_0 \in (0, \pi)$, we does not see the multiplicity of eigenvalues for (1.2).

Thus, in this paper, we set

$$\theta_0 = \pi - \epsilon$$

and only consider a sufficiently small $\epsilon > 0$, that is, $B_{\pi-\epsilon}$ is $\mathbf{S}^2$ with a small hole. Then we can exactly prove the multiplicity of eigenvalues λ. Hereafter we use the notation

$$(a)_k := a(a+1)(a+2)...(a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)}, \tag{1.8}$$

where $k \geq 0$ is an integer and $\Gamma(x)$ is the *gamma function* ($(a)_0 = 1$ and $(1)_k = (k-1)!$). The result on eigenvalues of (1.2) is as follows:

Theorem 1.1 *Assume $N=2$, $\theta_0=\pi-\epsilon$, $1<p<\infty$, and arbitrarily fix an integer $j\geq 0$. Then there exist $(j+1)$ positive values $\{\nu_{j,\epsilon}^m\}_{m=0}^j$ such that each $\lambda_{j,\epsilon}^m := \nu_{j,\epsilon}^m(\nu_{j,\epsilon}^m+1)$ is an eigenvalue of* (1.2). *Moreover, as $\epsilon\to 0$, it holds that*

$$\lambda_{j,\epsilon}^m = \begin{cases} j(j+1)+\dfrac{2j+1}{2|\log\epsilon|}+o\left(\dfrac{1}{|\log\epsilon|}\right) & (m=0), \\ j(j+1)+(2j+1)c_{j,m}\epsilon^{2m}+o(\epsilon^{2m}) & (1\leq m\leq j), \end{cases} \tag{1.9}$$

where $c_{j,m}=(j+m)!/[4^m m!(m-1)!(j-m)!]$.

The asymptotic formula (1.9) implies that, for sufficiently small $\epsilon>0$, it holds that

$$j(j+1)<\lambda_{j,\epsilon}^j<\lambda_{j,\epsilon}^{j-1}<\lambda_{j,\epsilon}^{j-2}<\ldots<\lambda_{j,\epsilon}^0,$$

and each $\lambda_{j,\epsilon}^m$ is located near $j(j+1)$. In addition the eigenspace corresponding to $\lambda_{j,\epsilon}^0$ is spanned by $P^0_{\nu^0_{j,\epsilon}}(\cos\theta)$. On the other hand, for $1\leq m\leq j$, the eigenspace corresponding to $\lambda_{j,\epsilon}^m$ is spanned by $P^m_{\nu^m_{j,\epsilon}}(\cos\theta)\cos m\varphi$ and $P^m_{\nu^m_{j,\epsilon}}(\cos\theta)\sin m\varphi$. Therefore the multiplicity of eigenvalues of (1.2) is 1 or 2.

By Theorem 1.1, we can apply the Lyapunov–Schmidt reduction method, and the following result is proved:

Theorem 1.2 *Assume the same assumptions as in Theorem* 1.1. *Let $\mu:=\lambda-\lambda_{j,\epsilon}^m$. Then the following statements hold*:

(i) *For $m=0$, there exist a constant $\delta_j>0$ and a non-trivial solution*

$$u_{j,\epsilon}^0(\ \cdot\ ,\ \cdot\ ;\mu+\lambda_{j,\epsilon}^0):=|\mu|^{\frac{1}{p-1}}\{t_{j,\epsilon}^0(\mu)v_*+l_{j,\epsilon}^0(\ \cdot\ ,\ \cdot\ ;\mu)\}$$

to (1.1) *for $\mu\in(-\delta_j,0)$, where $t_{j,\epsilon}^0(\mu)\in\mathbf{R}$ and $l_{j,\epsilon}^0(\ \cdot\ ,\ \cdot\ ;\mu)\in C^{2,\alpha}(B_{\pi-\epsilon})$ are of class C^1 with respect to μ. Here $t_{j,\epsilon}^0(0)=1$, $l_{j,\epsilon}^0(\ \cdot\ ,\ \cdot\ ;0)\equiv 0$ and*

$$v_*=M_0P^0_{\nu^0_{j,\epsilon}}(\cos\theta),$$

with some constant $M_0>0$.

(ii) *Arbitrarily fix $t_*\in\mathbf{R}$ and $s_*\in\mathbf{R}$ such that the condition $t_*^2+s_*^2=1$. For $1\leq m\leq j$, there exist a constant $\delta_j>0$ and a non-trivial solution*

$$u_{j,\epsilon}^m(\ \cdot\ ,\ \cdot\ ;\mu+\lambda_{j,\epsilon}^m):=|\mu|^{\frac{1}{p-1}}\{v(t_{j,\epsilon}^m(\mu),s_{j,\epsilon}^m(\mu))+l_{j,\epsilon}^m(\ \cdot\ ,\ \cdot\ ;\mu)\}$$

to (1.1) *for $\mu\in(-\delta_j,0)$, where $t_{j,\epsilon}^m(\mu),s_{j,\epsilon}^m(\mu)\in\mathbf{R}$ and $l_{j,\epsilon}^m(\ \cdot\ ,\ \cdot\ ;\mu)\in C^{2,\alpha}(B_{\pi-\epsilon})$ are of class C^1 with respect to μ. Here $t_{j,\epsilon}^m(0)=t_*$, $s_{j,\epsilon}^m(0)=s_*$, $l_{j,\epsilon}^m(\ \cdot\ ,\ \cdot\ ;0)\equiv 0$ and*

$$v(t,s)=M_m\left\{tP^m_{\nu^m_{j,\epsilon}}(\cos\theta)\cos m\varphi+sP^m_{\nu^m_{j,\epsilon}}(\cos\theta)\sin m\varphi\right\}\qquad(t,s\in\mathbf{R})$$

with some constant $M_m>0$.

Since $l_{j,\epsilon}^m(\ \cdot\ ,\ \cdot\ ;\mu)\to 0$ uniformly as $\mu\to 0$ and $P^0_{\nu^0_{0,\epsilon}}(\cos\theta)$ (the first eigenfunction of (1.2)) is positive on $(0,\pi-\epsilon)$, the bifurcating solution $u_{0,\epsilon}^0(\mu)$ is also positive. On the other hand, $u_{j,\epsilon}^m$ with $j\neq 0$ are not positive on $(0,\pi-\epsilon)$. Especially, for each $m\neq 0$, $u_{j,\epsilon}^m(\mu)$ is not radially symmetric since eigenfunctions $P^m_{\nu^m_{j,\epsilon}}(\cos\theta)\cos m\varphi$ and $P^m_{\nu^m_{j,\epsilon}}(\cos\theta)\sin m\varphi$ are not radially symmetric.

In Section 2 we investigate zeros of associated Legendre functions and show Theorem 1.1. In Section 3, by using this result, we prove Theorem 1.2.

2 Zeros of associated Legendre functions and Proof of Theorem 1.1 In this section we prove Theorem 1.1. In arguments below, we set $N = 2$ and $\theta_0 = \pi - \epsilon$. Since our aim in this paper is to investigate (1.1) with a sufficiently small $\epsilon > 0$, it suffices to investigate zeros of $P_\nu^m(x)$ near $x = -1$. In fact the following proposition holds:

Proposition 2.1 *Assume j and m be integers satisfying $0 \le m \le j$. If $j < \nu < j+1$, then $P_\nu^m(x)$ has $j - m + 1$ zeros in $(-1, 1)$. Moreover let $z_j^m(\nu)$ be the smallest zero of $P_\nu^m(x)$ in $(-1, 1)$. Then $z_j^m(\nu) \in C^1((j, j+1))$ and $z_j^m(\nu) \searrow -1$ as $\nu \searrow j$. Furthermore it holds that*

$$z_j^m(\nu) = \begin{cases} -1 + 2(1 + o(1)) \exp\left(-\dfrac{1}{\nu - j}\right) & (m = 0), \\ -1 + (d_{j,m} + o(1))(\nu - j)^{\frac{1}{m}} & (1 \le m \le j), \end{cases} \tag{2.1}$$

where $d_{j,m} = 2[m!(m-1)!(j-m)!/(j+m)!]^{1/m}$ and $o(1) \to 0$ as $\nu \searrow j$.

By the monotonicity of $z_j^m(\nu)$, we obtain a unique solution $\nu = \nu_{j,\epsilon}^m$ to $z_j^m(\nu) = \cos(\pi - \epsilon)$ for any sufficiently small $\epsilon > 0$. Moreover (1.9) follows from (2.1). Therefore, to show Theorem 1.1, we prove Proposition 2.1. Before the proof of Proposition 2.1, we state some preliminaries.

First we define

$$\psi(x) := \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

For $\Gamma(x)$ and $\psi(x)$, the following lemma holds:

Lemma 2.1 (Theorem 2.1.1 in [10]) *Functions $\Gamma(x)$ and $\psi(x)$ are analytic on $\mathbf{R}$ except for non-positive integers, that is, $x = 0, -1, -2, \dots$. Moreover, for an integer $k \ge 0$, it holds that*

$$\lim_{x \to -k} (x+k)\Gamma(x) = \frac{(-1)^k}{k!} \quad \text{and} \quad \lim_{x \to -k} (x+k)\psi(x) = -1.$$

Second we state some properties of $P_\nu^m(x)$. For $P_\nu^m(x)$, the following two lemmas hold:

Lemma 2.2 *It holds that $P_0^0(x) \equiv 1$ and $P_m^m(x) \equiv 0$ in $(-1, 1)$ $(m \ge 1)$. Moreover if $m > \nu$, then $P_\nu^m(x) > 0$ in $(-1, 1)$ (ν is not an integer) or $P_\nu^m(x) \equiv 0$ in $(-1, 1)$ (ν is an integer).*

Lemma 2.3 *Assume that $\nu \ge 0$ and ν is not an integer. If m is an integer satisfying $0 \le m < \nu$, then it holds that, for any $x \in (0, 1)$,*

$$\begin{aligned} P_\nu^m(-1+2x) = {} & \frac{(-1)^m \Gamma(\nu + m + 1)}{\Gamma(\nu - m + 1) m!} x^{\frac{m}{2}} (1-x)^{\frac{m}{2}} \\ & \times \left\{ \frac{(-1)^{m+1}}{\Gamma(-\nu)\Gamma(\nu+1)} \left[P_\nu^m(1-2x) \log x \right. \right. \\ & + \sum_{k=0}^{\infty} \frac{(-\nu+m)_k (\nu+m+1)_k}{(m+1)_k k!} \{\psi(-\nu+m+k) \\ & \left. + \psi(\nu+m+k+1) - \psi(k+1) - \psi(m+k+1)\} x^k \right] \\ & \left. + \frac{(m-1)! m!}{\Gamma(-\nu+m)\Gamma(\nu+m+1)} x^{-m} \sum_{k=0}^{m-1} \frac{(-\nu)_k (\nu+1)_k}{(1-m)_k k!} x^k \right\} \end{aligned}$$

with $(-1)! := 0$.

Concerning proofs of Lemmas 2.2 and 2.3, see Appendix. Lemma 2.3 implies that if ν is not an integer, then $P_\nu^m(x)$ tends to ∞ or $-\infty$ as $x \to -1$.

Next we prove a result on the number of zeros of $P_\nu^m(x)$. For integers $j \geq 0$ and $m = 0, 1, ..., j$, it is known that $P_j^m(x)$ has $j - m$ zeros in $(-1, 1)$ (e.g., see p.246 in Sansone [16]). On the other hand, when ν is not an integer, the following lemma holds:

Lemma 2.4 *Let j and m be integers satisfying $0 \leq m \leq j$. If $\nu \in (j, j+1)$ is fixed, then $P_\nu^m(x)$ has $j - m + 1$ zeros in $(-1, 1)$.*

Proof. Let $m = 0$. Then, from Lemma 2.3, we see that, as $x \to 0$, the leading term of $P_\nu^0(-1+2x)$ is $\log x$. Hence, from $P_\nu^0(1) = 1$ and

$$\Gamma(y)\Gamma(-y+1) = \frac{\pi}{\sin \pi y}, \tag{2.2}$$

(e.g., see p.1 in [15] or Theorem 2.2.3 in Beals and Wong in [10]) it holds that, as $x \to 0$,

$$\begin{aligned} P_\nu^0(-1+2x) &= -\frac{1}{\Gamma(-\nu)\Gamma(\nu+1)} P_\nu^0(1-2x)\log x + o(|\log x|) \\ &= -\frac{\sin \nu\pi}{\pi}\,|\log x| + o(|\log x|). \end{aligned} \tag{2.3}$$

On the other hand, let $1 \leq m \leq j$. Then, from Lemma 2.3, we see that, as $x \to 0$, the leading term of $P_\nu^m(-1+2x)$ is $x^{-m/2}$. Hence, from (2.2), it holds that, as $x \to 0$,

$$\begin{aligned} P_\nu^m(-1+2x) &= \frac{(-1)^m(m-1)!(1-x)^{\frac{m}{2}}}{\Gamma(\nu-m+1)\Gamma(-\nu+m)} x^{-\frac{m}{2}} + o(x^{-\frac{m}{2}}) \\ &= (-1)^{m+1}\frac{(m-1)!\sin(\nu-m)\pi}{\pi} x^{-\frac{m}{2}} + o(x^{-\frac{m}{2}}). \end{aligned} \tag{2.4}$$

Hence, from (2.3), (2.4) and

$$\operatorname{sgn}(\sin(\nu-m)\pi) = (-1)^{j-m} \qquad \text{for } \nu \in (j, j+1),$$

it holds that

$$\operatorname{sgn}(P_\nu^m(-1+2x)) = (-1)^{j+1} \qquad \text{near } x = 0. \tag{2.5}$$

Now the number of zeros for $P_\nu^m(x)$ $(x \in (-1,1))$ is denoted by $\sharp(P_\nu^m)$, and we recall $\sharp(P_j^m) = j - m$. We apply the Sturm–Liouville theorem (e.g., see pp.224–225 in [13]) for (1.4), and, from $\nu \in (j, j+1)$, we see that $\sharp(P_\nu^m) \geq \sharp(P_j^m) = j - m$. Similarly, we compare $P_\nu^m(x)$ and $P_{j+1}^m(x)$, and hence it is proved that $\sharp(P_\nu^m) \leq \sharp(P_{j+1}^m) = j - m + 1$. Thus, for $\nu \in (j, j+1)$, $\sharp(P_\nu^m) = j - m$ or $j - m + 1$.

From (1.6), $\operatorname{sgn}(P_\nu^m(x)) = (-1)^m$ near $x = 1$. If $\sharp(P_\nu^m) = j - m$, then $\operatorname{sgn}(P_\nu^m(x)) = (-1)^j$ near $x = -1$, which is inconsistent with (2.5). Therefore $\sharp(P_\nu^m) = j - m + 1$, and Lemma 2.4 is proved. ■

Now we show Proposition 2.1.

Proof of Proposition 2.1. Fix an integer $j \geq 0$ and an integer $m \in [0, j]$. The number of zeros is already known by Lemma 2.4. Thus it suffices to prove the asymptotic formula (2.1). The following arguments are divided into three steps.

Step 1. We prove that the smallest zero $z_j^m(\nu)$ is of class C^1 with respect to $\nu \in (j, j+1)$.

Let $x = z_j^m(\nu)$ be the smallest zero of $P_\nu^m(x)$ for $\nu \in (j, j+1)$. By Lemma 2.4, the number of zeros for $P_\nu^m(x)$ ($x \in (-1,1)$ and $\nu \in (j, j+1)$) is identically $j - m + 1 \geq 1$. Hence $z_j^m(\nu)$ always exists for $\nu \in (j, j+1)$.

Arbitrarily fix $\nu = \nu_0 \in (j, j+1)$ with $z_0 := z_j^m(\nu_0)$, and we prove that $z_j^m(\nu)$ is of class C^1 near ν_0. In fact, by differentiability with respect to a parameter, $P_\nu^m(x)$ is differentiable with respect to x and ν. If $(P_{\nu_0}^m)_x(z_0) = 0$ holds, then, by the uniqueness of a solution to (1.4), $P_{\nu_0}^m(z_0) = 0$ and $(P_{\nu_0}^m)_x(z_0) = 0$ imply $P_{\nu_0}^m(x) \equiv 0$, and it is a contradiction to $P_{\nu_0}^m(x) \not\equiv 0$. Thus $(P_{\nu_0}^m)_x(z_0) \neq 0$ holds.

Hence, from the implicit function theorem, there exists an implicit function $x = z(\nu)$ such that $P_\nu^m(z(\nu)) = 0$ and $z(\nu)$ is of class C^1 near $\nu = \nu_0$ (for the implicit function theorem, e.g., see Theorem 2.3 in Chapter 2 of [1]). By the uniqueness of the implicit function $z(\nu)$ near $\nu = \nu_0$, $z_j^m(\nu) \equiv z(\nu)$ holds near $\nu = \nu_0$. Therefore, from the arbitrariness of $\nu_0 \in (j, j+1)$, $z_j^m(\nu) \in C^1((j, j+1))$ holds. Step 1 is finished.

Step 2. We prove that $z_j^m(\nu) \searrow -1$ as $\nu \searrow j$.

First we prove that $z_j^m(\nu) \to -1$ as $\nu \searrow j$. We recall the *Gauss hypergeometric function*, and $P_\nu^m(x)$ is expressed by using the function (see (4.1) and (4.3) in Appendix). Since the series (4.1) uniformly converges in any closed interval of $(-1, 1]$, $P_\nu^m(x)$ is of class C^1 as ν in $x \in [-1+\delta, 1]$ with arbitrarily fixed $\delta \in (0, 2)$. Hence, when ν varies sufficiently near j, the number of zeros of $P_\nu^m(x)$ in $[-1+\delta, 1]$ is equal to the number of zeros of $P_j^m(x)$ in $[-1+\delta, 1]$.

Assume that $z_j^m(\nu) \not\to -1$ as $\nu \searrow j$. Then we can take $\delta > 0$ such that

$$z_j^m(\nu) \in [-1+\delta, 1] \qquad \text{as } \nu \searrow j \tag{2.6}$$

and all of zeros of $P_j^m(\nu)$ are contained in $[-1+\delta, 1]$. Since the number of zeros of $P_\nu^m(x)$ does not varies near $\nu = j$, there exists $j - m$ zeros of $P_\nu^m(x)$ in $[-1+\delta, 1)$ near $\nu = j$.

On the other hand, by Lemma 2.4, $P_\nu^m(x)$ has $j - m + 1$ zeros in $(-1, 1)$ if $\nu \in (j, j+1)$. Thus $z_j^m(\nu) \in (-1, -1+\delta)$ for $\nu > j$ near j, and it is inconsistent with (2.6). Therefore $z_j^m(\nu) \to -1$ as $\nu \searrow j$.

Next we show that $z_j^m(\nu) \searrow -1$ as $\nu \searrow j$. Arbitrarily fix a zero z_0 of $P_\nu^m(x)$. By the uniqueness of a solution to (1.4), $P_\nu^m(x) \equiv 0$ if $(P_\nu^m)_x(z_0) = 0$. Hence if $P_\nu^m(x) \not\equiv 0$, then it holds that

$$(P_\nu^m)_x(z_0) \neq 0.$$

On the other hand, a solution ν of $P_\nu^m(z_0) = 0$ is simple (e.g., see p.241 in [13]), that is,

$$(P_\nu^m)_\nu(z_0) \neq 0.$$

Thus, since

$$0 = \frac{d}{d\nu}\left[P_\nu^m(z_j^m(\nu))\right] = (P_\nu^m)_\nu(z_j^m(\nu)) + (P_\nu^m)_x(z_j^m(\nu)) \cdot (z_j^m)_\nu(\nu),$$

it holds that

$$(z_j^m)_\nu(\nu) = -\frac{(P_\nu^m)_\nu(z_j^m(\nu))}{(P_\nu^m)_x(z_j^m(\nu))} \neq 0 \qquad \text{for } \nu \in (j, j+1). \tag{2.7}$$

Since $z_j^m(\nu) \to -1$ as $\nu \searrow j$, there exists some $\nu = \nu_a > j$ such that $(z_j^m)_\nu(\nu_a) > 0$. If there exists some $\nu = \nu_b \in (j, \nu_a)$ such that $(z_j^m)_\nu(\nu_b) < 0$, then, by the continuity of

$(z_j^m)_\nu(\nu)$, there exists some $\nu_c \in [\nu_b, \nu_a]$ such that $(z_j^m)_\nu(\nu_c) = 0$, and it is a contradiction to (2.7). Therefore Step 2 is finished.

Step 3. The asymptotic formula (2.1) is proved.

We define

$$\zeta_j^m(\nu) := \frac{1 + z_j^m(\nu)}{2}.$$

Since $z_j^m(\nu) \searrow -1$ as $\nu \searrow j$, it follows that $\zeta_j^m(\nu) \searrow 0$ as $\nu \searrow j$.

We assume $m = 0$. Then, from Lemma 2.3, it holds that ,

$$-P_\nu^0(1 - 2\zeta_j^0(\nu)) \log \zeta_j^0(\nu) = \psi(-\nu) + \psi(\nu+1) - 2\psi(1) + R_j^0 \zeta_j^0(\nu) \qquad \text{as } \nu \searrow j, \tag{2.8}$$

where

$$R_j^0 = \sum_{k=1}^{\infty} \frac{(-\nu)_k(\nu+1)_k}{(k-1)^2 k!} \{\psi(-\nu+k) + \psi(\nu+k+1) - 2\psi(k+1)\}(\zeta_j^0(\nu))^{k-1}. \tag{2.9}$$

Now we prove

$$|(\nu - j)R_j^0| \le C_0 \qquad \text{as } \nu \searrow j, \tag{2.10}$$

where $C_0 > 0$ is independent of ν near j.

To prove (2.10), we first show that

$$|(\nu - j)\psi(-\nu + k)| \le k + C \qquad \text{for } \nu \in \left(j, j + \frac{1}{2}\right), \tag{2.11}$$

where $C > 0$ is independent of ν and $k \ge 1$. In fact, from Lemma 2.1, we obtain

$$\begin{aligned} \lim_{\nu \searrow j} (\nu - j)\psi(-\nu + k) &= -\lim_{s \nearrow -(j-k)} \{s + (j - k)\}\psi(s) \\ &= \begin{cases} 1 & (k \le j) \\ 0 & (k > j) \end{cases} \end{aligned} \tag{2.12}$$

with an integer k. For $k \le j$, (2.11) immediately follows from (2.12).

On the other hand, we assume $k > j$. From the following equality (e.g., see p.34 in [10])

$$\psi(x+1) = \psi(x) + \frac{1}{x}, \tag{2.13}$$

it follows that

$$\psi(-\nu + k) = \psi(-\nu + j + 1) + \sum_{l=1}^{k-j-1} \frac{1}{-\nu + k - l} \tag{2.14}$$

with $\sum_{l=1}^{0}(-\nu + k - l)^{-1} = 0$. Here $(\nu - j)\psi(-\nu + j + 1) \to 0$ as $\nu \searrow j$. Moreover, from $l \in [1, k - j - 1]$, it holds that

$$\frac{\nu - j}{-\nu + k - l} \le \frac{\nu - j}{-\nu + j + 1} < 1 \qquad \text{for } \nu \in \left(j, j + \frac{1}{2}\right).$$

Hence we obtain (2.11) for $k > j$.

Similarly, from (2.13), it follows that

$$\psi(\nu+k+1)=\psi(\nu+1)+\sum_{l=0}^{k-1}\frac{1}{\nu+k-l}, \tag{2.15}$$

$$\psi(k+1)=\psi(1)+\sum_{l=0}^{k-1}\frac{1}{k-l}. \tag{2.16}$$

Since (2.15) and (2.16) do not have singularity as $\nu\searrow j$, it holds that, for $\nu\in(j,j+2-1)$,

$$|(\nu-j)\psi(\nu+k+1)|\le k+C, \tag{2.17}$$

$$|(\nu-j)\psi(k+1)|\le k+C, \tag{2.18}$$

where $C>0$ is independent of ν and $k\ge 1$.

Let

$$a_k(\nu):=\left|\frac{(-\nu)_k(\nu+1)_k}{(k-1)!k!}\{\psi(-\nu+k)+\psi(\nu+k+1)-2\psi(k+1)\}(\zeta_j^0(\nu))^{k-1}\right|. \tag{2.19}$$

Then, from (2.11), (2.17) and (2.18), it follows that

$$|(\nu-j)a_k(\nu)|^{\frac{1}{k}}\le\left|\frac{(-\nu)_k(\nu+1)_k}{(k-1)!k!}\right|^{\frac{1}{k}}\cdot 4^{\frac{1}{k}}(k+C)^{\frac{1}{k}}\zeta_j^0(\nu). \tag{2.20}$$

Now we prove that

$$\left|\frac{(-\nu)_k(\nu+1)_k}{(k-1)!k!}\right|^{\frac{1}{k}}\to 1 \qquad \text{as } k\to\infty \tag{2.21}$$

and the asymptotic formula (2.21) is uniform with respect to ν sufficiently near j. For the purpose we use the following result (e.g., see Corollary 2.1.4 in [10])

$$\lim_{k\to\infty}\frac{(a)_k}{(b)_k}k^{b-a}=\frac{\Gamma(b)}{\Gamma(a)}, \tag{2.22}$$

where $a,b\ne 0,-1,-2,....$ From (2.22), we obtain

$$\lim_{k\to\infty}\frac{(-\nu)_k(\nu+1)_k}{(k-1)!k!}=\frac{\Gamma(1)}{\Gamma(-\nu)\Gamma(\nu+1)},$$

and hence (2.21) holds. Moreover, since $|\Gamma(-\nu)|\to\infty$ as $\nu\searrow j$, (2.21) is uniform with respect to ν near j.

Therefore, from (2.20), $(k+C)^{1/k}\to 1$ $(k\to\infty)$ and $\zeta_j^0(\nu)\searrow 0$ $(\nu\searrow j)$, we obtain

$$|(\nu-j)a_k(\nu)|\le c^k \qquad \text{for sufficiently near } \nu=j, \tag{2.23}$$

where a constant $c\in(0,1)$ is independent of ν (near j) and k. Hence, by (2.9), (2.19), (2.23) and the majorant test, it holds that

$$|R_j^0(\nu-j)|\le\sum_{k=1}^{\infty}|(\nu-j)a_k(\nu)|\le C_0$$

with some $C_0 > 0$ which is independent of ν near j.

Hence, from $P_\nu^0(1) = 1$, (2.8), (2.10) and (2.12), the singularity coming from $\log \zeta_j^0(\nu)$ must be canceled by $\psi(-\nu)$ as $\nu \searrow j$. Namely we obtain

$$(\nu - j) \log \zeta_j^0(\nu) = -1 + o(1) \qquad \text{as } \nu \searrow j,$$

where $o(1) \to 0$ as $\nu \searrow$. Therefore, from $z_j^m(\nu) = -1 + 2\zeta_j^m(\nu)$, (2.1) holds for $m = 0$.

Next we assume $1 \leq m \leq j$. The proof is similar to that of the case $m = 0$. Namely, from Lemma 2.3 and $P_\nu^m(-1 + 2\zeta_j^m(\nu)) \equiv 0$, it follows that, as $\nu \searrow j$,

$$\begin{aligned} &\frac{m!(m-1)!}{\Gamma(-\nu+m)\Gamma(\nu+m+1)}(1 + L_j^m \zeta_j^m(\nu))(\zeta_j^m(\nu))^{-m} + \frac{(-1)^{m+1} P_\nu^m(1 - 2\zeta_j^m(\nu))}{\Gamma(-\nu)\Gamma(\nu+1)} \log \zeta_j^m(\nu) \\ &= \frac{(-1)^m}{\Gamma(-\nu)\Gamma(\nu+1)m!}\{\psi(-\nu+m) + \psi(\nu+m+1) - \psi(1) - \psi(m+1)\} + R_j^m \zeta_j^m(\nu), \end{aligned}$$

where

$$L_j^m = \sum_{k=1}^{m-1} \frac{(-\nu)_k(\nu+1)_k}{(1-m)_k k!}(\zeta_j^m(\nu))^{k-1},$$

and

$$\begin{aligned} R_j^m = \sum_{k=1}^{\infty} &\frac{(-\nu+m)_k(\nu+m+1)_k}{(m+1)_k k!} \\ &\times \{\psi(-\nu+m+k) + \psi(\nu+m+k+1) - \psi(k+1) - \psi(m+k+1)\}(\zeta_j^m(\nu))^{k-1}. \end{aligned}$$

By similar arguments to the case $m = 0$, we obtain

$$|L_j^m| \leq C, \tag{2.24}$$

$$|(\nu - j)R_j^m| \leq C, \tag{2.25}$$

where $C > 0$ is independent of ν near j (we apply (2.12), (2.13) and (2.22)).

Now we remark that, as $\nu \searrow j$, it holds that $(\zeta_j^m(\nu))^m \log \zeta_j^m(\nu) = o(1)$. Hence, from (2.12), (2.24), (2.25) and $P_\nu^m(1) = 0$, the singularity coming from $(\zeta_j^m(\nu))^{-m}$ must be canceled by $\psi(-\nu + m)$. Namely we obtain, as $\nu \searrow j$,

$$(\nu-j)(1+o(1))(\zeta_j^m(\nu))^{-m} = \frac{(-1)^m \Gamma(-\nu+m)\Gamma(\nu+m+1)}{m!(m-1)!\Gamma(-\nu)\Gamma(\nu+1)}(\nu-j)\psi(-\nu+m) + o(1), \tag{2.26}$$

where $o(1) \to 0$ as $\nu \searrow j$. From Lemma 2.1, it holds that

$$\lim_{\nu \searrow j} \frac{\Gamma(-\nu+m)}{\Gamma(-\nu)} = \frac{(-1)^{j-m+1}}{(j-m)!} \times \frac{j!}{(-1)^{j+1}} = \frac{(-1)^{-m} j!}{(j-m)!}. \tag{2.27}$$

Thus, from (2.12), (2.26) and (2.27) and $\Gamma(j + m + 1) = (j + m)!$, it holds that

$$\lim_{\nu \searrow j} (\nu - j)(\zeta_j^m(\nu))^{-m} = \frac{(j+m)!}{m!(m-1)!(j-m)!}$$

Therefore, from $z_j^m(\nu) = -1 + 2\zeta_j^m(\nu)$, we obtain (2.1) with $1 \leq m \leq j$. Now all of steps are finished, and Proposition 2.1 is completely shown. ■

Theorem 1.1 follows from Proposition 2.1.

Proof of Theorem 1.1. We prove the existence of ν satisfying

$$P_\nu^m(\cos(\pi-\epsilon))=0$$

and investigate its behavior as $\epsilon\to 0$. Here we take $z_j^m(\nu)$ which is the same definition as in Proposition 2.1.

Since $z_j^m(\nu)\searrow -1$ as $\nu\searrow j$ (see Proposition 2.1), the equation

$$z_j^m(\nu)=\cos(\pi-\epsilon) \tag{2.28}$$

has a unique solution $\nu=\nu_{j,\epsilon}^m$ for a sufficiently small $\epsilon>0$. Moreover $\lambda_{j,\epsilon}^m:=\nu_{j,\epsilon}^m(\nu_{j,\epsilon}^m+1)$ is an eigenvalue of (1.2).

Next if $m\geq\nu$, then, from Lemma 2.2, $P_\nu^m(\cos\theta)$ does not have a zero in $(0,\pi-\epsilon)$ or $P_\nu^m(\cos\theta)\equiv 0$. Therefore, for $\nu\in(j,j+1)$, there exist exactly $(j+1)$ eigenvalues.

Finally we show (1.9). From (2.28), it follows that

$$z_j^m(\nu_{j,\epsilon}^m)=-1+\frac{1}{2}\epsilon^2+O(\epsilon^4)\quad\text{as }\epsilon\to 0.$$

Thus, from Proposition 2.1, it holds that

$$\frac{1}{2}\epsilon^2+O(\epsilon^4)=\begin{cases}2(1+o(1))\exp\left(-\dfrac{1}{\nu_{j,\epsilon}^m-j}\right) & (m=0),\\ (d_{j,m}+o(1))(\nu_{j,\epsilon}^m-j)^{\frac{1}{m}} & (1\leq m\leq j),\end{cases}$$

where $o(1)\to 0$ as $\epsilon\to 0$. Hence, as $\epsilon\to 0$, we obtain

$$\nu_{j,\epsilon}^m=\begin{cases}j+\dfrac{1}{2|\log\epsilon|}+o\left(\dfrac{1}{|\log\epsilon|}\right) & (m=0),\\ j+c_{j,m}\epsilon^{2m}+o(\epsilon^{2m}) & (1\leq m\leq j),\end{cases}$$

where $c_{j,m}=(2d_{j,m})^{-m}=(j+m)!/[4^m m!(m-1)!(j-m)!]$. Recall $\lambda=\nu(\nu+1)$ (see (1.3)), and (1.9) is shown. Now Theorem 1.1 is proved. ■

3 Proof of Theorem 1.2 In Section 2, we investigated eigenvalues $\{\lambda_{j,\epsilon}^m\}$ of (1.2). In this section we prove that $\{\lambda_{j,\epsilon}^m\}$ are bifurcation points of (1.1).

We use the Lyapunov–Schmidt reduction method, that is, we consider our problem by dividing $C^{2,\alpha}(B_{\pi-\epsilon})$ into the eigenspace corresponding to $\lambda_{j,\epsilon}^m$ and its orthogonal complement.

We introduce the Banach spaces

$$\mathcal{X}:=\{u\in C^{2,\alpha}(B_{\pi-\epsilon})\mid u=0\text{ on }\partial B_{\pi-\epsilon}\},$$
$$\mathcal{Y}:=C^\alpha(B_{\pi-\epsilon}).$$

Then the following function

$$f(\lambda,u):=\Delta_{\mathbf{S}^2}u+\lambda u+|u|^{p-1}u$$

satisfies $f\in C^1(\mathbf{R}\times\mathcal{X};\mathcal{Y})$ for $1<p<\infty$. We see that $u\equiv 0$ is a solution to $f(\lambda,u)=0$. In arguments below, we show that there exists a non-trivial solution $u\in\mathcal{X}$ to $f(\lambda,u)=0$ near $\lambda=\lambda_{j,\epsilon}^m$ and $u\equiv 0$.

For convenience let $\lambda_* := \lambda^m_{j,\epsilon}$, and we define

$$\begin{aligned}
L &:= f_u(\lambda_*, 0) = \Delta_{\mathbf{S}^2} + \lambda_*,\\
\mathcal{V} &:= \mathrm{Ker}(L),\\
\mathcal{R} &:= \mathrm{Ran}(L),
\end{aligned}$$

where $\mathrm{Ker}(L)$ and $\mathrm{Ran}(L)$ denote the kernel of L and the range of L, respectively. From Theorem 1.1, the dimension of $\mathcal{V}$ is 1 or 2. Moreover $L : \mathcal{X} \to \mathcal{Y}$ is continuous, and $\mathcal{V}$ and $\mathcal{R}$ are closed subspaces of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Moreover we define the following inner product

$$\langle u, v\rangle := \int_0^{\pi-\epsilon}\int_0^{2\pi} uv \sin\theta d\theta d\varphi \qquad (u, v \in \mathcal{Y}).$$

For this inner product, the orthogonal complement $\mathcal{W}$ of $\mathcal{V}$ is defined, and it holds that

$$\mathcal{X} = \mathcal{V} \oplus \mathcal{W}.$$

We remark that $L : \mathcal{W} \to \mathcal{R}$ is one-to-one and onto.

On the other hand, as for $\mathcal{R}$, the following lemma holds:

Lemma 3.1 *Let $u \in \mathcal{Y}$. Then u belongs to $\mathcal{R} \subset \mathcal{Y}$ if and only if*

$$\langle u, v\rangle = 0 \qquad \text{for any } v \in \mathcal{V}.$$

Lemma 3.1 is proved in Appendix. Lemma 3.1 implies that $\mathcal{Y}$ is expressed as

$$\mathcal{Y} = \mathcal{R} \oplus \mathcal{V}. \tag{3.1}$$

By (3.1), we define orthogonal projections

$$\begin{aligned}
Q &: \mathcal{Y} \to \mathcal{R},\\
P &: \mathcal{Y} \to \mathcal{V}.
\end{aligned}$$

From preliminaries above we show the existence of non-trivial solutions to $f(\lambda, u) = 0$. Let $\mu := \lambda - \lambda_*$, and we seek a solution whose form is

$$u = |\mu|^{\frac{1}{p-1}}(v + w) \qquad \text{with } v \in \mathcal{V} \text{ and } w \in \mathcal{W}. \tag{3.2}$$

Namely we solve

$$f(\mu + \lambda_*, |\mu|^{\frac{1}{p-1}}(v + w)) = 0,$$

We define

$$\begin{aligned}
h(\mu, v, w) &:= \frac{1}{|\mu|^{\frac{1}{p-1}}} f(\mu + \lambda_*, |\mu|^{\frac{1}{p-1}}(v + w))\\
&= Lw + \mu(v + w) + |\mu||v + w|^{p-1}(v + w)
\end{aligned}$$

For $\mu \neq 0$, $h(\mu, v, w) = 0$ is equivalent to $f(\mu + \lambda_*, \mu^{1/(p-1)}(v + w)) = 0$.

In fact we can find a non-trivial solution to $h(\mu, v, w) = 0$ for $\mu < 0$, where

$$h(\mu, v, w) = Lw + \mu(v + w) - \mu|v + w|^{p-1}(v + w).$$

On the other hand, by arguments below, we cannot find a non-trivial solution to $f(\mu + \lambda_*, u) = 0$ with $\mu > 0$ (in fact, for $\mu > 0$, we only obtain a trivial solution by the implicit function theorem in proofs of Proposition 3.2 or 3.3 below). Therefore we only consider the case $\mu \le 0$.

By orthogonal projections Q and P, $h(\mu, v, w) = 0$ is equivalent to the following simultaneous equations

$$Qh(\mu, v, w) = 0, \tag{3.3}$$

$$Ph(\mu, v, w) = 0. \tag{3.4}$$

Arguments below are divided into the following two steps:

(s1) we show that there exists a function $l(\mu, v)$ such that $Qh(\mu, v, l(\mu, v)) = 0$,

(s2) we solve $Ph(\mu, v, l(\mu, v)) = 0$ and find a non-trivial solution $v = v(\mu)$.

Step (s1). Let

$$J(\mu, v, w) := Qh(\mu, v, w) = Lw + \mu w - \mu Q|v + w|^{p-1}(v + w). \tag{3.5}$$

For $J(\mu, v, w)$, the following proposition holds:

Proposition 3.1 *Assume $v_* \in \mathcal{V}$. Then there exists $l \in C^1((-\delta_0, 0] \times \mathcal{V}_*; \mathcal{W}_0)$ such that $J(\mu, v, w) = 0$ implies $w = l(\mu, v)$, where $\delta_0 > 0$ is some constant, $\mathcal{V}_* \subset \mathcal{V}$ and $\mathcal{W}_0 \subset \mathcal{W}$ are neighborhoods of $v_* \in \mathcal{V}$ and $0 \in \mathcal{W}$, respectively. Namely $l(0, v_*) = 0$ holds.*

Proof. From (3.5), we obtain

$$J(0, v_*, 0) = 0,$$

and

$$J_w(0, v_*, 0)[\xi] = L\xi \qquad \text{for any } \xi \in \mathcal{W}.$$

Therefore, by the implicit function theorem, Proposition 3.1 holds. ■

Here we remark that v_* is arbitrarily fixed, and $l(\mu, v)$ depends on $v_* \in \mathcal{V}$. In addition, for $l(\mu, v)$, the following Lemma 3.2 holds:

Lemma 3.2 *For $l(\mu, v)$ defined in Proposition* 3.1*, it holds that*

$$l_v(0, v_*) = 0.$$

Proof. From $J(\mu, v, l(\mu, v)) = 0$ and (3.5), it holds that

$$Ll(\mu, v) + \mu l(\mu, v) - \mu Q|v + l(\mu, v)|^{p-1}(v + l(\mu, v)) = 0. \tag{3.6}$$

Differentiating (3.6) by v, we obtain

$$Ll_v(\mu, v)[\xi] + \mu l_v(\mu, v)[\xi] - \mu p Q|v + l(\mu, v)|^{p-1}(\xi + l_v(\mu, v)[\xi]) = 0 \qquad \text{for any } \xi \in \mathcal{V}.$$

Hence, by substituting $\mu = 0$ and $v = v_*$, we obtain

$$Ll_v(0, v_*)[\xi] = 0 \qquad \text{for any } \xi \in \mathcal{V}. \tag{3.7}$$

From (3.7) and Proposition 3.1, it follows that $l_v(0, v_*)[\xi] \in \mathcal{V} \cap \mathcal{W} = \{0\}$ for any $\xi \in \mathcal{V}$. Therefore, from the arbitrariness of $\xi \in \mathcal{V}$, Lemma 3.2 is shown. ■

Lemma 3.2 is required in arguments below.

Step (s2). Next we consider (3.4). Hereafter let $w = l(\mu, v)$ which is defined in Proposition 3.1. Let

$$K(\mu, v) := Ph(\mu, v, l(\mu, v)) = \mu v - \mu P|v + l(\mu, v)|^{p-1}(v + l(\mu, v)). \tag{3.8}$$

We show the existence of $v = v(\mu) \not\equiv 0$ satisfying $K(\mu, v) = 0$ by dividing a proof into two cases, that is, the dimension of $\mathcal{V}$ is 1 or 2.

First we consider the case that the dimension of $\mathcal{V}$ is 1. Then, from Theorem 1.1, it follows that

$$\mathcal{V} = \{tP_\nu^0(\cos\theta) \mid t \in \mathbf{R}\}, \tag{3.9}$$

where $\nu_{j,\epsilon}^m$ is abbreviated to ν. Then the following proposition holds.

Proposition 3.2 *Assume* $1 < p < \infty$ *and* $\nu = \nu_{j,\epsilon}^0$. *Let* $M_0 > 0$ *satisfy*

$$\int_0^{\pi-\epsilon} \left\{|P_\nu^0(\cos\theta)|^2 - M_0^{p-1}|P_\nu^0(\cos\theta)|^{p+1}\right\} \sin\theta d\theta = 0. \tag{3.10}$$

Then there exist a constant $\delta > 0$ *and a* C^1*-function* $t(\mu)$ *such that*

$$K(\mu, t(\mu)M_0P_\nu^m(\cos\theta)) = 0 \qquad \text{for } \mu \le 0,$$

where $t(0) = 1$ *and* $|t(\mu) - 1| + |\mu| < \delta$.

Proof. Let

$$v_* := M_0P_\nu^0(\cos\theta).$$

Then, by Proposition 3.1, there exists an implicit function $w = l(\mu, v)$ satisfying (3.3) and $l(0, v_*) = 0$. From (3.8) and (3.9), it follows that

$$\begin{aligned} K(\mu, tv_*) &= \mu a_\nu^0 N(\mu, t)P_\nu^0(\cos\theta) \\ N(\mu, t) &:= \langle P_\nu^0(\cos\theta), h(\mu, tv_*, l(\mu, tv_*))\rangle \\ &= \langle P_\nu^0(\cos\theta), tv_* - |tv_* + l(\mu, tv_*)|^{p-1}(tv_* + l(\mu, tv_*)[v_*])\rangle. \end{aligned}$$

Here $(a_\nu^0)^{-1} = 2\pi \int_0^{\pi-\epsilon} |P_\nu^0(\cos\theta)|^2 \sin\theta d\theta$. We remark $N(\mu, t) = 0$ ($\mu \ne 0$) is equivalent to $K(\mu, tv_*) = 0$. We show that there exists some $t(\mu)$ satisfying $N(\mu, t(\mu)) = 0$ by the implicit function theorem. For the purpose it suffices to prove that $N(0, 1) = 0$ and $N_t(0, 1) \ne 0$.

From (3.10) and $l(0, v_*) = 0$, we obtain

$$\begin{aligned} N(0, 1) = 2\pi M_0 &\int_0^{\pi-\epsilon} \left\{|P_\nu^0(\cos\theta)|^2 - M_0^{p-1}|P_\nu^0(\cos\theta)|^{p+1}\right\} \sin\theta d\theta \\ = 0&. \end{aligned}$$

Moreover, from direct calculation, it follows that

$$N_t(\mu, t) = \langle P_\nu^0(\cos\theta), v_* - p|tv_* + l(\mu, tv_*)|^{p-1}(v_* + l_v(\mu, tv_*)[v_*])\rangle.$$

Hence, from $l_v(0, v_*)$ (see Lemma 3.2) and (3.10), we obtain

$$\begin{aligned} N_t(0, 1) &= 2\pi M_0 \int_0^{\pi-\epsilon} \left\{|P_\nu^0(\cos\theta)|^2 - pM_0^{p-1}|P_\nu^0(\cos\theta)|^{p+1}\right\} \sin\theta d\theta \\ &= -2\pi(p-1)M_0^p \int_0^{\pi-\epsilon} |P_\nu^0(\cos\theta)|^{p+1} \sin\theta d\theta < 0. \end{aligned}$$

Therefore, by the implicit function theorem, there exist a constant $\delta > 0$ and a C^1-function $t(\mu)$ such that

$$N(\mu, t(\mu)) = 0 \qquad \text{for } \mu \leq 0,$$

where $t(0) = 1$ and $|t(\mu) - 1| + |\mu| < \delta$. Since $K(\mu, t(\mu)v_*) = 0$ is equivalent to $N(\mu, t(\mu)) = 0$ $(\mu \neq 0)$ and $t(0) = 1$, Proposition 3.2 is proved. ■

Second we consider the case that the dimension of $\mathcal{V}$ is 2. Then, from Theorem 1.1, it follows that

$$\mathcal{V} = \{P_\nu^m(\cos\theta)(t\cos m\varphi + s\sin m\varphi) \mid t, s \in \mathbf{R}\} \qquad \text{with } 1 \leq m \leq j. \tag{3.11}$$

Then, for $K(\mu, v)$ defined in (3.8), the following lemma holds:

Proposition 3.3 *Assume* $1 < p < \infty$, $1 \leq m \leq j$ *and* $\nu = \nu_{j,\epsilon}^m$. *Let* $M_m > 0$ *satisfy*

$$\begin{aligned} \int_0^{\pi-\epsilon} \Big\{ & |P_\nu^m(\cos\theta)|^2 \\ & -M_m^{p-1}\left[\frac{1}{\pi}\int_0^{2\pi} |\cos m\varphi|^{p+1} d\varphi\right] |P_\nu^m(\cos\theta)|^{p+1}\Big\} \sin\theta d\theta = 0. \end{aligned} \tag{3.12}$$

Then there exist a constant $\delta > 0$ *and* C^1*-functions* $t(\mu), s(\mu)$ *such that*

$$K(\mu, t(\mu)M_m P_\nu^m(\cos\theta)\cos m\varphi + s(\mu)M_m P_\nu^m(\cos\theta)\cos m\varphi) = 0 \quad \text{for } \mu \leq 0,$$

where $0 \leq m \leq j$, $t(0) = t_*$ $s(0) = s_*$ *and* $|t(\mu) - t_*| + |s(\mu) - s_*| + |\mu| < \delta$. *Here* $t_*, s_* \in \mathbf{R}$ *is arbitrarily taken such that* $t_*^2 + s_*^2 = 1$, $t_* \neq 0$ *and* $s_* \neq 0$.

Proof. Let

$$v(t, s) := M_m\left(tP_\nu^m(\cos\theta)\cos m\varphi + sP_\nu^m(\cos\theta)\sin m\varphi\right).$$

Moreover we arbitrary fix t_* and s_* satisfying $t_*^2 + s_*^2 = 1$. In addition let $v_* := v(t_*, s_*)$. Then, by Proposition 3.1, there exists an implicit function $l(\mu, v)$ for v_*.

From (3.8) and (3.11), $K(\mu, v)$ is expressed as

$$K(\mu, v) = \mu a_\nu^m\{N_1(\mu, t)P_\nu^m(\cos\theta)\cos m\varphi + N_2(\mu, t)P_\nu^m(\cos\theta)\sin m\varphi\},$$

and

$$N_1(\mu, t, s) := \langle P_\nu^m(\cos\theta)\cos m\varphi, v - |v + l(\mu, v)|^{p-1}(v + l(\mu, v))\rangle, \tag{3.13}$$

$$N_2(\mu, t, s) := \langle P_\nu^m(\cos\theta)\sin m\varphi, v - |v + l(\mu, v)|^{p-1}(v + l(\mu, v))\rangle, \tag{3.14}$$

where $(a_\nu^m)^{-1} = \int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2 \sin\theta d\theta \int_0^{2\pi} |\cos m\varphi|^2 d\varphi = \pi\int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2 \sin\theta d\theta$. The equation $K(\mu, v) = 0$ $(\mu \neq 0)$ is equivalent to

$$N_1(\mu, t, s) = N_2(\mu, t, s) = 0. \tag{3.15}$$

Thus it suffices to show the existence of non-trivial solutions $t = t(\mu)$ and $s = s(\mu)$ to (3.15), and we prove it by using the implicit function theorem.

From direct calculation, it follows that

$$
\begin{aligned}
&N_1(0,t_*,s_*)\\
&= M_m \int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2 \sin\theta d\theta \int_0^{2\pi} (t_*\cos^2 m\varphi + s_*\sin m\varphi\cos m\varphi)d\varphi\\
&\quad - M_m^p \int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^{p+1}\sin\theta d\theta\\
&\quad \times \int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}(t_*\cos^2 m\varphi + s_*\sin m\varphi\cos m\varphi)d\varphi.
\end{aligned}
\tag{3.16}
$$

We see that $\int_0^{2\pi}\cos^2 m\varphi d\varphi = \pi$ and $\int_0^{2\pi}\sin m\varphi\cos m\varphi d\varphi = 0$. Moreover let

$$\cos\beta := t_* \quad \text{and} \quad \sin\beta := s_*.$$

In addition, since $|\cos m\varphi|^{p-1}\cos m\varphi\sin m\varphi$ is odd and periodic, it holds that

$$\int_0^{2\pi} |\cos m\varphi|^{p-1}\cos m\varphi\sin m\varphi d\varphi = \int_{-\pi}^{\pi} |\cos m\varphi|^{p-1}\cos m\varphi\sin m\varphi d\varphi = 0. \tag{3.17}$$

Hence, from (3.17), it holds that

$$
\begin{aligned}
&\int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}(t_*\cos^2 m\varphi + s_*\sin m\varphi\cos m\varphi)d\varphi\\
&= \int_0^{2\pi} |\cos(m\varphi-\beta)|^{p-1}\cos(m\varphi-\beta)\cos m\varphi d\varphi\\
&= \int_0^{2\pi} |\cos m\varphi|^{p-1}\cos m\varphi\cos(m\varphi+\beta)d\varphi\\
&= \int_0^{2\pi} |\cos m\varphi|^{p-1}\cos m\varphi(t_*\cos m\varphi - s_*\sin m\varphi)d\varphi\\
&= t_* D_p\pi + s_*\int_0^{2\pi} |\cos m\varphi|^{p-1}\cos m\varphi\sin m\varphi d\varphi\\
&= t_* D_p \pi
\end{aligned}
$$

with

$$D_p := \frac{1}{\pi}\int_0^{2\pi} |\cos m\varphi|^{p+1}d\varphi = \frac{1}{\pi}\int_0^{2\pi}|\cos\varphi|^{p+1}d\varphi \quad (m\geq 1).$$

Thus, from (3.12) and (3.16), we obtain

$$
\begin{aligned}
N_1(0,t_*,s_*) &= t_* M_m\pi \int_0^{\pi-\epsilon}\left\{|P_\nu^m(\cos\theta)|^2 - M_m^{p-1}D_p|P_\nu^m(\cos\theta)|^{p+1}\right\}\sin\theta d\theta\\
&= 0.
\end{aligned}
$$

Similarly, since it holds that

$$\int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}(t_*\cos m\varphi + s_*\sin m\varphi)\sin m\varphi d\varphi = s_* D_p\pi,$$

we obtain

$$
\begin{aligned}
N_2(0,t_*,s_*) &= s_* M_m\pi \int_0^{\pi-\epsilon}\left\{|P_\nu^m(\cos\theta)|^2 - M_m^{p-1}D_p|P_\nu^m(\cos\theta)|^{p+1}\right\}\sin\theta d\theta\\
&= 0.
\end{aligned}
$$

Next, from direct calculation, it holds that

$$\begin{aligned}(N_1)_t(\mu,t,s) =&\langle P_\nu^m(\cos\theta)\cos m\varphi, M_m P_\nu^m(\cos\theta)\cos m\varphi \\ &- pM_m|v+l(\mu,v)|^{p-1}(1+l_v(\mu,v))P_\nu^m(\cos\theta)\cos m\varphi\rangle.\end{aligned}$$

Thus it holds that

$$\begin{aligned}(N_1)_t(0,t_*,s_*) = M_m\pi &\int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2 \sin\theta d\theta \\ &- pM_m^p\pi \int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^{p+1}\sin\theta d\theta \\ &\times \int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}\cos^2 m\varphi d\varphi.\end{aligned} \tag{3.18}$$

Here, from (3.17), it follows that

$$\begin{aligned}&\int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}\cos^2 m\varphi d\varphi \\ &= \int_0^{2\pi} |\cos m\varphi|^{p-1}\cos^2(m\varphi+\beta)d\varphi \\ &= \int_0^{2\pi} |\cos m\varphi|^{p-1}(t_*\cos m\varphi - s_*\sin m\varphi)^2 d\varphi \\ &= t_*^2\int_0^{2\pi} |\cos m\varphi|^{p-1}\cos^2 m\varphi d\varphi + s_*^2\int_0^{2\pi} |\cos m\varphi|^{p-1}\sin^2 m\varphi d\varphi.\end{aligned}$$

Since

$$\begin{aligned}&\int_0^{2\pi} |\cos m\varphi|^{p-1}\sin^2 m\varphi d\varphi \\ &= \left[-\frac{1}{mp}|\cos m\varphi|^{p-1}\cos m\varphi\sin m\varphi\right]_0^{2\pi} + \frac{1}{p}\int_0^{2\pi}|\cos m\varphi|^{p+1}d\varphi \\ &= \frac{1}{p}\int_0^{2\pi}|\cos m\varphi|^{p+1}d\varphi,\end{aligned}$$

we obtain

$$\int_0^{2\pi} |t_*\cos m\varphi + s_*\sin m\varphi|^{p-1}\cos^2 m\varphi d\varphi = D_p\pi\left(t_*^2+\frac{s_*^2}{p}\right).$$

Hence, from (3.18), it holds that

$$\begin{aligned}(N_1)_t(0,t_*,s_*) = M_m\pi &\int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2\sin\theta d\theta \\ &- \left(pt_*^2+s_*^2\right)M_m^p D_p\pi\int_0^{\pi-\epsilon}|P_\nu^m(\cos\theta)|^{p+1}\sin\theta d\theta.\end{aligned}$$

From (3.12) and $pt_*^2+s_*^2>1$, we obtain $(N_1)_t(0,t_*,s_*)<0$.

On the other hand, it holds that

$$\begin{aligned}(N_1)_s(\mu,t,s) =&\langle P_\nu^m(\cos\theta)\cos m\varphi, M_m P_\nu^m(\cos\theta)\sin m\varphi \\ &- p|v+l(\mu,v)|^{p-1}(1+l_v(\mu,v))M_m P_\nu^m(\cos\theta)\sin m\varphi\rangle.\end{aligned}$$

Thus, by similar calculation above, we obtain

$$
\begin{aligned}
(N_1)_s(0,t_*,s_*) &= M_m\pi\int_0^{\pi-\epsilon}|P_\nu^m(\cos\theta)|^2\sin\theta d\theta\int_0^{2\pi}\cos m\varphi\sin m\varphi d\varphi\\
&\quad - pM_m^pD_p\pi\int_0^{\pi-\epsilon}|P_\nu^m(\cos\theta)|^{p+1}\sin\theta d\theta\\
&\quad \times\int_0^{2\pi}|t_*\cos m\varphi+s_*\sin m\varphi|^{p-1}\cos m\varphi\sin m\varphi d\varphi\\
&= 0.
\end{aligned}
$$

Similarly, since (3.12) and $t_*^2+ps_*^2>1$, it follows that

$$
\begin{aligned}
(N_2)_t(0,t_*,s_*) &= 0,\\
(N_2)_s(0,t_*,s_*) &= M_m\pi\int_0^{\pi-\epsilon}|P_\nu^m(\cos\theta)|^2\sin\theta d\theta\\
&\quad - \left(t_*^2+ps_*^2\right)M_m^pD_p\pi\int_0^{\pi-\epsilon}|P_\nu^m(\cos\theta)|^{p+1}\sin\theta d\theta\\
&< 0.
\end{aligned}
$$

Therefore, by the implicit function theorem, there exists a constant $\delta>0$ and C^1-functions $t(\mu)$ and $s(\mu)$ such that, for $|t(\mu)-t_*|+|s(\mu)-s_*|+|\mu|<\delta$, it holds that

$$
N_1(\mu,t(\mu),s(\mu)) = N_2(\mu,t(\mu),s(\mu)) = 0 \tag{3.19}
$$

with $t(0)=t_*$ and $s(0)=s_*$. Since $K(\mu,t(\mu)v(t(\mu),s(\mu)))=0$ is equivalent to (3.19) with $\mu\neq 0$, Proposition 3.3 is proved. ■

Proposition 3.3 is considered in the case of $t_*\neq 0$ and $s_*\neq 0$. If $t_*=0$ (or $s_*=0$), then $(N_1)_t(0,0,1)=0$ (or $(N_2)_s(0,1,0)=0$), and hence the argument in Proposition 3.3 is not valid (we cannot apply the implicit function theorem). Thus we show Proposition 3.3 in the case of $t_*=0$ or $s_*=0$ by another method. Namely we prove that $N_1(\mu,0,s)\equiv 0$ ($N_2(\mu,t,0)\equiv 0$) holds if $t_*=0$ ($s_*=0$), and we follows the same argument as in the proof of Proposition 3.2.

We prepare for a proof of the case $t_*=0$ or $s_*=0$. For $J(\mu,v,w)$ defined in (3.5), we remark that

$$
J(\mu,-v,-w) = -Lw-\mu w-\mu Q|v+w|^{p-1}(-v-w) = -J(\mu,v,w). \tag{3.20}
$$

From Proposition 3.1 and (3.20), it holds that

$$
J(\mu,-v,-l(\mu,v)) = J(\mu,v,l(\mu,v)) = 0.
$$

Hence we extend $l(\mu,v)$ for $-v$ ($v\in\mathcal{V}_*$) by

$$
l(\mu,-v) := -l(\mu,v). \tag{3.21}
$$

and then $J(\mu,-v,l(\mu,-v))=0$ holds. Thus, from (3.8) and (3.21), it follows that

$$
K(\mu,-v) = Ph(\mu,-v,l(\mu,-v)) = -Ph(\mu,v,l(\mu,v)) = -K(\mu,v).
$$

Hence if $K(\mu,v)=0$ holds for the extended $l(\mu,v)$, then $K(\mu,-v)=0$ also holds.

Now we prove the following lemma.

Lemma 3.3 *Assume the same assumption as in Proposition* 3.3 *and extend* $l(\mu, v)$ *by* (3.21). *If* $t_* = 0$ ($s_* = 0$) *for* $N_1(\mu, t, s)$ *and* $N_2(\mu, t, s)$ *as in* (3.13) *and* (3.14), *then it holds that* $N_1(\mu, 0, s) \equiv 0$ ($N_2(\mu, t, 0) \equiv 0$) *near* $\mu = 0$ *and* $s = 1$ ($\mu = 0$ *and* $t = 1$).

Proof. First we assume $t_* = 0$. Let $v_1 = v_1(\theta, \varphi; s) := sM_m P_\nu^m(\cos\theta)\sin m\varphi$. Then, from $\int_0^{2\pi} \cos m\varphi \sin m\varphi d\varphi = 0$, it follows that

$$\begin{aligned} N_1(\mu, 0, s) &= sM_m \int_0^{\pi-\epsilon} |P_\nu^m(\cos\theta)|^2 \sin\theta d\theta \int_0^{2\pi} \cos m\varphi \sin m\varphi d\varphi \\ &\quad - M_m^p \int_0^{\pi-\epsilon} \left\{ \int_0^{2\pi} V_1(\theta, \varphi; s) \cos m\varphi d\varphi \right\} P_\nu^m(\cos\theta) \sin\theta d\theta \\ &= -M_m^p \int_0^{\pi-\epsilon} \left\{ \int_0^{2\pi} V_1(\theta, \varphi; s) \cos m\varphi d\varphi \right\} P_\nu^m(\cos\theta) \sin\theta d\theta \end{aligned}$$

with

$$V_1(\theta, \varphi; s) := |v_1(\theta, \varphi; s) + l(\mu, v_1(\theta, \varphi; s))|^{p-1}(v_1(\theta, \varphi; s) + l(\mu, v_1(\theta, \varphi; s)))$$

The function $v_1(\theta, \varphi; s)$ is odd and periodic with respect to φ. Moreover, from (3.21), it holds that

$$\begin{aligned} l(\mu, v_1(\theta, -\varphi; s)) &= l(\mu, sM_m P_\nu^m(\cos\theta) \sin m(-\varphi)) \\ &= l(\mu, -sM_m P_\nu^m(\cos\theta) \sin m\varphi) \\ &= -l(\mu, v_1(\theta, \varphi; s)). \end{aligned}$$

Hence $V_1(\theta, \varphi; s)$ is odd and periodic with respect to φ, and it holds that

$$\int_0^{2\pi} V_1(\theta, \varphi; s) \cos m\varphi d\varphi = \int_{-\pi}^{\pi} V_1(\theta, \varphi; s) \cos m\varphi d\varphi = 0.$$

Therefore we obtain

$$N_1(\mu, 0, s) \equiv 0 \qquad \text{near } (\mu, s) = (0, 1).$$

Next we assume $s_* = 0$. Let $v_2 = v_2(\theta, \varphi; t) := tM_m P_\nu^m(\cos\theta)\cos m\varphi$. Then, from $\int_0^{2\pi} \cos m\varphi \sin m\varphi d\varphi = 0$, it follows that

$$N_2(\mu, t, 0) = -M_m^p \int_0^{\pi-\epsilon} \left\{ \int_0^{2\pi} V_2(\theta, \varphi; s) \sin m\varphi d\varphi \right\} P_\nu^m(\cos\theta) \sin\theta d\theta$$

with

$$V_2(\theta, \varphi; t) := |v_2(\theta, \varphi; s) + l(\mu, v_2(\theta, \varphi; s))|^{p-1}(v_2(\theta, \varphi; s) + l(\mu, v_2(\theta, \varphi; s)))$$

Since $v_2(\theta, \varphi; t)$ is even and periodic with respect to φ, $l(\mu, v_2(\theta, \varphi; t))$ is also even and periodic with respect to φ. Therefore, since $\cos m\varphi \sin m\varphi$ and $V_2(\theta, \varphi; t) \sin m\varphi$ are odd and periodic with respect to φ, we obtain

$$N_2(\mu, t, 0) \equiv 0 \qquad \text{near } (\mu, t) = (0, 1).$$

Lemma 3.3 is proved. ■

Lemma 3.3 implies that if $\mathcal{V}$ is restricted to $\mathcal{V}_c := \{tP_\nu^m(\cos\theta)\cos m\varphi \mid t \in \mathbf{R}\}$ (or $\mathcal{V}_s := \{tP_\nu^m(\cos\theta)\sin m\varphi \mid t \in \mathbf{R}\}$), then the dimension of the range of $P(\mu, v, l(\mu, v))$ ($v \in \mathcal{V}_c$) is 1. Hence, by similar arguments to Proposition 3.2, we can prove the existence of a non-trivial solution:

Proposition 3.4 *Assume the same assumptions as in Proposition* 3.3. *Then, for* $t_* = 0$ *or* $s_* = 0$, *there exists a nontrivial solution of* $K(\mu, v) = 0$.

Proof. First we assume $s_* = 0$, and then $t_* = 1$ ($t_*^2 + s_*^2 = 1$). We define

$$\mathcal{V}_c := \{tP_\nu^m(\cos\theta)\cos m\varphi \mid t \in \mathbf{R}\}.$$

For $\mathcal{V}_c$, we define

$$\mathcal{X}_c := \mathcal{V}_c \oplus \mathcal{W},$$
$$\mathcal{Y}_c := \mathcal{V}_c \oplus \mathcal{R}.$$

Let $v_c(t) := tP_\nu^m(\cos\theta)\cos m\varphi$. We restrict $f(\lambda, u) : \mathcal{X}_c \to \mathcal{Y}_c$. Then, by almost the same arguments as in the proof of Proposition 3.1, we can prove that there exists an implicit function $l(\mu, v)$ near $(\mu, v) = (0, P_\nu^m(\cos\theta)\cos m\varphi)$ such that

$$N_1(\mu, t) = 0 \qquad \text{for } |\mu| + |t - 1| < \delta \tag{3.22}$$

with some $\delta > 0$. On the other hand, by Lemma 3.3, $N_2(\mu, t) = 0$ also holds. Since $N_1(\mu, t) = N_2(\mu, t) = 0$ is equivalent to $K(\mu, v_c(t)) = 0$, Proposition 3.4 with $s_* = 0$ is proved.

Next we assume $t_* = 0$, and then $s_* = 1$. Let $v_s(t) := sP_\nu^m(\cos\theta)\sin m\varphi$, and we define

$$\mathcal{V}_s := \{tP_\nu^m(\cos\theta)\sin m\varphi \mid t \in \mathbf{R}\},$$
$$\mathcal{X}_s := \mathcal{V}_s \oplus \mathcal{W},$$
$$\mathcal{Y}_s := \mathcal{V}_s \oplus \mathcal{R}.$$

Now we replace $\mathcal{V}_c$, $\mathcal{X}_c$ and $\mathcal{Y}_c$ with $\mathcal{V}_s$, $\mathcal{X}_s$ and $\mathcal{Y}_s$, respectively. Then, from almost the same arguments above, Proposition 3.4 with $t_* = 0$ is proved. ∎

From Propositions 3.2–3.4, Theorem 1.2 follows.

Proof of Theorem 1.2. Recall (3.2) and Proposition 3.1. If $m = 0$, then, from arguments above and Proposition 3.2, we see that, for $\mu \le 0$ near $\mu = 0$,

$$u(\theta, \varphi; \mu + \lambda_{j,\epsilon}^0) = |\mu|^{\frac{1}{p-1}}\{t(\mu)v_* + l(\mu, t(\mu)v_*)\}$$

is a solution to (1.1). Here $t(0) = 1$, $l(0, v_*) = 0$ and

$$v_* = M_0 P^0_{\nu^0_{j,\epsilon}}(\cos\theta).$$

Similarly if $1 \le m \le j$, then, from arguments above and Propositions 3.3 and 3.4, we see that, for $\mu \le 0$ near $\mu = 0$,

$$u(\theta, \varphi; \mu + \lambda_{j,\epsilon}^m) = |\mu|^{\frac{1}{p-1}}\{v(t(\mu), s(\mu)) + l(\mu, v(t(\mu), s(\mu))\}$$

is a solution to (1.1). Here $t(0) = t_*$, $s(0) = s_*$, $l(0, t_*, s_*) = 0$, $t_*^2 + s_*^2 = 1$,

$$v(t, s) = tM_m P_\nu^m(\cos\theta)\cos m\varphi + sM_m P_\nu^m(\cos\theta)\sin m\varphi,$$

and $t(\mu) \equiv 0$ ($s(\mu) \equiv 0$) when $t_* = 0$ ($s_* = 0$). Theorem 1.2 is proved. ∎

4 Appendix In Appendix A we show Lemmas 2.2 and 2.3. Additionally, in Appendix B, we prove Lemma 3.1.

4.1 Proofs of Lemmas 2.2 and 2.3 In this appendix, we prove Lemmas 2.2 and 2.3. Arguments below follows results in [10].

Recall (1.8), and we define

$$F(a,b,c;x) := \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k k!} x^k, \tag{4.1}$$

where c is a non-positive integer. The function $F(a,b,c;x)$ is said to be the *Gauss hypergeometric function.* The radius of convergence of (4.1) is 1, and $F = F(a,b,c;x)$ satisfies the *Gauss hypergeometric equation*

$$x(1-x)\frac{d^2F}{dx^2} + \{c - (a+b+1)x\}\frac{dF}{dx} - abF = 0 \qquad (-1 < x < 1). \tag{4.2}$$

In Appendix we use some properties of $F(a,b,c;x)$, and those results follow from [10].

First, for $P_\nu^m(x)$ and $F(a,b,c;x)$, the following relation holds (p.319 in [10]):

$$\begin{aligned} P_\nu^m(x) = {} & \frac{(-1)^m \Gamma(\nu+m+1)(1-x^2)^{\frac{m}{2}}}{2^m \Gamma(\nu-m+1)m!} \\ & \times F\left(-\nu+m, \nu+m+1, m+1; \frac{1-x}{2}\right). \end{aligned} \tag{4.3}$$

By (4.3), we can prove Lemma 2.2.

Proof of Lemma 2.2. First we assume that ν is an integer. If $-\nu + m = -n$ (n be a non-negative integer), then

$$\begin{aligned} P_\nu^m(x) = {} & \frac{\Gamma(n+2\nu+1)(x^2-1)^{\frac{m}{2}}}{2^m \Gamma(-n+1)m!} \\ & \times F\left(-n, n+2\nu+1, m+1; \frac{1-x}{2}\right). \end{aligned}$$

If $n = 0$ ($\nu = m$), then, from (4.1), it follows that

$$\begin{aligned} F\left(0, 2\nu+1, m+1; \frac{1-x}{2}\right) &= \sum_{k=0}^{+\infty} \frac{(0)_k (2\nu+1)_k}{(m+1)_k k!} \left(\frac{1-x}{2}\right)^k \\ &\equiv 0 \end{aligned}$$

Hence $P_m^m(x) \equiv 0$ holds. On the other hand, if $n \geq 1$ ($\nu > m$), then, since $\Gamma(x)$ has singularity at $x = 0, -1, -2, ...,$ the identity $P_{m+n}^m(x) \equiv 0$ holds (see (4.3)).

Next we assume that ν is not an integer. Then, from (4.1), it holds that

$$F\left(-\nu+m, \nu+m+1, m+1, \frac{1-x}{2}\right) > 0 \qquad \text{for } -1 < x < 1.$$

Therefore, from (4.3), $P_\nu^m(x)$ does not have a zero for $-1 < x < 1$, and Lemma 2.2 is proved. ∎

Next we show Lemma 2.3 by using properties of $F(a,b,c;x)$. To investigate the behavior of $P_\nu^m(x)$ near $x=-1$, it suffices to consider that of $F(a,b,c;x)$ near $x=1$. For the purpose we use the following formula (see p.273 in [10])

$$\begin{aligned} F(a,b,a+b+1-c;1-x) &= \frac{\Gamma(a+b+1-c)\Gamma(1-c)}{\Gamma(a+1-c)\Gamma(b+1-c)}F(a,b,c;x) \\ &\quad + \frac{\Gamma(a+b+1-c)\Gamma(c-1)}{\Gamma(a)\Gamma(b)}x^{1-c}F(a+1-c,b+1-c,2-c;x). \end{aligned} \tag{4.4}$$

In addition we define

$$\begin{aligned} U(a,b,c;x) :=& \frac{\Gamma(1-c)}{\Gamma(a+1-c)\Gamma(b+1-c)}F(a,b,c;x) \\ &+ \frac{\Gamma(c-1)}{\Gamma(a)\Gamma(b)}x^{1-c}F(a+1-c,b+1-c,2-c;x). \end{aligned} \tag{4.5}$$

The function $U(a,b,c;x)$ is also a solution to (4.2), and moreover $F(a,b,c;x)$ and $U(a,b,c;x)$ are linearly independent (see p.274 in [10]). Thus, from (4.3)–(4.5), it follows that

$$\begin{aligned} P_\nu^m(-1+2x) = & \frac{(-1)^m\Gamma(\nu+m+1)\Gamma(m+1)x^{\frac{m}{2}}(1-x)^{\frac{m}{2}}}{\Gamma(\nu-m+1)m!} \\ & \times U\left(-\nu+m,\nu+m+1,m+1;x\right). \end{aligned} \tag{4.6}$$

Furthermore if c is an integer ($c=n$), then, for $U(a,b,n;x)$, the following formula holds (see p.275 in [10]):

$$\begin{aligned} U(a,b,n;x) = & \frac{(-1)^n}{\Gamma(a+1-n)\Gamma(b+1-n)(n-1)!}\Bigg[F(a,b,n;x)\log x \\ & + \sum_{k=0}^{+\infty}\frac{(a)_k(b)_k}{(n)_k k!}\{\psi(a+n)+\psi(b+n)-\psi(k+1)-\psi(n+k)\}x^k\Bigg] \\ & + \frac{(m-2)!}{\Gamma(a)\Gamma(b)}x^{1-n}\sum_{k=0}^{n-2}\frac{(a+1-n)_k(b+1-n)_k}{(2-n)_k k!}x^k. \end{aligned} \tag{4.7}$$

Proof of Lemma 2.3. Lemma 2.3 follows from (4.6) and (4.7). ■

4.2 Proof of Lemma 3.1 In this section, $H^k(\Omega|\mathbf{R}^2)$ denotes the usual Sobolev space on $\Omega\subset\mathbf{R}^2$. Before beginning to prove Lemma 3.1, we introduce the Sobolev spaces $H^k(B_{\pi-\epsilon})$ on $B_{\pi-\epsilon}\subset\mathbf{S}^2$ ($k=0,1,2$ and $H^0(B_{\pi-\epsilon})=L^2(B_{\pi-\epsilon})$). Now we introduce the stereographic projection from $\mathbf{S}^2$ to $\mathbf{R}^2$. Namely let

$$\left(\frac{2x_1}{1+|x|^2},\frac{2x_2}{1+|x|^2},\frac{1-|x|^2}{1+|x|^2}\right)\in\mathbf{S}^2,$$

where $(x_1,x_2)\in\Omega_{R_\epsilon}:=\{x\in\mathbf{R}^2 \mid |x|<R_\epsilon\}$ ($R_\epsilon:=\tan[(\pi-\epsilon)/2]$). Then norms of $H^k(B_{\pi-\epsilon})$ are expressed as, respectively,

$$\|u\|^2_{H^k(B_{\pi-\epsilon})} = \sum_{s=0}^{k}\int_{\Omega_{R_\epsilon}}|D^s_{\mathbf{S}^2}u|^2q^{2-2s}dx_1dx_2. \tag{4.8}$$

Here

$$|D^0_{\mathbf{S}^2} u|^2 = |u|^2, \tag{4.9}$$

$$|D^1_{\mathbf{S}^2} u|^2 = \left|\frac{\partial u}{\partial x_1}\right|^2 + \left|\frac{\partial u}{\partial x_2}\right|^2, \tag{4.10}$$

and

$$|D^2_{\mathbf{S}^2} u|^2 = \sum_{i,j=1}^{2} \left|\frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{n=1}^{2} \Gamma^n_{ij} \frac{\partial u}{\partial x_n}\right|^2, \tag{4.11}$$

where

$$q := \frac{2}{1+|x|^2},$$

$-\Gamma^1_{11} = -\Gamma^2_{12} = -\Gamma^1_{22} = 2x_1/(1+|x|^2)$ and $\Gamma^2_{11} = -\Gamma^1_{12} = -\Gamma^2_{22} = 2x_2/(1+|x|^2)$ (e.g., see Definitions 2.2 and 2.3 in [2] or Definition 2.1 in Hebey [12]). Moreover we define $H^k_0(B_{\pi-\epsilon})$ as the closure of $C^\infty_0(B_{\pi-\epsilon})$ in $H^k(B_{\pi-\epsilon})$.

From the definition above, it follows that

$$\Delta_{\mathbf{S}^2} u = q^{-2} \Delta u, \tag{4.12}$$

where Δ is the Laplace operator on $\mathbf{R}^2$.

The relation between the stereographic projection and the polar coordinates is as follows:

$$x_1 = \tan\left(\frac{\theta}{2}\right)\cos\varphi,$$
$$x_2 = \tan\left(\frac{\theta}{2}\right)\sin\varphi.$$

Thus, from

$$\begin{aligned} q^2 dx_1 dx_2 &= \frac{4\tan\left(\frac{\theta}{2}\right)}{\left(1+\tan^2\left(\frac{\theta}{2}\right)\right)^2} \cdot \frac{1}{2\cos^2\left(\frac{\theta}{2}\right)} d\theta d\varphi \\ &= 2\tan\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right) d\theta d\varphi \\ &= \sin\theta d\theta d\varphi, \end{aligned}$$

it follows that

$$\langle u, v\rangle = \int_{\Omega_{R_\epsilon}} uvq^2 dx_1 dx_2.$$

Therefore we prove Lemma 3.1 with the stereographic projection. Now recall definitions of the operator L and subspaces of $\mathcal{V}$, $\mathcal{W}$, $\mathcal{R}$ (see Section 3).

Proof of Lemma 3.1. We remark that, for any $g \in L^2(B_{\pi-\epsilon})$, there exists a unique solution $u \in H^1_0(B_{\pi-\epsilon}) \cap H^2(B_{\pi-\epsilon})$ to

$$\begin{cases} \Delta_{\mathbf{S}^2} u = g & \text{in } B_{\pi-\epsilon} \\ u = 0 & \text{on } \partial B_{\pi-\epsilon} \end{cases}$$

(e.g., see Theorem 4.8 in [2]). Hence there exists a unique inverse operator $\mathcal{G} : L^2(B_{\pi-\epsilon}) \to H^2(B_{\pi-\epsilon}) \subset L^2(B_{\pi-\epsilon})$ of $\Delta_{\mathbf{S}^2} : H^2(B_{\pi-\epsilon}) \to L^2(B_{\pi-\epsilon})$. Now we prove that $\mathcal{G}$ is compact and self-adjoint.

First we show that $\mathcal{G}$ is compact. From (4.11), the Minkowski inequality and $(a+b)^2 \leq 2(a^2+b^2)$, it holds that

$$
\begin{aligned}
|D^2_{\mathbf{S}^2}u|^2 &\leq \left[\left\{ \sum_{i,j=1}^{2} \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{i,j=1}^{2} \left| \sum_{n=1}^{2} \Gamma^n_{ij} \frac{\partial u}{\partial x_n} \right|^2 \right\}^{\frac{1}{2}} \right]^2 \\
&\leq 2 \sum_{i,j=1}^{2} \left\{ \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^2 + 2 \sum_{n=1}^{2} |\Gamma^n_{ij}| \left| \frac{\partial u}{\partial x_n} \right|^2 \right\}
\end{aligned} \tag{4.13}
$$

Hence, from (4.8)–(4.10) and (4.13), it follows that

$$\|u\|_{H^2(B_{\pi-\epsilon})} \leq K_1 \|u\|_{H^2(\Omega_{R_\epsilon}|\mathbf{R}^2)}, \tag{4.14}$$

where $K_1 > 0$ is some constant. In addition, from (4.8) and (4.9), it holds that

$$\|u\|_{L^2(\Omega_{R_\epsilon}|\mathbf{R}^2)} \leq K_2 \|u\|_{L^2(B_{\pi-\epsilon})}, \tag{4.15}$$

where some constant $K_2 > 0$. Moreover we apply the regularity theorem of elliptic equations for

$$
\begin{cases}
q^{-2}\Delta u = g & \text{in } \Omega_{R_\epsilon}, \\
u = 0 & \text{on } \partial\Omega_{R_\epsilon},
\end{cases}
$$

where $g \in L^2(\Omega_{R_\epsilon}|\mathbf{R}^2)$ (e.g., see Theorem 8.12 in Gilbarg and Trudinger [11]). Then we obtain

$$\|u\|_{H^2(\Omega_{R_\epsilon}|\mathbf{R}^2)} \leq K_3 \|g\|_{L^2(\Omega_{R_\epsilon}|\mathbf{R}^2)} = K_3 \|q^{-2}\Delta u\|_{L^2(\Omega_{R_\epsilon}|\mathbf{R}^2)} \tag{4.16}$$

with some constant $K_3 > 0$. Hence, from (4.8) and (4.14)–(4.16), we obtain

$$\|u\|_{H^2(B_{\pi-\epsilon})} \leq K_1 K_2 K_3 \|q^{-2}\Delta u\|_{L^2(\Omega_{R_\epsilon}|\mathbf{R}^2)} \leq 2K_1 K_2 K_3 \|\Delta_{\mathbf{S}^2} u\|_{L^2(B_{\pi-\epsilon})}.$$

for $u \in H^1_0(B_{\pi-\epsilon}) \cap H^2(B_{\pi-\epsilon})$ (see (4.12)). Thus the operator $\mathcal{G} : L^2(B_{\pi-\epsilon}) \to H^2(B_{\pi-\epsilon})$ is bounded. Furthermore the imbedding $H^2(B_{\pi-\epsilon}) \hookrightarrow L^2(B_{\pi-\epsilon})$ is compact by the *Rellich–Kondrachov theorem* (e.g., see Theorem 2.34 in [2]). Thus $\mathcal{G} : L^2(B_{\pi-\epsilon}) \to L^2(B_{\pi-\epsilon})$ is compact.

Second we show that $\mathcal{G}$ is self-adjoint. Let $\mathcal{G}^*$ be the adjoint operator of $\mathcal{G}$. Then, for any $u, v \in L^2(B_{\pi-\epsilon})$, it holds that

$$\langle u, \mathcal{G}^*(\Delta_{\mathbf{S}^2} v) \rangle = \langle \Delta_{\mathbf{S}^2}(\mathcal{G}u), v \rangle = \langle u, v \rangle.$$

Thus, by the uniqueness of the inverse operator $\mathcal{G}$, it holds that $\mathcal{G} = \mathcal{G}^*$ on $L^2(B_{\pi-\epsilon})$.

Thus we can apply the *Fredholm alternative theorem* (e.g., see Theorem 3 in p.284 of Yosida [18]) for $\mathcal{G}$. Namely, for any $a \in L^2(B_{\pi-\epsilon})$,

$$\lambda^{-1}u + \mathcal{G}(u) = a \qquad \text{for } u \in H^1_0(B_{\pi-\epsilon}) \cap H^2(B_{\pi-\epsilon})$$

has a solution u if and only if $\langle a, b \rangle = 0$ for any $b \in \mathrm{Ker}(\lambda^{-1} + \mathcal{G})$. Especially, for any $w \in \mathcal{R} \subset L^2(B_{\pi-\epsilon})$, $a := \lambda^{-1}\mathcal{G}(w) \in L^2(B_{\pi-\epsilon})$ holds. Hence it holds that

$$0 = \langle a, b \rangle = \langle \lambda^{-1}\mathcal{G}(w), b \rangle = \lambda^{-2}\langle w, \lambda\mathcal{G}(b) \rangle = -\lambda^{-2}\langle w, b \rangle.$$

Thus, for any $w \in \mathcal{R}$ and $b \in \mathrm{Ker}(\lambda^{-1} + \mathcal{G})$, we obtain $\langle w, b \rangle = 0$.

On the other hand, if $b \in \mathrm{Ker}(L)(= \mathrm{Ker}(\Delta_{\mathbf{S}^2} + \lambda))$, then, from $\lambda^{-1} b = \mathcal{G}(b)$, we obtain

$$\mathcal{G}(\Delta_{\mathbf{S}^2} b + \lambda b) = \lambda(\lambda^{-1} b + \mathcal{G}(b)) = 0.$$

Thus it holds that $b \in \mathrm{Ker}(\lambda^{-1} + \mathcal{G})$. Similarly if $b \in \mathrm{Ker}(\lambda^{-1} + \mathcal{G})$, then $b \in \mathrm{Ker}(L)$, and hence $\mathrm{Ker}(L) = \mathrm{Ker}(\lambda^{-1} + \mathcal{G})$. Therefore, since $\langle w, b \rangle = 0$ holds for any $w \in \mathcal{R}$ and $b \in \mathrm{Ker}(L)$. Lemma 3.1 is proved. ■

Acknowledgement. Professors Yoshitsugu Kabeya and Tatsuki Kawakami gave me significant advices in the research for the problem in this paper. I would like to express my gratitude to them.

References

[1] A. Ambrosetti and G. Prodi, "A primer of nonlinear analysis", Cambridge University Press (1993), Cambridge.

[2] T. Aubin, "Some nonlinear problems in Riemannian geometry", Springer Monographs in Mathematics, Springer-Verlag (1998), Berlin.

[3] F. E. Baginski, Ordering the zeros of Legendre functions $P_\nu^m(z_0)$ when considered as a function of ν, J. Math. Anal. Appl. **147** (1990), 296–308.

[4] F. E. Baginski, Comparison theorems for the ν-zeros of Legendre functions $P_\nu^m(z_0)$ when $-1 < z_0 < 1$, Proc. Amer. Math. Soc. **111** (1991), 395–402.

[5] C. Bandle, A. Brillard and M. Flucher, Green's function, harmonic transplantation, and best Sobolev constant in spaces of constant curvature, Trans. Amer. Math. Soc. **350** (1998), 1103–1128.

[6] C. Bandle and R. Benguria, The Brezis–Nirenberg problem on $\mathbf{S}^3$, J. Differential Equations **178** (2002), 264–279.

[7] C. Bandle, Y. Kabeya and H. Ninomiya, Imperfect bifurcations in nonlinear elliptic equations on spherical caps, Comm. Pure Appl. Anal. **9** (2010) 1189–1208.

[8] C. Bandle and L. A. Peletier, Best Sobolev constants and Emden equations for the critical exponent in $\mathbf{S}^3$, Math. Ann. **313** (1999), 83–93.

[9] H. Brezis and L. A. Peletier, Elliptic equations with critical exponent on spherical caps of $\mathbf{S}^3$. J. Anal. Math. 98 (2006), 279–316.

[10] R. Beals and R. Wong, "Special Functions", Cambridge University Press (2010), New York.

[11] D. Gilbarg and N. S. Trudinger, "Elliptic partial differential equations of second order. Reprint of the 1998 edition", Springer-Verlag (2001), Berlin.

[12] E. Hebey, "Sobolev spaces on Riemannian manifolds", Lecture Notes in Mathematics **1635**, Springer-Verlag (1996), Berlin.

[13] E. L. Ince, "Ordinary Differential Equations", Dover (1956), New York.

[14] A. Kosaka, Emden equation involving the critical Sobolev exponent with the third-kind boundary condition in $\mathbf{S}^3$, Kodai. Math. J. **35** (2012), 613–628.

[15] S. Moriguchi, K. Udagawa and S. Hitotsumatsu, "Mathematical Formulas III. Special Functions", Iwanami Shoten (1960), Tokyo (in Japanese).

[16] G. Sansone, "Orthogonal functions", Interscience Publishers (1959), New York.

[17] M. Struwe, "Variational methods. Applications to nonlinear partial differential equations and Hamiltonian systems" Fourth edition, Springer-Verlag (2008), Berlin.

[18] K. Yosida, "Functional Analysis" Reprint of the sixth 1980 edition, Springer-Verlag (1995), Berlin.

Communicated by *Yoshitsugu Kabeya*

Osaka Prefecture University
1-1 Gakuen-cho, Naka-ku, Sakai, Osaka
599-8531, Japan
e-mail: kosaka@ms.osakafu-u.ac.jp

RINGS WITH IDEAL CENTRES

Stephen M. Buckley and Desmond MacHale

Received October 23,2012

Abstract. We discuss the condition that the centre of a ring is an ideal. We also show that some classical commutativity results of Jacobson and Herstein have elementary proofs under the added assumption that the centre is an ideal.

1 Introduction Since in a group G, the centre $Z(G)$ is always a normal subgroup, one might expect that the centre $Z(R)$ of a ring would be a (two-sided) ideal in R. This is not true in general, though it can be true in some cases: at one extreme, it is trivially true if R is commutative, and at the other extreme it is also trivially true if R is "extremely non-commutative" in the sense that $Z(R) = \{0\}$. Thus rings where it fails to hold are in some sense "moderately non-commutative". We say that a ring R has an *ideal centre* if its centre is an ideal.

One of the main aims of this paper is to show that some classical ring commutativity results of Jacobson and Herstein, which we now state, have elementary proofs if we restrict to rings with ideal centres. Jacobson [12] proved that rings R satisfying an identity of the form $x^{n(x)} = x$ are commutative. Rings satisfying such an identity are rather special, but Herstein showed that commutativity is equivalent to the weaker condition $x^{n(x)} - x \in Z(R)$; see [9]. Herstein then generalized this result further:

Theorem A (Herstein [10])**.** *A ring R is commutative if and only if for each $x \in R$ there exists $f(X) \in X^2\mathbb{Z}[X]$ such that $f(x) - x \in Z(R)$.*

In the above result, $f(X) \in X^2\mathbb{Z}[X]$ means that $f(X)$ is a formal polynomial with integer coefficients (in the indeterminate X) which is formally divisible by X^2. We view f as a function on R in the natural way.

Subsequently, Herstein gave the following quite different generalization of Jacobson's theorem; here and later, $[x, y] = xy - yx$ is the *commutator* of x and y.

Theorem B (Herstein [11])**.** *A ring R is commutative if and only if for each $x, y \in R$ there exists an integer $n(x, y) > 1$ such $[x, y]^{n(x,y)} = [x, y]$.*

Known proofs of Theorems A and B require Jacobson's structure theory of rings, but we give elementary proofs of the following variants of these results.

Theorem 1. *A ring R with ideal centre is commutative if and only if for each $x \in R$ there exists $f(X) \in X^2\mathbb{Z}[X]$ such that $f(x) - x \in Z(R)$.*

Theorem 2. *The following conditions are equivalent for rings R with ideal centres.*

(a) R is commutative.

(b) For each $x, y \in R$ there exists an integer $n(x, y) > 1$ such $[x, y]^{n(x,y)} = [x, y]$.

2010 *Mathematics Subject Classification.* 16R50.
Key words and phrases. ring, commutativity condition.

(c) For each $x, y \in R$ there exists $f(X) \in X^2\mathbb{Z}[X]$ such that $f([x,y]) = [x,y]$.

Obviously Theorem 1 follows immediately from Theorem A, and the equivalence of (a) and (b) in Theorem 2 follows immediately from Theorem B, so the important feature of these results is that our proofs avoid structure theory. We do not however know of any proof that condition (c) in Theorem 2 implies commutativity for general rings.

We can view the above theorems as stating in particular that a polynomial $g(X) \in X\mathbb{Z}[X]$ whose X-coefficient equals ± 1 "forces" commutativity of R if $g(R) = 0$ (meaning that $g(x) = 0$ for all $x \in R$), or more generally $g(R) \subseteq Z(R)$ (meaning that $g(x) \in Z(R)$ for all $x \in R$), or $g([R,R]) = 0$ (meaning that $g([x,y]) = 0$ for all $x, y \in R$). The following result classifies the polynomials that force rings with ideal centre to be commutative in any of these three senses.

Theorem 3. *Let $g(X) := \sum_{i=1}^{n} a_i X^i \in \mathbb{Z}[X]$. Then*

(a) All rings R with ideal centre satisfying $g(R) = 0$ are commutative if and only if either:

- *(i) $a_1 = \pm 1$, or*
- *(ii) $a_1 = \pm 2$, a_2 is odd, and $a_2 + a_3 + \cdots + a_n$ is odd.*

(b) All rings R with ideal centre satisfying $g(R) \subseteq Z(R)$ are commutative if and only $a_1 = \pm 1$.

(c) All rings R with ideal centre satisfying $g([R,R]) = 0$ are commutative if and only $a_1 = \pm 1$.

Each part of the above result follow easily from the corresponding results without the ideal centre assumption; for these, see the main theorem in [13] for (a), [5, Proposition 4] for (b), and [6, Theorem 2] for (c). Thus the main value of these parts of Theorem 3 is again that the proof is elementary, although (c) also lead to the investigation of [6, Theorem 2].

We prove Theorems 1–3 in Section 3, but first in Section 2 we give some examples of noncommutative rings in which the centre is an ideal, and also answer the following pair of questions:

> *What is the order of the smallest finite ring/non-unital ring whose centre is not an ideal?*

2 Examples The concept of a ring with an ideal centre is mainly of interest for non-unital rings, since clearly a unital ring has an ideal centre if and only if it is commutative.

If we define a *good example* of a ring R with an ideal centre to be one where $Z(R)$ is both nonzero and proper, then all good examples are non-unital. The following pair of propositions give some families of good examples. In these propositions and later, $M(n, l, r, R_0)$ is the ring of $n \times n$ matrices $A = (a_{i,j})$ over a base ring R_0 such that $a_{i,j} = 0$ if $i > n - l$ or $j \le r$, and $U(n, m, R_0)$ is the ring of $n \times n$ matrices $A = (a_{i,j})$ over R_0 such that $a_{i,j} = 0$ if $j < i + m$. We use the more common notation $M(n, R_0)$ and $U(n, R_0)$ in place of $M(n, 0, 0, R_0)$ and $U(n, 0, R_0)$, respectively.

Proposition 4. *Suppose R_0 is a commutative unital ring with $1 \ne 0$, and that $n, l, r \in \mathbb{N}$ satisfy $n \ge 3$ and $l + r < n$. Then $R := M(n, l, r, R_0)$ is non-commutative, and $Z(R) = M(n, n - r, n - l, R_0)$ is a nontrivial proper ideal in R.*

Proof. Let $S := M(n, n-r, n-l, R_0)$. Because $n > n-r > l$ and $n > n-l > r$, it is clear that S is a proper and nontrivial subring of R. Let $\Sigma \in M(n, R_0)$ be the matrix corresponding to the shift map $(x_1, \ldots, x_n) \mapsto (x_2, \ldots, x_n, 0)$ in R_0^n, so that $\Sigma = (\sigma_{i,j})$, where

$$\sigma_{i,j} = \begin{cases} 1\,, & \text{if } 2 \le j = i+1 \le n\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Note that $R = \Sigma^l M(n, R_0) \Sigma^r$. Since the matrix Σ^n corresponds to the zero map and $S = \Sigma^{n-r} M(n, R_0) \Sigma^{n-l}$, it follows that $AB = BA = 0$ whenever $A \in R$ and $B \in S$. Thus S is an ideal and $S \subseteq Z(R)$.

Taking $A = (a_{i,j}) \in R \setminus S$, it remains to show that $A \notin Z(R)$. For $1 \le i, j \le n$, define $M_{i,j} \in M(n, R_0)$ to be the matrix whose (i, j)th entry is 1 and all other entries are 0. Suppose first that $a_{p,q} \ne 0$ for some $1 \le p, q \le n$ such that $p > r$. Now $B = M_{1,p} \in R$ and $\Sigma B = 0$, so $AB = 0$. On the other hand, the $(1, q)$th entry of BA is $a_{p,q}$ so $BA \ne 0$. Thus $A \notin Z(R)$.

The other way that A can fail to be in S is if $a_{p,q} \ne 0$ for some $1 \le p, q \le n$ such that $q \le n-l$. Now $B = M_{q,n} \in R$ and $B\Sigma = 0$, so $BA = 0$. On the other hand, the (p, n)th entry of AB is $a_{p,q}$, so $AB \ne 0$. Thus again $A \notin Z(R)$, and we are done. $\square$

The equation $Z(R) = M(n, n-r, n-l, R_0)$ proved above, and the fact that $Z(R)$ is an ideal, is true under weaker assumptions on n, l, r: it suffices that $n \ge 2$ and $0 < l + r < n$. However, note that if either $l = 0$ or $r = 0$, then $Z(R) = \{0\}$.

Our second proposition says various families of strictly upper triangular matrices also provide good examples.

Proposition 5. *Suppose R_0 is a commutative unital ring with $1 \ne 0$, and that $n, m \in \mathbb{N}$ satisfy $n \ge 3$ and $m < n/2$. Then $R := U(n, m, R_0)$ is non-commutative, and $Z(R) = M(n, n-m, n-m, R_0)$ is a nontrivial proper ideal in R.*

Proof. It is readily verified that $R \subset M(n, m, m, R_0)$ and that $S := M(n, n-m, n-m, R_0)$ is a proper subset of R. Most of the result now follows from Proposition 4, but we need to verify that if $A \in R \setminus S$ then $A \notin Z(R)$. The matrices B used to prove the corresponding result in Proposition 4 also lie in this ring R, so the same proof works. $\square$

In contrast with the above propositions, $M(n, R_0)$ and $U(n, R_0)$ are unital, so they have an ideal centre only if they are commutative, i.e. only if $n = 1$.

We now turn our attention to rings without ideal centres. Our first result is the following non-existence result.

Theorem 6. *Suppose R is a non-unital ring of order p^n, where p is prime and $n \in \mathbb{N}$, $n \le 3$. Then R has an ideal centre.*

Let us introduce some notation that will be useful in this proof and later: if x is an element of a ring R, then $\langle x; Z \rangle$ and $\langle\langle x; Z \rangle\rangle$ are the additive subgroup and the subring, respectively, generated in both cases by x and all $z \in Z(R)$; the ring R will be understood whenever we use such notation. Note that if $x \notin Z(R)$, then $Z(R) \subsetneq \langle x; Z \rangle \subseteq \langle\langle x; Z \rangle\rangle$, and that $\langle\langle x; Z \rangle\rangle$ is commutative.

Proof of Theorem 6. Suppose for the sake of contradiction that R is a non-unital ring of order p^n, $n \le 3$, and that $Z(R)$ is not an ideal. In particular $Z(R)$ is neither $\{0\}$ nor R so, as an additive subgroup of R, it must have order p^k for some $0 < k < n$. In particular $n > 1$. We can also quickly rule out $n = 2$, since then necessarily $k = 1$, and if $x \in R \setminus Z(R)$, then $\langle x; Z\rangle$ is commutative and strictly contains $Z(R)$, so it must have order p^2. Thus R is commutative, contradicting the assumption that $Z(R)$ is not an ideal.

Finally, suppose $n = 3$. We can rule out $k = 2$ in the same way as we ruled out $k = 1$ for $n = 2$, so we must have $k = 1$. Let z be a generator of $Z(R)$ as an additive group. Suppose first that $z^2 = 0$. Since $Z(R)$ is not an ideal, there exists some $u \in R \setminus Z(R)$ such that $zu \notin Z(R)$. Then $\langle u; Z\rangle$ has order at least p^2 and, since R cannot be commutative, $\langle\langle u; Z\rangle\rangle = \langle u; Z\rangle$ must have order p^2. Thus $zu = iz + ju$, where $i, j \in \mathbb{Z}_p$ and $j \neq 0$. But then

$$0 = z^2u = z(zu) = z(iz + ju) = ijz + j^2u$$

which gives a contradiction because $j^2 \neq 0$ and $u \notin Z(R)$.

Suppose instead that $z^2 \neq 0$, and so $z^2 = sz$ for some $s \in \mathbb{Z}_p$, $s \neq 0$. By distributivity, we see that $z^{t+1} = s^t z$ for all $t \in \mathbb{N}$, and so in particular $ez = z$, where $e = z^{p-1}$. Thus e is an identity on $Z(R)$ and in particular $e^2 = e$. Since $Z(R)$ is not an ideal, there exists $u \in R \setminus Z(R)$ such that $eu \notin Z(R)$. As before, $\langle u; Z\rangle$ has order at least p^2, and $\langle\langle u; Z\rangle\rangle$ cannot have order p^3 lest R be commutative, so $\langle\langle u; Z\rangle\rangle = \langle u; Z\rangle$ has order p^2. Thus $eu = ie + ju$ for some $i, j \in \mathbb{Z}_p$, $j \neq 0$. Now

$$ie + ju = eu = e^2u = e(ie + ju) = (i + ij)e + j^2u$$

so $i = i + ij$ and $j = j^2$. Since $j \neq 0$, we have $j = 1$, and hence $i = 0$. Thus $eu = u = ue$.

Now $x \mapsto ex$ is an additive homomorphism on R. Suppose it has trivial kernel. Then this map is a permutation on R, and so some iterate of it is the identity map. Of course the nth iterate of this map is just $x \mapsto e^n x$, and so $x \mapsto ex$, since $e^2 = e$. It follows that e is an identity for R, contradicting the assumption that R is non-unital. Thus there exists $v \in R \setminus \{0\}$ such that $ev = 0$. We deduce that

$$uv = (ue)v = u(ev) = 0 = (ev)u = v(eu) = vu\,,$$

so the subring S generated by $Z(R)$, u, and v is commutative. But $ex = x$ for $x = e, u$, so $ex = x$ for all $x \in \langle\langle u; Z\rangle\rangle$. Since $ev \neq v$, we see that $v \notin \langle\langle u; Z\rangle\rangle$. But $\langle\langle u; Z\rangle\rangle$ has order p^2, so S must have order p^3 and equal R. Thus R is commutative, contradicting our assumptions. □

It is easily proved that a finite ring can be decomposed as a direct sum of rings of prime power order. Indeed if $n = \prod_{p|n} p^{k_p}$ is the prime factorization of n, and $m_p = n/p^{k_p}$ for each $p \mid n$, then R is the direct sum of the ideals $R_p := m_p R$; see [8].

Clearly the centre of a direct sum is a direct sum of the centres, and a ring has an ideal centre if and only if each direct summand has an ideal centre, so to find a ring of minimal order where the centre is not an ideal it suffices to consider prime powers. It is now a straightforward matter to find the minimal order of (non-unital) rings in which the centre is not an ideal. In fact, we get the following result in which $N(2)$ is the ring of order 2 in which all products are 0.

Theorem 7.

(a) *Suppose R is a unital ring of order p^n, where p is prime and $n \le 3$. If R does not have an ideal centre, then $n = 3$ and R is isomorphic to $U(2, \mathbb{Z}_p)$.*

(b) *If R is a non-unital ring of order p^n, where p is prime and $n \le 3$, then R has an ideal centre. However, $R_{16} := U(2, \mathbb{Z}_2) \oplus N(2)$ is a non-unital ring of order 2^4 that fails to have an ideal centre.*

Consequently, the order of the smallest unital ring failing to have an ideal centre is 8*, and the order of the smallest non-unital ring failing to have an ideal centre is* 16*.*

Proof. By the comments above, the minimal orders must be prime powers. If a ring R is unital, then $1 \in Z(R)$, and so $Z(R)$ is an ideal if and only if R is commutative. Thus the unital ring without an ideal centre of minimal order is just the noncommutative unital ring of minimal order. This minimal order is known to be 8, and any such ring of order 8 must be isomorphic to the upper triangular matrix ring $U(2, \mathbb{Z}_2)$; see [7].

Since all prime powers of order less than 16 are of the form p^n for some $n \leq 3$, all such non-unital rings have ideal centres according to Theorem 6. A direct sum has an ideal centre if and only if all of its direct summands have ideal centres, so R_{16} fails to have an ideal centre by (a). The presence of the $N(2)$ summand prevents R_{16} from being unital. □

Remark 8. Unlike our unital ring of order 8, the non-unital example in the above proof is decomposable as a direct sum of smaller rings. We prove in [4] that all indecomposable non-unital rings of order p^k for $k \leq 4$ have an ideal centre (p being any prime), and that the smallest indecomposable non-unital ring without an ideal centre has order 32.

Remark 9. The previous remark suggests that perhaps finite indecomposable non-unital rings rarely fail to have an ideal centre. On the other hand, we will see in the next section that the assumption that a ring has an ideal centre is of great use for proving commutativity results. This suggests that the ideal centre assumption may be useful for formulating conjectures regarding conditions that may imply commutativity: if we can prove a commutativity result for rings with ideal centres, then it seems reasonable to search for a proof of the corresponding result without the ideal centre assumption. This technique has already lead to one success: we have been able to drop the ideal centre assumption in Theorem 2(c) leading to [6, Theorem 2]. Going beyond this, we would like to know if the three conditions in Theorem 2 are equivalent in the class of all rings.

We conclude this section by strengthening the previously mentioned fact that a noncommutative ring with an ideal centre cannot have a unity, but first we record a simple proposition.

Proposition 10. *A ring R has an ideal centre if and only if $cz = 0$ whenever c is a commutator and $z \in Z(R)$.*

Proof. If $x, y \in R$, and $z \in Z(R)$ then $(xy)z = x(zy)$ and $(yx)z = (yz)x = (zy)x$. Thus $[x, y]z = 0$ for all such x, y, z if and only if $zy \in Z(R)$ for all such y, z. □

Theorem 11. *Suppose a noncommutative ring R has an ideal centre. Then $R/Z(R)$ does not have a unity.*

Proof. Suppose for the sake of contradiction that $R/Z(R)$ has a unity $e + Z(R)$, where $e \in R$. It follows that for all $x \in R$, $x = exe + z_x$, for some $z_x \in Z(R)$. Also $e^2 = e + w$ for some $w \in Z(R)$. Thus

$$\begin{aligned} ex - xe &= e(exe + z_x) - (exe + z_x)e \\ &= e^2xe - exe^2 = wxe - exw = [x, e]w = 0\,, \end{aligned}$$

where the last equation follows from Proposition 10. Since x is arbitrary, it follows that $e \in Z(R)$, so the unity of $R/Z(R)$ is also the zero element. This contradicts the assumption that $Z(R)$ is not all of R. □

3 Elementary commutativity results Many results in the literature give elementary proofs of special cases of the results of Jacobson and Herstein mentioned in the introduction; in all cases, we use *elementary* to refer to proofs that do not appeal to Jacobson's structure theory of rings. The typical special case involves assuming that $n(x)$ or $n(x,y)$ takes on a particular constant value n. Let us review a few such results.

In the case of the identity $x^n = x$, elementary commutativity proofs were given by Morita [17] for all odd $n \le 25$ and all even $n \le 50$. MacHale [16] gave an elementary proof of commutativity for all even numbers n that can be written as sums or differences of two powers of 2, but are not themselves powers of 2. Also notable is the proof by Wamsley [19] of Jacobson's result which uses only a weak form of structure theory (specifically, the fact that a finite commutative ring can be written as a direct sum of fields).

For the condition $x^n - x \in Z(R)$, elementary proofs of commutativity are well known for $n = 2$ (see e.g. [1] and [14]), and such a proof for $n = 3$ can be found in [15, Theorem 2] and [18, Theorem 1]. Elementary proofs for odd $n < 10$, and for infinitely many even values of n, can be found in [3]. In the case of the condition $[x,y]^n = [x,y]$, an elementary proof for $n = 3$ (and a fortiori for $n = 2$) is given in [3, Theorem 17].

Our results are rather different from the above elementary theorems since we do not restrict $n(x)$ or $n(x,y)$—indeed we consider more general polynomial conditions—but instead we add the assumption that $Z(R)$ is an ideal. The one result of a similar type in the literature of which we are aware is Theorem 6 of [16] which implies in particular that if R is a ring, $Z(R)$ is an ideal, and there exists an even number $n > 1$ such $x^n - x \in Z(R)$ for all $x \in R$, then R is commutative. This implication of course follows from a special case of Theorem 1.

We begin with a well-known lemma, and include a proof for completeness.

Lemma 12. *Let R be a ring in which $xy = 0$ implies $yx = 0$. If e is an idempotent in R, then $e \in Z(R)$.*

Proof. For all $r \in R$, $e(r - er) = er - eer = er - er = 0$, so $(r - er)e = 0$, and so $re = ere$. By considering $(r - re)e$, we similarly deduce that $er = ere$. Thus $er = re$, and so $e \in Z(R)$. □

We now prove one of our main results.

Proof of Theorem 2. Trivially, (a) implies (b), and (b) implies (c), so we need only prove that (c) implies (a). Suppose $xy = 0$ for some $x, y \in R$. Then $[y,x] = yx$ and so there exists $f(X) \in X^2\mathbb{Z}[X]$ such that $yx = f(yx)$. Each term of the polynomial expression $f(yx)$ is an integer multiple of $(yx)^n$ for some $n > 1$. But $(yx)^n$ can be written in the form $y(xy)x'$ (where $x' = x$ or $x' = x(yx)^{n-2}$, depending on whether $n = 2$ or $n > 2$), and so $yx = y(xy)x' = y(0)x' = 0$. By Lemma 12, idempotents are central.

Let us now fix an arbitrary pair of elements $u, v \in R$, and write $w = [u,v]$. Let $G(X) \in X^2\mathbb{Z}[X]$ be such that $w = G(w)$. Factorizing $G(X) = Xg(X)$, we have $wg(w) = w$, and so $g(w)$ is an identity in the subring generated by w. In particular, $g(w)$ is an idempotent, and so central. Since $w = wg(w)$ and $Z(R)$ is an ideal, it follows that $w \in Z(R)$. Thus all commutators are central.

In an arbitrary ring, the identity $y[x,y] = [yx,y]$ follows immediately from the definition of commutators. Using this identity and the centrality of commutators, we see that for all $a, b \in R$,

$$ab[a,b] = a[ba,b] = [ba,b]a = b[a,b]a = ba[a,b]\,,$$

and so $[a,b]^2 = 0$. Choosing $h(X) \in X^2\mathbb{Z}[X]$ such that $h([a,b]) = [a,b]$, the identity $[a,b]^2 = 0$ readily implies that $h([a,b]) = 0$, and so $[a,b] = 0$. Thus R is commutative. □

For convenience, we now make the following definition: a ring R is a *H-ring* if for every $x \in R$ there exists $f(X) \in X^2\mathbb{Z}[X]$ for which $f(x) - x \in Z(R)$. ("H" is in honor of Herstein who proved that H-rings are commutative; see Theorem A.)

The key step in proving Theorem 1 is to prove the following special case.

Lemma 13. *Suppose that R is a H-ring with an ideal centre. Then $Z(R)$ contains all $e \in R$ such that $e^2 - e \in Z(R)$.*

Proof. Let $y \in R$ be arbitrary, and define $d := eye - ye$ and $z := e^2 - e$. Suppose that $z \in Z(R)$. Since R has an ideal centre, $ed = zye \in Z(R)$ and consequently $(de)^2 = d(ed)e \in Z(R)$. It then follows that $(de)^k \in Z(R)$ for all $k \in \mathbb{N}$, $k > 1$, and so $f(de) \in Z(R)$ whenever $f(X) \in X^2\mathbb{Z}[X]$. Using the H-ring property, we see that $de \in Z(R)$. Thus $eye - ye = de - (ey - y)z \in Z(R)$. By symmetry[1], $eye - ey \in Z(R)$, and so $ey - ye \in Z(R)$.

Now $(ey)(ey - ye) = (ey - ye)(ey)$, so

$$ey^2e = ye^2y = yey + zy^2\,. \tag{14}$$

Next, we show that

$$(ye)^2 = (ey^2e - zy^2)e = ey^2e\,. \tag{15}$$

The first equation in (15) follows immediately from (14). Because $z, zy^2 \in Z(R)$, we deduce that $zy^2e = ezy^2 = ey^2z$, and the second equation in (15) now follows immediately from the equation $z = e^2 - e$. We deduce from (15) and symmetry that $(ey)^2 = ey^2e$, and so

$$(ye)^2 = (ey)^2\,. \tag{16}$$

Now $e(e + x) = e + z + ex$ and $(e + x)e = e + z + xe$, so

$$(e(e + x))^2 = e^2 + z^2 + (ex)^2 + 2ez + 2zex + e^2x + exe$$

and

$$((e + x)e)^2 = e^2 + z^2 + (xe)^2 + 2ez + 2zxe + xe^2 + exe$$

But the expressions on the left of the last two displays are equal by (16) with $y = e + x$, and $zxe = zex$ since $z \in Z(R)$ and $Z(R)$ is an ideal, so we conclude that $e^2x = xe^2$. Since $e^2 - e \in Z(R)$, we finally get $ex = xe$ for all $x \in R$, as required. □

Proof of Theorem 1. It suffices to prove that a H-ring with ideal centre R is necessarily commutative. Fixing an arbitrary $x \in R$, let $F(X) \in X^2\mathbb{Z}[X]$ be such that $F(x) - x \in Z(R)$. We factorize $F(X) = Xf(X)$, and write $e := f(x)$. By assumption, $xe - x \in Z(R)$. Since $Z(R)$ is an ideal, we deduce inductively that $x^ne - x^n \in Z(R)$ for all $n \in \mathbb{N}$, and so $g(x)e - g(x) \in Z(R)$ for all $g(X) \in X\mathbb{Z}[X]$. In particular, $e^2 - e \in Z(R)$. Lemma 13 now implies that $e \in Z(R)$, and so $x = xe - (xe - x) \in Z(R)$. Since x is arbitrary, we are done. □

Proof of Theorem 3. Sufficiency of the coefficient conditions in (a) and (b) follows trivially from the proof of the corresponding result for general rings (i.e. the main result in [13] for (a), and [5, Proposition 4] for (b)). The only non-elementary parts of the earlier proofs are the use in both cases of results of Herstein mentioned in the introduction: in fact, an

[1]Note that the hypotheses are also satisfied by the opposite ring R^{op} (this is the ring with the same addition as R and multiplication $*$ given in terms of R-multiplication by $x * y = yx$), so appealing to symmetry to reverse the order of the elements here and later is justified.

appeal to Theorem A suffices in both cases. For rings with ideal centre, we therefore get an elementary proof of these implications simply by appealing to Theorem 1 instead of Theorem A. As for the converse implications in (a) and (b), these are established in [13] and [5] by giving counterexamples for each of the various situations in which the coefficient conditions fail. Since it is readily verified that all of these counterexamples are rings with ideal centres, these same counterexamples establish the converse implications in the current result.

Finally, we tackle (c). Sufficiency of the coefficient condition follows from Theorem 2. As for necessity, suppose f is a polynomial such that its coefficient a_1 is not ± 1. Thus a_1 has a prime factor p. Consider the ring R_p of 3×3 matrices over $\mathbb{Z}_p$ of the form

$$\begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

As is readily verified, $Z(R_p)$ consists of all matrices of the above form with $a = c = 0$ and $R_p \cdot Z(R_p) = \{0\}$, so $Z(R_p)$ is an ideal. Moreover, the set of commutators C_p equals $Z(R_p)$, so it follows from the equation $R_p \cdot C_p = \{0\}$ that $f(x) = 0$ for all $x \in C_p$. However, R_p is not commutative, so we are done. □

References

[1] R. Ayoub, C. Ayoub, *On the commutativity of rings*, Amer. Math. Monthly **71** (1964), 267–271.

[2] S.M. Buckley and D. MacHale, *Centrifiers and ring commutativity*, J. Math. Sci. Adv. Appl. **15** (2012), 139–161.

[3] S.M. Buckley and D. MacHale, *Variations on a theme: rings satisfying $x^3 = x$ are commutative*, Amer. Math. Monthly **120** (2013), 430–440.

[4] S.M. Buckley and D. MacHale, *Small rings in which the centre is not an ideal*, Bull. Irish Math. Soc. **70** (2012), 41–49.

[5] S.M. Buckley and D. MacHale, *Z-polynomials and ring commutativity*, Math. Proc. R. Ir. Acad. **112** (2012), 51–57.

[6] S.M. Buckley and D. MacHale, *A characterization of commutator-forcing polynomials*, Math. Proc. R. Ir. Acad., to appear. (Available at http://archive.maths.nuim.ie/staff/sbuckley/Papers/bm$_c$*omm.pdf*)

[7] K.E. Eldridge, *Orders for finite noncommutative rings with unity*, Amer. Math. Monthly **75** (1968), 512–514.

[8] L. Fuchs, *Ringe und ihre additive Gruppe*, Publ. Math. Debrecen **4** (1956), 488-508.

[9] I.N. Herstein, *A generalization of a theorem of Jacobson III*, Amer. J. Math. **75** (1953), 105–111.

[10] I.N. Herstein, *The structure of a certain class of rings*, Amer. J. Math. **75** (1953), 864–871.

[11] I.N. Herstein, *A condition for the commutativity of rings*, Canad. J. Math. **9** (1957), 583–586.

[12] N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. Math. **46** (1945), 695–707.

[13] T.J. Laffey and D. MacHale, *Polynomials that force a ring to be commutative*, Proc. Royal Irish Acad. **92A** (1992), 277–280.

[14] D. MacHale, *Rings that are nearly Boolean*, Proc. Roy. Irish Acad. Sect. A **80** (1980), 41–46.

[15] D. MacHale, *Rings that are nearly Boolean II*, Proc. Roy. Irish Acad. Sect. A **83** (1983), 165–167.

[16] D. MacHale, *Rings that are nearly Boolean III*, Proc. Roy. Irish Acad. Sect. A **86** (1986), 165–167.

[17] Y. Morita, *Elementary proofs of the commutativity of rings satisfying $x^n = x$*, Memoirs Def. Acad. Jap. **XVIII** (1978), 1–23.

[18] S. Vernon, *On commutativity in certain classes of rings*, Proc. Roy. Irish Acad. Sect. A **83** (1983), 197–200.

[19] J.W. Wamsley, *On a condition for commutativity of rings*, Proc. London Math. Soc. (2) **4** (1971), 331–332.

Communicated by *Klaus Denecke*

S.M. Buckley:
Department of Mathematics and Statistics, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland.
Email: mailto:stephen.buckleymaths.nuim.iestephen.buckley@maths.nuim.ie

D. MacHale:
School of Mathematical Sciences, University College Cork, Cork, Ireland.
E-mail: mailto:d.machaleucc.ied.machale@ucc.ie

Scientiae Mathematicae Japonicae, **78**, No. 1 (2015), 53–59

HYPERSPACES AND COMPLETE INVARIANCE PROPERTY

Saurabh Chandra Maury

Received June 15, 2013

Abstract. In this paper, the uniform flow over the hyperspaces 2^X of nonempty compact subsets of a noncompact metric space X with uniform flow, and $F_n(X)$ of nonempty subsets of a compact metric space X with uniform flow containing atmost n points is introduced and used to show that the hyperspace 2^X has the CIP and the hyperspace $F_n(X)$ has the CIPH.[1]

1. Introduction

A topological space X is said to possess the complete invariance property (CIP) if each of its nonempty closed subsets is the fixed point set, for some self continuous map f on X [14]. In case, f can be found to be a homeomorphism, we say that the space enjoys the complete invariance property with respect to homeomorphism ($CIPH$) [6].

A survey of results concerning the CIP may be found in [11] and a number of nonmetric results may be found in [9]. Some spaces known to have the $CIPH$ are even-dimensional Euclidean balls [13], compact surfaces and positive-dimensional spheres [12], the Hilbert cube and metrizable product spaces which have the real line or an odd-dimensional sphere as a factor [6]. In [9], it is shown that an uncountable self product of circles, real lines or two point spaces has the CIP and that connected subgroups of the plane and compact groups need not have the CIP.

In this paper, it is investigated as to when the hyperspace 2^X of nonempty compact subsets of a metric space X enjoys the notion of complete invariance property (CIP). It is also shown that the hyperspace $F_n(X)$ of nonempty subsets of a compact metric space X with uniform flow containing atmost n points has the complete invariance property with respect to homeomorphism ($CIPH$).

2. Pre-requisites

A. Hyperspaces

For a metric space $(X,\ d)$, the *hyperspace* 2^X of nonempty compact subsets and the *hyperspace* $C(X)$ of nonempty compact connected subsets are topologized by the *Hausdorff*

[1] *2010 AMS subject Classification*: 55M20, 54H25, 54B20

Keywords and Phrases: Hyperspaces, Hausdorff metric, CIP, CIPH.

metric, defined by

$$d_H(A,\ B) = max\{sup_{a\in A}d(a,\ B),\ sup_{b\in\ B}d(b,\ A)\}.$$

If $(X,\ d)$ is a compact metric space, then for each $n \in \mathbb{N}$ the *hyperspace* $F_n(X) = \{A \in 2^X : cardA \leq n\}$ is called the *n-symmetric* product of X. 1-*symmetric* product of X is *the hyperspace* $F_1(X)$ of singletons of X and $F_1(X) \approx X$. For a compact metric space $(X,\ d)$ *the hyperspace* 2^X is compact with respect to the *Hausdorff metric* and $F_n(X) \subset 2^X$ is closed for every n.

Definition 2.1.[10] A *continuum* is a nonempty, compact, connected metric space. A locally connected continuum is called *Peano continuum.* A continuum that contains more than one point is called *nondegenerate.*

Definition 2.2.[10] Any space homeomorphic to the closed interval $[0, 1]$ is called an *arc.* A *free arc* in the continuum X is an arc α with end points a and b such that $\alpha - \{a,\ b\}$ is open in X.

Result 2.3. 2^X *is compact metric space if and only if* X *is a compact metric space.* $C(X)$ *is closed in* 2^X*, hence also compact.*

Result 2.4.[5] (Curtis - Shori theorem, a famous result concerning the topology of hyperspaces) *The hyperspace* 2^X *is homeomorphic to the Hilbert cube* $(\mathcal{Q})$ *if and only if* X *is a non-degenerate Peano continuum and* $C(X)$ *is homeomorphic to* $\mathcal{Q}$ *if and only if* X *is a non-degenerate Peano continuum with no free arcs.*

B. CIP and CIPH

A topological space X is said to possess *the complete invariance property* (CIP) [14] if every nonempty closed subset of X is the fixed point set of some continuous self map f on X. In case, f can be chosen to be a homeomorphism, the space is said to possess *the complete invariance property with respect to homeomorphism* (CIPH) [6].

Definition 2.5.[3] A continuous function $\varphi : X \times \mathbb{R} \to X$ on a metric space $(X,\ d)$ is called *uniform flow* if it satisfies the following conditions :

i. $\varphi(x,\ 0) = x$, for all $x \in X$.

ii. $\varphi(\varphi(x,\ s),\ t) = \varphi(x,\ s + t)$, for all $s,\ t \in \mathbb{R}$ and $x \in X$.

iii. $d(x,\ \varphi(x,\ t)) \leq C|t|$, for some positive C and for all $x \in X,\ t \in \mathbb{R}$.

iv. There is a real number $p \geq 0$ such that for all $t \in \mathbb{R}$ and $x \in X$, $\varphi(x,\ t) = x$ iff $t \in p\mathbb{Z}$.

The map $\varphi_t : X \to X$ *defined by* $\varphi_t(x) = \varphi(x,\ t)$ *is a homeomorphism.*

Definition 2.6.[8](1) A space X has *property* Q if for every nonempty closed subset K of X there is a point $p \in K$, a retract R of X containing K and a deformation $H : X \times I \to R$ such that $H(x,t) \neq x$ if $x \neq p$ and $t > 0$.

(2) If in (1) we omit p and stipulate that $H(x,t) \neq x$ if $x \notin K$ and $t > 0$, then we say that X has *property* $Q(weak)$.

(3) A space X has *property* W if for every point $p \in X$, there is a deformation H : $X \times I \to X$ such that $H(x,t) \neq x$ if $x \neq p$ and $t > 0$.

(4) If in (3) $H(x,t) \neq x$, whenever $t > 0$, we say that X has *property* $W(strong)$.

Remark 2.7. $W(strong) \Rightarrow W \Rightarrow Q \Rightarrow Q(weak)$.

Result 2.8.[8] *A metric space* (X, d) *with property* W *has the CIP.*

Result 2.9.[6] *The Hilbert cube* $\mathcal{Q}$ *has the CIPH.*

Result 2.10.[3] L*et* (X, d) *be a compact metric space with a uniform flow* φ*. Then every nonempty closed subset of* X *is the fixed point set of an orbit-preserving autohomeomorphism of* X*. In particular,* X *has the CIPH.*

3. CIP and CIPH over hyperspaces

In this section, the notion of uniform flow over the hyperspaces 2^X and $F_n(X)$ is introduced and using the property of uniform flow we study the CIP and the CIPH over hyperspaces with respect to the Hausdorff metric.

Theorem 3.1. *Let* $(X,\ d)$ *be a noncompact metric space with uniform flow* φ *satisfying* $\varphi_t(A) = A$ *if and only if* $t = 0$*, for all nonempty compact subsets* A *of* X*. Then the map* $\Phi : 2^X \times \mathbb{R} \to 2^X$ *on the hyperspace of all nonempty compact subsets of* X*, defined by* $\Phi(A,\ t) = \{\varphi(a,\ t) : a \in A\}$ *is a uniform flow.*

Proof. Let constants of uniform flow φ be p and C. Now we show that the map $\Phi : 2^X \times \mathbb{R} \to 2^X$ defined by $\Phi(A,\ t) = \{\varphi(a,\ t) : a \in A\}$ is a uniform flow.

i. $\Phi(A,\ 0) = \{\varphi(a,\ 0) : a \in A\} = A$ for all $A \in 2^X$.

ii. $\Phi(A,\ s + t) = \Phi(\Phi(A,\ s),\ t)$ for all $A \in 2^X$ and $s, t \in \mathbb{R}$.

iii. Consider

$$d_H(A,\ \Phi(A,\ t)) = \max\{\sup_{a\in A} d(a,\ \Phi(A,\ t)),\ \sup_{\varphi(a,\ t)\in\Phi(A,\ t)} d(A,\ \varphi(a,\ t))\}.$$

By noting that

$$d(a,\ \Phi(A,\ t)) = \inf_{\varphi(a',\ t)\in\Phi(A,\ t)} d(a,\ \varphi(a',\ t)) \leq C|t|,$$

we have

$$\sup_{a\in A} d(a,\ \Phi(A,\ t)) \leq C|t|,$$

and also

$$d(A,\ \varphi(a,\ t)) = \inf_{a' \in A} d(a',\ \varphi(a,\ t)) \leq C|t|,$$

provides

$$\sup_{\varphi(a,\ t) \in \Phi(A,\ t)} d(A,\ \varphi(a,\ t)) \leq C|t|.$$

Thus, we have

$$d_H(A,\ \Phi(A,\ t)) \leq C|t|,$$

for all $A \in 2^X$.

iv. From the condition $\varphi_t(A) = A$ if and only if $t = 0$ we have $\Phi(A,\ t) = \{\varphi(a,\ t) : a \in A\} = \varphi_t(A) = A$, if and only if $t = 0$.

Remark. If X is compact metric space, then $X \in 2^X$. In this case $\varphi_t(X) = X$ for all $t \in \mathbb{R}$.

Example 3.2. Let $(S^1,\ d_1)$ be a metric space where S^1 is the unit circle and d_1 is arc length metric. If $(X,\ d_2)$ is any metric space, then the product $X \times S^1$ is a metric space with the metric D defined by

$$D((x_1,\ y_1),\ (x_2,\ y_2)) = max\{d_2(x_1,\ x_2),\ d_1(y_1,\ y_2)\}.$$

The map $\varphi : (X \times S^1) \times \mathbb{R} \to X \times S^1$ defined by $\varphi((x,\ e^{i\alpha}),\ t) = (x,\ e^{i(\alpha+t)}),\ \alpha \in [0,\ 2\pi), t \in \mathbb{R}$ is a uniform flow with $p = 2\pi$ and $C = 1$.

If we take $A = \{x_0\} \times S^1$, then A is a nonempty compact subset of $X \times S^1$ such that $\varphi_t(A) = A$ for all $t \in \mathbb{R}$.

Example 3.3. The map $\varphi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2$ defined by $\varphi((x,\ y),\ t) = (x + t,\ y),\ t \in \mathbb{R}$ is a uniform flow with $p = 0$ and $C = 1$.

If A is a nonempty compact subset of $\mathbb{R}^2$, then $\varphi_t(A) = A$ if and only if $t = 0$.

Example 3.4. Let $(S^1,\ d)$ be a metric space where S^1 is the unit circle and d is the arc length metric. If $\mathbb{T}^2 = S^1 \times S^1$, then $\mathbb{T}^2$ is a metric space with the metric D defined by

$$D((e^{2\pi ix_1},\ e^{2\pi iy_1}),\ (e^{2\pi ix_2},\ e^{2\pi iy_2})) = \sqrt{d^2(e^{2\pi ix_1},\ e^{2\pi ix_2}) + d^2(e^{2\pi iy_1},\ e^{2\pi iy_2})},$$

where $x_1, x_2, y_1, y_2 \in [0, 1]$.

Define a map φ on $\mathbb{T}^2$ by

$$\varphi((e^{2\pi ix},\ e^{2\pi iy}),\ t) = (e^{2\pi i(x+t)},\ e^{2\pi i(y+\sqrt{2}t)}).$$

Then

$D((e^{2\pi ix},\ e^{2\pi iy}),\ \varphi((e^{2\pi ix},\ e^{2\pi iy}),\ t)) = D((e^{2\pi ix},\ e^{2\pi iy}),\ (e^{2\pi i(x+t)},\ e^{2\pi i(y+\sqrt{2}t)})) \leq C\,|t|$, where $C = 2\sqrt{3}\pi$ and $\varphi((e^{2\pi ix},\ e^{2\pi iy}),\ t) = (e^{2\pi ix},\ e^{2\pi iy})$ if and only if $t = 0$, shows that φ is a uniform flow.

Again the map $\Phi : (\mathbb{T}^2 \times \mathbb{R}) \times \mathbb{R} \longrightarrow \mathbb{T}^2 \times \mathbb{R}$ defined by

$$\Phi(((e^{2\pi ix},\ e^{2\pi iy}),\ a),\ t) = ((e^{2\pi i(x+t)},\ e^{2\pi i(y+\sqrt{2}t)}),\ a),$$

is a uniform flow with respect to the metric $\mathcal{D}$ given by

$$\mathcal{D}(((e^{2\pi ix_1},\ e^{2\pi iy_1}),\ a_1),\ ((e^{2\pi ix_2},\ e^{2\pi iy_2}),\ a_2)) = \max\{D((e^{2\pi ix_1},\ e^{2\pi iy_1}),\ (e^{2\pi ix_2},\ e^{2\pi iy_2})),\ |a_1 - a_2|\}.$$

Consider $A = \{e^{2\pi ix_0}\} \times S^1 \times \{a\}$, the nonempty compact set in $\mathbb{T}^2 \times \mathbb{R}$. We have $\{\Phi((((e^{2\pi ix_0},\ e^{2\pi iy}),\ a),\ 1)) : y \in [0,\ 1]\} = \{((e^{2\pi i(x_0+1)},\ e^{2\pi i(y+\sqrt{2})}),\ a) : y \in [0,\ 1]\} = \{((e^{2\pi ix_0},\ e^{2\pi i(y+\sqrt{2})}),\ a) : y \in [0,\ 1]\} = A$.
Thus, in this case, $\Phi_t(A) = A$ if and only if $t = 0$, is not true.

Theorem 3.5. *Let $(X,\ d)$ be a compact metric space with uniform flow φ satisfying $\varphi_t(A) = A$ if and only if $t \in p\mathbb{Z}$, for all subsets A of X containing atmost n points, then the map $\Phi : F_n(X) \times \mathbb{R} \to F_n(X)$ on the hyperspace of subsets of X containing atmost n points, defined by $\Phi(A,\ t) = \{\varphi(a,\ t) : a \in A\}$ is a uniform flow.*

Proof. Since (X, d) is a compact metric space and $A \in F_n(X)$ is closed in X, hence $A \in F_n(X)$ is a compact set in X. Since continuous image of a compact metric space is compact, we have the result from the proof of Theorem 3.1.

Theorem 3.6.[3] *Any metric space with uniform flow has property $W(strong)$.*

Proof. Let (X, d) be a metric space with uniform flow ϕ.
If $p = 0$, then define $H : X \times [0, 1] \to X$ by

$$H(x, t) = \phi(x, t), \qquad x \in X,\ t \in [0, 1].$$

Then H is continuous and $H(x, t) = x$ if and only if $t = 0$.

If $p > 0$, then $H : X \times [0, 1] \to X$ given by

$$H(x, t) = \phi(x, t/2p), \qquad x \in X,\ t \in [0, 1]$$

is a homotopy and $H(x, t) = x$ if and only if $t = 0$.

Theorem 3.7. *Any metric space with uniform flow has the CIP.*

Proof. It follows from the fact that a metric space having property $W(strong)$ has the CIP.

Theorem 3.8. *If $(X,\ d)$ is a noncompact metric space with uniform flow φ satisfying $\varphi_t(A) = A$ if and only if $t = 0$, for all nonempty compact subsets A of X, then the hyperspace 2^X of all nonempty compact subsets of X has the CIP.*

Proof. Since 2^X is a metric space with respect to the Hausdorff metric. The proof is immediate by Theorem 3.1 and Theorem 3.7.

Theorem 3.9. *If (X, d) is a compact metric space with uniform flow φ satisfying $\varphi_t(A) = A$ if and only if $t \in p\mathbb{Z}$, for all subsets A of X containing atmost n points, then the hyperspace $F_n(X)$ of subsets of X containing atmost n points has the CIPH.*

Proof. Since for a compact metric space X, the hyperspace $F_n(X)$ is compact. From the Theorem 3.5 we get that $F_n(X)$ has uniform flow. Thus $F_n(X)$ is a compact metric space with uniform flow and hence the result follows from Theorem 2.10.

Theorem 3.10 *For a nondegenerate Peano continuum X, the hyperspace 2^X of nonempty compact subsets of X and the hyperspace $C(X)$ of nonempty compact connected subsets of X with no free arcs have the CIPH.*

Proof. On account of Result 2.3, 2^X and $C(X)$ are homeomorphic to the Hilbert cube $\mathcal{Q}$. Thus from 2.9 we get the result.

Example 3.11 The closed unit interval I, the Hilbert cube $\mathcal{Q}$ and closed unit sphere S^n are nondegenerate Peano continuum. Thus from Theorem 3.10 the hyperspaecs 2^I, $2^{\mathcal{Q}}$ and 2^{S^n} have the CIPH.

Example 3.12 For the closed unit interval I and the unit circle S^1, hyperspaecs $C(I)$ and $C(S^1)$ are homeomorphic to the product I^2. Hence $C(I)$ and $C(S^1)$ have the CIPH.

ACKNOWLEDGEMENT: The author gratefully acknowledges the financial support provided by the CSIR, New Delhi, India.

References

[**1**] T. Banakh, R. Voytsitskyy, Characterizing metric spaces whose hyperspaces are absolute neighborhood retracts, Topology Appl. 154, 10 (2007), 2009-2025.

[**2**] J.J. Charatonik, Recent research in hyperspace theory, Extracta. Math. 18, 2(2003), 235-262.

[**3**] A. Chigogidze, K.H. Hofmann, J.R. Martin, Compact groups and fixed point sets, Trans. Amer. Math. Soc. 349 (1997), 4537-4554.

[**4**] D.W. Curtis, Hyperspaces of noncompact metric spaces, Compositio Math. 40 (1980), 126-130

[**5**] D.W. Curtis R.M. Schori, Hyperspaces of Peano continua are Hilbert cubes, Fund. Math. 101 (1978), 19-38.

[**6**] J.R. Martin, Fixed point sets of homeomorphisms of metric products, Proc. Amer. Math. Soc. 103 (1988), 1293-1298.

[**7**] J.R. Martin and S.B. Nadler Jr., Example and questions in the theory of fixed point sets, Canad. J. Math. 31(1997), 1017-1032.

[**8**] J.R. Martin, L.G. Oversteegen, E.D. Tymchatyn, Fixed point sets of products and cones, Pacific J. Math. 101 (1982), 133-139

[**9**] J.R. Martin and W.A.R. Weiss, Fixed point sets of metric and nonmetric spaces, Trans. Amer.Math. Soc. 284 (1984) 337-353.

[**10**] S.B. Nadler Jr., Continuum theory, M. Dekker, New York, Basel and Hong Kong, 1992.

[**11**] H. Schirmer, Fixed point sets of continuous selfmaps, in: Fixed Point Theory, Conf. Proc., Sherbrooke, 1980, Lecture Notes in Math. 866 (1981) 417-428.

[**12**] H. Schirmer, Fixed point sets of homeomorphisms of compact surfaces, Israel J. Math. 10 (1971) 373-378.

[**13**] H. Schirmer, On fixed point sets of homeomorphisms of the n-ball, Israel J. Math. 7 (1969) 46-50.

[**14**] L.E. Ward, Jr., Fixed point sets, Pacific J. Math. 47 (1973), 553-565.

Communicated by *Ioan A. Rus*

Saurabh Chandra Maury,
Department of Mathematics,
University of Allahabad, Allahabad-211 002, India
E-mail: smaury94@gmail.com

ELEMENTARY PROOFS OF OPERATOR MONOTONICITY OF SOME FUNCTIONS II

Saichi Izumino and Noboru Nakamura

Received December 27, 2013 ; revised January 28, 2014

Abstract. In the previous paper we gave elementary proofs of operator monotonicity of the representing function of the weighted arithmetic mean and some other related functions. In this note, we show some extensions and applications of those results.

1 Introduction. A (bounded linear) operator A acting on a Hilbert space H is said to be positive, denoted by $A \geq 0$, if $(Av, v) \geq 0$ for all $v \in H$. The definition of positivity induces the order $A \geq B$ for self-adjoint operators A and B on H. A real-valued function f on $(0, \infty)$ is *operator monotone,* if $0 \leq f(A) \leq f(B)$ for operators A and B on H such that $0 \leq A \leq B$. Thus, throughout this paper, we assume that operator monotone functions are positive and their domains are $(0, \infty)$. As a typical example, $x \mapsto x^p$ $(0 \leq p \leq 1)$ is an operator monotone function, which is well-known as *Löwner-Heinz theorem* (*LH*).

For convenience sake, we state the main facts shown in our previous paper with elementary proofs:

Proposition 1.1 (cf. [11, Theorem 1.2], [1], [2], [3], [4], [5], [8], [9], [13]). *The function*

$$a_p(x) = \left(\frac{1+x^p}{2}\right)^{\frac{1}{p}}, \ p \neq 0 \quad \left(a_0(x) = x^{\frac{1}{2}}\right)$$

is operator monotone if (and only if) $-1 \leq p \leq 1$.

Proposition 1.2 (cf. [11, Theorem 1.1], [13], [1]). *The function*

$$s_p(x) = \left(\frac{p(x-1)}{x^p - 1}\right)^{\frac{1}{1-p}}, \ p \neq 0, 1 \quad \left(s_0(x) \left(= \lim_{p \to 0} s_p(x)\right) = \frac{x-1}{\log x}, \ s_1(x) = \frac{1}{e} x^{\frac{x}{x-1}}\right)$$

is operator monotone if $-2 \leq p \leq 2$.

Proposition 1.3 ([11, Theorem 1.3], [5], [9], [6], [2], [3]). *The function*

$$k_p(x) = \frac{p-1}{p} \cdot \frac{x^p - 1}{x^{p-1} - 1}, \ p \neq 0, 1 \ \left(k_0(x) = \frac{x \log x}{x-1}, \ k_1(x) = \frac{x-1}{\log x}\right)$$

is operator monotone if $-1 \leq p \leq 2$.

In this paper, we give some extensions of those propositions and their applications. As an application of the extension of Proposition 1.2, we give a slight extensions of Uchiyama's example in [15] related to Petz-Hasegawa theorem [14].

2010 *Mathematics Subject Classification.* 47A63, 47A64.
Key words and phrases. operator monotone function, operator mean.

2 Preliminaries. By Kubo-Ando theory [12], an operator mean σ is defined as a binary relation of positive operators, satisfying the following properties in common:

(monotonicity)	$A \leq C, B \leq D \Longrightarrow A\sigma B \leq C\sigma D,$
(transformer inequality)	$C(A\sigma B)C \leq (CAC)\sigma(CBC),$
(normality)	$A\sigma A = A,$
(strong operator semi-continuity)	$A_n \downarrow A, B_n \downarrow B \Longrightarrow A_n\sigma B_n \downarrow A\sigma B.$

Sometimes for the definition of an operator mean we must assume operators to be invertible. Without any assumption for invertibility every mean is well-defined as the (strong operator) limits of $(A+\varepsilon I)\sigma(B+\varepsilon I)$ as $\varepsilon \downarrow 0$ instead of $A\sigma B$. (I is the identity operator.)

Every operator mean σ corresponds a unique operator monotone function, that is, its representing function f_σ which is defined by $f_\sigma(x) = 1\sigma x$. Conversely, if f is an operator monotone function with $f(1) = 1$, then the definition of the operator mean corresponding to f is given by

$$A\sigma B = A^{\frac{1}{2}} f\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) A^{\frac{1}{2}}$$

for positive invertible operators A and B.

For our discussion, we use the following basic facts:

(I) For an operator mean σ and for two operator monotone functions g and h, if we define $g\sigma h$ by

$$(g\sigma h)(x) = g(x) f_\sigma\left(\frac{h(x)}{g(x)}\right),$$

then $g\sigma h$ is operator monotone.

(II) For a strictly positive function f on $(0,\infty)$, define $f^\circ(x) := xf(1/x)$ (transpose), $f^*(x) := 1/f(1/x)$ (adjoint) and $f^\perp(x) := x/f(x)$ (dual), then the four functions $f, f^\circ, f^*, f^\perp$ are equivalent to one another with respect to operator monotonicity ([12], [10]).

(III) For a continuous path f_t $(0 \leq t \leq 1)$ of operator monotone functions, its integral mean $\tilde{f}$ defined by

$$\tilde{f}(x) = \int_0^1 f_t(x)dt$$

is an operator monotone function ([2], [3]).

3 Main results. Applying (I) to the operator mean σ_{a_p} corresponding to the operator monotone function $a_p(x)$ (*notice* $a_p(1) = 1$), as an extension of Proposition 1.1, we showed in [11]:

Lemma 3.1 (cf. [11, Lemma 3.1], [13])**.** *Let f, g be operator monotone functions, then $f\sigma_{a_p}g = \left(\frac{f^p+g^p}{2}\right)^{\frac{1}{p}}$ (or equivalently, $(f^p+g^p)^{\frac{1}{p}}$) is operator monotone for $-1 \leq p \leq 1$, $p \neq 0$. Further, if $f_1, ..., f_n$ are operator monotone functions, then $(\sum_{i=1}^n f_i^p)^{\frac{1}{p}}$ is operator monotone. In particular, $(\sum_{i=1}^n (\alpha_i + \beta_i x)^p)^{\frac{1}{p}}$ $(\alpha_i, \beta_i \geq 0)$ is operator monotone.*

Similarly as σ_{a_p}, let σ_{s_p} and σ_{k_p} be the operator means corresponding to the operator monotone functions s_p and k_p, respectively. Then we obtain the following result:

Theorem 3.2 (cf. [11, Theorem 1.1], [13], [1])**.** *For operator monotone functions f, g $(f \neq g)$, the function*

$$f\sigma_{s_p}g = \left(\frac{p(f-g)}{f^p - g^p}\right)^{\frac{1}{1-p}}, \quad p \neq 0, 1 \quad \left(f\sigma_{s_0}g = \frac{f-g}{\log f - \log g}, f\sigma_{s_1}g = \frac{f}{e} \cdot \left(\frac{g}{f}\right)^{\frac{f}{g-f}}\right)$$

is operator monotone if $-2 \le p \le 2$.

Proof. By Proposition 1.2 (for $p \neq 0, 1,$) we have

$$f\sigma_{s_p} g = f \cdot \left(1\sigma_{s_p} \frac{g}{f}\right) = f \cdot \left(\frac{p(\frac{g}{f}-1)}{(\frac{g}{f})^p - 1}\right)^{\frac{1}{1-p}} = \left(\frac{p(f-g)}{f^p - g^p}\right)^{\frac{1}{1-p}}.$$

□

Similarly, we can show:

Theorem 3.3 (cf. [11, Theorem 1.3], [5], [9], [6], [2], [3])**.** *For operator monotone functions* f, g $(f \neq g)$*, the function*

$$f\sigma_{k_p} g = \frac{p-1}{p} \cdot \frac{f^p - g^p}{f^{p-1} - g^{p-1}}, \quad p \neq 0, 1, \left(f\sigma_{k_0} g = \frac{f(\log f - \log g)}{f-g}, \ f\sigma_{k_1} g = \frac{f-g}{\log f - \log g}\right)$$

is operator monotone if $-1 \le p \le 2$.

In [11], the following fact was shown, as an extension of Proposition 1.3:

Lemma 3.4 (cf. [11, Theorem 3.2])**.** *For* $-1 \le p \le 1,\ 0 \le s \le 1,$ *the function*

$$u_{p,s}(x) = \frac{p}{p+s} \cdot \frac{x^{p+s} - 1}{x^p - 1}, \ p \neq 0, -s \ \left(u_{0,s}(x) = \frac{x^s - 1}{\log x^s}, \ u_{-s,s}(x) = \frac{\log x^{-s}}{x^{-s} - 1}\right)$$

is operator monotone.

For the operator mean corresponding to the function $u_{p,s}$, we can obtain the following theorem which is an extension of Theorem 3.3 (and also Lemma 3.4):

Theorem 3.5. *For operator monotone functions* f, g $(f \neq g)$*, and for* $-1 \le p \le 1,\ 0 \le s \le 1,$ *the function*

$$(*) \quad f\sigma_{u_{p,s}} g = \frac{p}{p+s} \cdot \frac{f^{p+s} - g^{p+s}}{f^p - g^p}, \ p \neq 0, -s$$

$$\left(f\sigma_{u_{0,s}} g = \frac{f^s - g^s}{\log f^s - \log g^s}, \ f\sigma_{u_{-s,s}} g = \frac{f^{-s} - g^{-s}}{\log f^{-s} - \log g^{-s}}\right) \quad \textit{is operator monotone.}$$

Example (cf. [15, Example 2.4])**.** For $-1 \le p \le 1,\ 0 \le q - p \le 1,\ p \neq 0,\ q \neq 0$ (and for $a \ge 0$),

$$\frac{p}{q} \cdot \frac{x^q - a^q}{x^p - a^p} \quad \text{is operator monotone.}$$

We can obtain this fact, by putting $f = x,\ g = a,$ and $q = p + s$ in $(*)$.

As an application of Proposition 1.2, we showed an alternative proof of the following result due to Petz and Hasegawa [14], [6]:

Proposition 3.6 (cf. [11, Theorem 3.4])**.** *For* $-1 \le p \le 2$

$$h_p(x) = \frac{p(1-p)(x-1)^2}{(x^p - 1)(x^{1-p} - 1)}, \ p \neq 0, 1 \ \left(h_0(x) = h_1(x) = \frac{x-1}{\log x}\right)$$

is operator monotone.

As an extension of this fact and an application of Theorem 3.2, though the range of p is reduced, we have:

Theorem 3.7. *If f, g, k, l $(f \neq g, k \neq l)$ are operator monotone functions, then for $0 < p < 1$,*

$$\frac{(f-g)(k-l)}{(f^p-g^p)(k^{1-p}-l^{1-p})} \text{ is operator monotone.}$$

Proof. Since $f\sigma_{s_p}g$ and $k\sigma_{s_{1-p}}l$ are operator monotone, we see $\frac{1}{p(1-p)} \cdot (f\sigma_{s_p}g)\sharp_p(k\sigma_{s_{1-p}}l) = \frac{(f-g)(k-l)}{(f^p-g^p)(k^{1-p}-l^{1-p})}$ is operator monotone. □

Example (cf. [15, Theorem 2.7])**.** Putting $f = k = x$ and $g = a,\ l = b$ $(a, b \geq 0)$, we see that $\frac{(x-a)(x-b)}{(x^p-a^p)(x^{1-p}-b^{1-p})}$ is operator monotone.

Further, we have:

Theorem 3.8. *For $-1 \leq p \leq 2,\ a, b \geq 0$*

$$(**) \qquad h_p(a, b; x) = \frac{p(1-p)(x-a)(x-b)}{(x^p-a^p)(x^{1-p}-b^{1-p})} \text{ is operator monotone.}$$

Proof. We may prove the theorem for $p \neq 0, \pm 1, 2$ and $a, b > 0$. For the case $0 < p < 1$, then $(**)$ is clear. There remain the two cases:

(i) If $1 < p < 2$, then we put $p = q + 1$, so that $0 < q < 1$. We have:

$$h_p(a, b; x) = h_{q+1}(a, b; x) = (-q)(q+1) \cdot \frac{(x-a)(x-b)}{(x^{q+1}-a^{q+1})(x^{-q}-b^{-q})}$$

$$= \frac{q(q+1)b^q x^q (x-a)(x-b)}{(x^{q+1}-a^{q+1})(x^q-b^q)}.$$

Now since $0 < q < 1$, we see that $\left(\frac{q(x-b)}{x^q-b^q}\right)^{\frac{1}{1-q}}$ is operator monotone by Proposition 1.2. Further, since $1 < q + 1 < 2$, we see that

$$(\eta(a, b; x) :=) \ \left(\frac{(q+1)(x-a)}{x^{q+1}-a^{q+1}}\right)^{\frac{1}{1-(q+1)}} = \left(\frac{(q+1)(x-a)}{x^{q+1}-a^{q+1}}\right)^{-\frac{1}{q}}$$

is operator monotone by Proposition 1.2, so that its dual $(\eta^{\perp}(a, b; x) =) \ x \cdot \left(\frac{(q+1)(x-a)}{x^{q+1}-a^{q+1}}\right)^{\frac{1}{q}}$ is operator monotone. Hence

$$\left\{ \left(\frac{q(x-b)}{x^q-b^q}\right)^{\frac{1}{1-q}} \sharp_q \ x \cdot \left(\frac{(q+1)(x-a)}{x^{q+1}-a^{q+1}}\right)^{\frac{1}{q}} \right\} \times b^q = h_p(a, b; x)$$

is operator monotone.

(ii) If $-1 < p < 0$, then putting $p = -q$, we can similarly prove $(**)$. □

Acknowledgment. The authors would like to express their hearty thanks to the referee for valuable advice.

References

[1] A. Besenyei and D. Petz, *Completely positive mappings and mean matrices*, Linear Algebra Appl., **435** (2011), 984-997.

[2] J.I. Fujii, *Interpolationality for symmetric operator means*, Sci. Math. Japon., **75**, No.3 (2012), 267-274.

[3] J.I. Fujii and M. Fujii, *Upper estimations on integral operator means*, Sci. Math. Japon., **75**, No.2 (2012), 217-222.

[4] J.I. Fujii, M. Nakamura and S. -E. Takahashi, *Cooper's approach to chaotic operator means*, Sci. Math. Japon., **63** (2006), 319-324.

[5] J.I. Fujii and Y. Seo, *On parametrized operator means dominated by power ones*, Sci. Math. **1** (1998), 301-306.

[6] T. Furuta, *Concrete examples of operator monotone functions obtained by an elementary method without appealing to Löwner integral representation*, Linear Algebra Appl., **429** (2008), 972-980.

[7] T. Furuta, *Elementary proof of Petz-Hasegawa Theorem*, Lett. Math. Phys., **101** (2012), 355-359.

[8] T. Furuta, J. Mićič Hot, J. Pečarić and Y. Seo, Mond-Pečarić Method in Operator Inequalities, Element, Zagreb, 2005.

[9] F. Hiai and H. Kosaki, *Means for matrices and comparison of their norms*, Indiana Univ. Math. J., **48** (1999), 899-936.

[10] F. Hiai and K. Yanagi, *Hilbert spaces and linear operators*, Makino Syoten, (1995), (in Japanese).

[11] S. Izumino and N. Nakamura, *Elementary proofs of operator monotonicity of some functions*, Sci. Math. Japon., Online, **e-2013**, 679-686.

[12] F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann., **246** (1980), 205-224.

[13] Y. Nakamura, *Classes of operator monotone functions and Stieltjes functions*, In: Dym H. et al., (eds) The Gohberg Anniversary Collection, Vol. II: Topics in Analysis and Operator Theory, Operator Theory: Advances and Appl., **Vol. 41** Birkhäuser, Basel, (1989), 395-404.

[14] D. Petz and H. Hasegawa, *On the Riemannian metric of α-entropies of density matrices*, Lett. Math. Phys., **38** (1996), 221-225.

[15] M. Uchiyama, *Majorizaition and some operator functions*, Linear Algebra and Appl., **432** (2010), 1867-1872.

Communicated by *Jun Ichi Fujii*

Saichi Izumino:
University of Toyama, Gofuku, Toyama, 930-8555, Japan
Email: saizumino@h5.dion.ne.jp

Noboru Nakamura:
Toyama National College of Technology, Hongo-machi 13, Toyama, 939-8630, Japan
Email: n-nakamu@nc-toyama.ac.jp

APPROXIMATELY DERIVATIVE IN A VECTOR LATTICE

Toshiharu Kawasaki

Received September 29, 2013 ; revised February 7, 2014

Abstract. In previous paper we defined the derivative of mappings from a vector lattice into a complete vector lattice. In this paper we define an approximately derivative of mappings from a vector lattice into a complete vector lattice. Moreover we consider a relation between these two derivatives.

1 Introduction The purpose of our researches is to consider some derivatives and some integrals of mappings in vector spaces and to study their relations, for instance, the fundamental theorem of calculus, inclusive relations between integrals and so on; see [9–17].

When we consider extending from restricted Denjoy integral to improper Denjoy integral for real valued functions, the derivative is transposed to more general derivative, called approximately derivative. Therefore in this paper we consider approximately derivative for mappings from a vector lattice into a vector lattice.

In [15] we defined the derivative of mappings from a vector lattice into a complete vector lattice. In [12] we defined the approximately derivative in the case where the domain is finite dimension. This derivative seemed to be a subset of bounded linear mappings generally, however in [14] it was proved that the subset consists of a single point. In this paper we consider an approximately derivative of mappings from a vector lattice into a complete vector lattice. Moreover we consider a relation between these two derivatives.

In this paper we use notation and definitions in [15, 16]. Let X be a vector lattice. An element $e \in X$ is said to be a unit if $e \wedge x > 0$ for any $x \in X$ with $x > 0$. Let $\mathcal{K}_X$ be the class of units of X. Let $\mathcal{I}_X$ be the class of intervals of X and $\mathcal{IK}_X$ the class of intervals $[a, b]$ with $b - a \in \mathcal{K}_X$. Let $\mathcal{L}(X, Y)$ be the class of bounded linear mappings from X into a vector lattice Y. If Y is complete, then $\mathcal{L}(X, Y)$ is also so [2, 20, 24, 25]. A subset $D \subset X$ is said to be open if for any $x \in D$ and for any $e \in \mathcal{K}_X$ there exists $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ such that $[x - \varepsilon e, x + \varepsilon e] \subset D$. Let $\mathcal{O}_X$ be the class of open subsets of X. For an interval $[a, b]$ and $e \in \mathcal{K}_X$ let

$$[a, b]^e = \{x \mid \text{ there exists } \varepsilon \in \mathcal{K}_{\mathbb{R}} \text{ such that } x - a \geq \varepsilon e \text{ and } b - x \geq \varepsilon e\}.$$

Let Λ be an upward directed set. Then let $\mathcal{U}_X(\Lambda)$ be the class of $\{v_\lambda \mid \lambda \in \Lambda\}$ which satisfies the following conditions:

(U1) $v_\lambda \in X$ with $v_\lambda > 0$;

$(\mathrm{U2})^u$ $v_{\lambda_1} \geq v_{\lambda_2}$ if $\lambda_1 \leq \lambda_2$;

(U3) $\bigwedge_{\lambda \in \Lambda} v_\lambda = 0$.

Moreover we consider the following condition:

(M) There exists an interval function $q : \mathcal{I}_X \longrightarrow [0, \infty)$ such that

2010 *Mathematics Subject Classification.* 46G05, 46G12.
Key words and phrases. derivative, approximately derivative, vector lattice, Riesz space.

(M1) $q(I_1) \leq q(I_2)$ if $I_1 \subset I_2$;

(M2) $q(I) > 0$ if $I \in \mathcal{IK}_X$;

(M3) For any $x \in X$, for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $\delta \in \mathcal{K}_{\mathbb{R}}$ such that $q([x, x + \delta e]) \leq \varepsilon$ and $q([x - \delta e, x]) \leq \varepsilon$.

Example 1.1. Let X be a Banach lattice, that is, it satisfies that $|a| \leq |b|$ implies $\|a\| \leq \|b\|$. Suppose that $\mathcal{K}_X \neq \emptyset$. For any $a, b \in X$ with $a \leq b$ let $q([a, b]) = \|b - a\|$. Then X endowed with q satisfies (M). Indeed, if $[a, b] \subset [c, d]$, then $0 \leq b - a \leq d - c$ and hence $q([a, b]) = \|b - a\| \leq \|d - c\| = q([c, d])$. If $b - a \in \mathcal{K}_X$, then $a \neq b$ and hence $q([a, b]) = \|b - a\| > 0$. Moreover for any $x \in X$, for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$, taking $\delta \leq \frac{\varepsilon}{\|e\|}$, then it holds that $q([x, x + \delta e]) = \delta\|e\| \leq \varepsilon$ and $q([x - \delta e, x]) = \delta\|e\| \leq \varepsilon$. For instance, since $C(K)$, where K is a compact Hausdorff space, and L^p, which $1 \leq p \leq \infty$, are Banach lattices with unit, these spaces endowed with the above q satisfy (M).

Example 1.2. Let $X = \mathbb{R}^d \times X_1$, where X_1 is any vector lattice with unit. For any $a = ((a_1, \ldots, a_d), a'), b = ((b_1, \ldots, b_d), b') \in X$ we define $a \leq b$ whenever $a_i \leq b_i$ for any $i = 1, \ldots, d$ and $a' \leq b'$. Then $\mathcal{K}_X = \{((e_1, \ldots, e_d), e') \mid e_i > 0 \text{ for any } i = 1, \ldots, d \text{ and } e' \in \mathcal{K}_{X_1}\}$. Moreover for any $a = ((a_1, \ldots, a_d), a'), b = ((b_1, \ldots, b_d), b') \in X$ with $a \leq b$ let $q([a, b]) = \prod_{i=1}^{d}(b_i - a_i)$. Then X endowed with q satisfies (M). Indeed, if $[a, b] \subset [c, d]$, then $b_i - a_i \leq d_i - c_i$ for any $i = 1, \ldots, d$ and hence $q([a, b]) \leq q([c, d])$. If $b - a \in \mathcal{K}_X$, then $a_i < b_i$ for any $i = 1, \ldots, d$ and hence $q([a, b]) > 0$. Moreover for any $x \in X$, for any $e = ((e_1, \ldots, e_d), e') \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$, taking $\delta \leq \frac{\varepsilon}{\prod_{i=1}^{d} e_i}$, then it holds that $q([x, x + \delta e]) = \delta \prod_{i=1}^{d} e_i \leq \varepsilon$ and $q([x - \delta e, x]) = \delta \prod_{i=1}^{d} e_i \leq \varepsilon$. For instance, since $\mathbb{R}^S$, where S is an arbitrary nonempty set, is such a space, this space endowed with the above q satisfies (M).

In general a lot of interval functions satisfying (M) in X can be considered. Hereafter in the case of $X = \mathbb{R}^d$ we always consider the Lebesgue measure as an interval function q.

2 Definitions

Definition 2.1. Let X be a vector lattice with unit, $x_0 \in D \in \mathcal{O}_X$ and $E \subset D$. Suppose that X satisfies (M).

x_0 is said to be a right density point of E if for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_1 \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies the following conditions:

(RDS) $E^C \cap [x_0, x_0 + h] \subset \bigcup_{k=1}^{\infty}[a_k, b_k]^e$.

(RD) $\sum_{k=1}^{\infty} q([a_k, b_k]) \leq \varepsilon q([x_0, x_0 + h])$.

x_0 is said to be a left density point of E if for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_1 \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies the following conditions:

(LDS) $E^C \cap [x_0 - h, x_0] \subset \bigcup_{k=1}^{\infty}[a_k, b_k]^e$.

(LD) $\sum_{k=1}^{\infty} q([a_k, b_k]) \leq \varepsilon q([x_0 - h, x_0])$.

x_0 is said to be a density point of E if it is a right density point and a left density point.

x_0 is said to be a right dispersion point of E if for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_1 \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies (RD) and the following condition:

(RDP) $E \cap [x_0, x_0 + h] \subset \bigcup_{k=1}^{\infty} [a_k, b_k]^e$.

x_0 is said to be a left dispersion point of E if for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_1 \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies (LD) and the following condition:

(LDP) $E \cap [x_0 - h, x_0] \subset \bigcup_{k=1}^{\infty} [a_k, b_k]^e$.

x_0 is said to be a dispersion point of E if it is a right dispersion point and a left dispersion point.

Definition 2.2. Let X be a vector lattice with unit, Y a complete vector lattice, $D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies (M).

For any $l \in \mathcal{L}(X, Y)$ and for any right density point x_0 of $\{x \mid x \in D, x - x_0 \in \mathcal{K}_X\}$ let

$$\begin{aligned} E^+_{sup}(l; F, x_0) &= \{x \mid x \in D, x - x_0 \in \mathcal{K}_X, F(x) - F(x_0) \not< l(x - x_0)\}, \\ L^+_{sup}(F, x_0) &= \left\{ l \;\middle|\; \begin{array}{l} l \in \mathcal{L}(X, Y), \\ x_0 \text{ is a right dispersion point of } E^+_{sup}(l; F, x_0) \end{array} \right\} \end{aligned}$$

and $o\text{-}\overline{AD}^+ F(x_0)$ the class of $l \in \mathcal{L}(X, Y)$ which satisfies the following conditions:

(a-S1$_R$) For any $l' \in \mathcal{L}(X, Y)$ with $l' > 0$ there exists $l'' \in L^+_{sup}(F, x_0)$ such that $l \leq l'' < l + l'$.

(a-S2$_R$) $l'' \not< l$ for any $l'' \in L^+_{sup}(F, x_0)$.

Let

$$\begin{aligned} E^+_{inf}(l; F, x_0) &= \{x \mid x \in D, x - x_0 \in \mathcal{K}_X, F(x) - F(x_0) \not> l(x - x_0)\}, \\ L^+_{inf}(F, x_0) &= \left\{ l \;\middle|\; \begin{array}{l} l \in \mathcal{L}(X, Y), \\ x_0 \text{ is a right dispersion point of } E^+_{inf}(l; F, x_0) \end{array} \right\} \end{aligned}$$

and $o\text{-}\underline{AD}^+ F(x_0)$ the class of $l \in \mathcal{L}(X, Y)$ which satisfies the following conditions:

(a-I1$_R$) For any $l' \in \mathcal{L}(X, Y)$ with $l' > 0$ there exists $l'' \in L^+_{inf}(F, x_0)$ such that $l \geq l'' > l - l'$.

(a-I2$_R$) $l'' \not> l$ for any $l'' \in L^+_{inf}(F, x_0)$.

For any $l \in \mathcal{L}(X, Y)$ and for any left density point x_0 of $\{x \mid x \in D, x_0 - x \in \mathcal{K}_X\}$ let

$$\begin{aligned} E^-_{sup}(l; F, x_0) &= \{x \mid x \in D, x_0 - x \in \mathcal{K}_X, F(x_0) - F(x) \not< l(x_0 - x)\}, \\ L^-_{sup}(F, x_0) &= \left\{ l \;\middle|\; \begin{array}{l} l \in \mathcal{L}(X, Y), \\ x_0 \text{ is a left dispersion point of } E^-_{sup}(l; F, x_0) \end{array} \right\} \end{aligned}$$

and $o\text{-}\overline{AD}^- F(x_0)$ the class of $l \in \mathcal{L}(X, Y)$ which satisfies the following conditions:

(a-S1$_L$) For any $l' \in \mathcal{L}(X, Y)$ with $l' > 0$ there exists $l'' \in L^-_{sup}(F, x_0)$ such that $l \leq l'' < l + l'$.

(a-S2$_L$) $l'' \not< l$ for any $l'' \in L^-_{sup}(F, x_0)$.

Let

$$\begin{aligned} E^-_{inf}(l;F,x_0) &= \{x \mid x \in D, x_0 - x \in \mathcal{K}_X, F(x_0) - F(x) \not> l(x_0 - x)\}, \\ L^-_{inf}(F,x_0) &= \left\{ l \;\middle|\; \begin{array}{l} l \in \mathcal{L}(X,Y), \\ x_0 \text{ is a left dispersion point of } E^-_{inf}(l;F,x_0) \end{array} \right\} \end{aligned}$$

and $o\text{-}\underline{AD}^- F(x_0)$ the class of $l \in \mathcal{L}(X,Y)$ which satisfies the following conditions:

(a-I1$_L$) For any $l' \in \mathcal{L}(X,Y)$ with $l' > 0$ there exists $l'' \in L^-_{inf}(F,x_0)$ such that $l \geq l'' > l - l'$.

(a-I2$_L$) $l'' \not> l$ for any $l'' \in L^-_{inf}(F,x_0)$.

F is said to be approximately right upper differentiable, approximately right lower differentiable, approximately left upper differentiable and approximately left lower differentiable at x_0 if $o\text{-}\overline{AD}^+ F(x_0)$, $o\text{-}\underline{AD}^+ F(x_0)$, $o\text{-}\overline{AD}^- F(x_0)$ and $o\text{-}\underline{AD}^- F(x_0)$ are not empty, respectively. If $o\text{-}AD^+ F(x_0) = o\text{-}\overline{AD}^+ F(x_0) \cap o\text{-}\underline{AD}^+ F(x_0)$ and $o\text{-}AD^- F(x_0) = o\text{-}\overline{AD}^- F(x_0) \cap o\text{-}\underline{AD}^- F(x_0)$ are not empty, then F is said to be approximately right differentiable and approximately left differentiable at x_0, respectively. If $o\text{-}ADF(x_0) = o\text{-}AD^+ F(x_0) \cap o\text{-}AD^- F(x_0)$ is not empty, then F is said to be approximately differentiable at x_0.

3 Properties

Theorem 3.1. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies* (M)*.*

(1) *If F is approximately right upper differentiable at right density point x_0 of $\{x \mid x \in D, x - x_0 \in \mathcal{K}_X\}$, then any two different elements in $o\text{-}\overline{AD}^+ F(x_0)$ are incomparable.*

(2) *If F is approximately right lower differentiable at right density point x_0 of $\{x \mid x \in D, x - x_0 \in \mathcal{K}_X\}$, then any two different elements in $o\text{-}\underline{AD}^+ F(x_0)$ are incomparable.*

(3) *If F is approximately left upper differentiable at left density point x_0 of $\{x \mid x \in D, x_0 - x \in \mathcal{K}_X\}$, then any two different elements in $o\text{-}\overline{AD}^- F(x_0)$ are incomparable.*

(4) *If F is approximately left lower differentiable at left density point x_0 of $\{x \mid x \in D, x_0 - x \in \mathcal{K}_X\}$, then any two different elements in $o\text{-}\underline{AD}^- F(x_0)$ are incomparable.*

Proof. Assume that $l_1 < l_2$ for $l_1, l_2 \in o\text{-}\overline{AD}^+ F(x_0)$. By (a-S1$_R$) for l_1 there exists $l'' \in L^+_{sup}(F,x_0)$ such that $l_1 \leq l'' < l_1 + (l_2 - l_1) = l_2$. However it is a contradiction to (a-S2$_R$) for l_2, that is, $l'' \not< l_2$ for any $l'' \in L^+_{sup}(F,x_0)$. Therefore any two different elements in $o\text{-}\underline{AD}^+ F(x_0)$ must be incomparable. The rest can be proved similarly. □

Lemma 3.1. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies* (M)*.*

(1) *If $l \in L^{\pm}_{sup}(F,x_0)$ and $l' \in \mathcal{L}(X,Y)$ with $l' > l$, then $l' \in L^{\pm}_{sup}(F,x_0)$.*

(2) *If $l \in L^{\pm}_{inf}(F,x_0)$ and $l' \in \mathcal{L}(X,Y)$ with $l' < l$, then $l' \in L^{\pm}_{inf}(F,x_0)$.*

Proof. It is clear by definition. □

Let X be a vector lattice and $A, B \subset X$. We write $A \leq B$ if $a \leq b$ for any $a \in A$ and for any $b \in B$. Similarly we write $A < B$ and $A \not\leq B$ if $a < b$ and $a \not\leq b$, respectively, for any $a \in A$ and for any $b \in B$, and so on. Moreover we write $A \preceq B$ if for any $a \in A$ there exists $b \in B$ such that $a \leq b$ and if for any $b \in B$ there exists $a \in A$ such that $a \leq b$.

Lemma 3.2. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies* (M)*.*

Then

(1) $\quad L^{\pm}_{inf}(F, x_0) \cap L^{\pm}_{sup}(F, x_0) = \emptyset.$

(2) $\quad L^{\pm}_{inf}(F, x_0) \not> L^{\pm}_{sup}(F, x_0).$

(3) $\quad L^{\pm}_{inf}(F, x_0) \preceq L^{\pm}_{sup}(F, x_0).$

Proof. (1) Assume that $L^{+}_{inf}(F, x_0) \cap L^{+}_{sup}(F, x_0) \neq \emptyset$. Let $l \in L^{+}_{inf}(F, x_0) \cap L^{+}_{sup}(F, x_0)$. Then x_0 is a right dispersion point of $E^{+}_{inf}(l; F, x_0)$ and $E^{+}_{sup}(l; F, x_0)$, that is, for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_{inf} \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_{inf}$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies

$$\begin{aligned} E^{+}_{inf}(l; F, x_0) \cap [x_0, x_0 + h] &\subset \bigcup_{k=1}^{\infty} [a_k, b_k]^e, \\ \sum_{k=1}^{\infty} q([a_k, b_k]) &\leq \varepsilon q([x_0, x_0 + h]), \end{aligned}$$

and there exists $e_{sup} \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_{sup}$ there exists $\{[c_k, d_k] \mid k = 1, 2, \ldots\}$ which satisfies

$$\begin{aligned} E^{+}_{sup}(l; F, x_0) \cap [x_0, x_0 + h] &\subset \bigcup_{k=1}^{\infty} [c_k, d_k]^e, \\ \sum_{k=1}^{\infty} q([c_k, d_k]) &\leq \varepsilon q([x_0, x_0 + h]). \end{aligned}$$

Let $e_1 = e_{inf} \wedge e_{sup}$. Then the above two inequalities are true for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$. Since

$$E^{+}_{inf}(l; F, x_0) \cup E^{+}_{sup}(l; F, x_0) = \{x \mid x \in D, x - x_0 \in \mathcal{K}_X\},$$

it holds that

$$\begin{aligned} &(E^{+}_{inf}(l; F, x_0) \cup E^{+}_{sup}(l; F, x_0)) \cap [x_0, x_0 + h] \\ &\quad = \{x \mid x \in D, x - x_0 \in \mathcal{K}_X\} \cap [x_0, x_0 + h]. \end{aligned}$$

Therefore

$$\begin{aligned} \{x \mid x \in D, x - x_0 \in \mathcal{K}_X\} \cap [x_0, x_0 + h] &\subset \bigcup_{k=1}^{\infty} ([a_k, b_k]^e \cup [c_k, d_k]^e), \\ \sum_{k=1}^{\infty} q([a_k, b_k]) + \sum_{k=1}^{\infty} q([c_k, d_k]) &\leq 2\varepsilon q([x_0, x_0 + h]). \end{aligned}$$

It is a contradiction to that x_0 is a right density point of $\{x \mid x \in D, x-x_0 \in \mathcal{K}_X\}$. Therefore $L^+_{inf}(F,x_0) \cap L^+_{sup}(F,x_0) = \emptyset$. It can be proved similarly that $L^-_{inf}(F,x_0) \cap L^-_{sup}(F,x_0) = \emptyset$.
(2) Assume that $l_1 \in L^\pm_{inf}(F,x_0)$, $l_2 \in L^\pm_{sup}(F,x_0)$ and $l_1 > l_2$. By Lemma 3.1 it holds that $l_1 \in L^\pm_{sup}(F,x_0)$ and $l_2 \in L^\pm_{inf}(F,x_0)$. However it is a contradiction to (1). Therefore $L^\pm_{inf}(F,x_0) \not\succ L^\pm_{sup}(F,x_0)$.
(3) Let $l_1 \in L^\pm_{inf}(F,x_0)$ and $l_2 \in L^\pm_{sup}(F,x_0)$. By Lemma 3.1 it holds that $l_1 \vee l_2 \in L^\pm_{sup}(F,x_0)$ and $l_1 \le l_1 \vee l_2$. By Lemma 3.1 it holds that $l_1 \wedge l_2 \in L^\pm_{inf}(F,x_0)$ and $l_1 \wedge l_2 \le l_2$. Therefore $L^\pm_{inf}(F,x_0) \preceq L^\pm_{sup}(F,x_0)$. □

Theorem 3.2. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies* (M).

(1) *If F is approximately right upper differentiable and approximately right lower differentiable at right density point x_0 of $\{x \mid x \in D, x-x_0 \in \mathcal{K}_X\}$, then $o\text{-}\overline{AD}^+F(x_0) \not< o\text{-}\underline{AD}^+F(x_0)$.*

(2) *If F is approximately left upper differentiable and approximately left lower differentiable at left density point x_0 of $\{x \mid x \in D, x_0-x \in \mathcal{K}_X\}$, then $o\text{-}\overline{AD}^-F(x_0) \not< o\text{-}\underline{AD}^-F(x_0)$.*

Proof. Assume that $l_1 \in o\text{-}\overline{AD}^+F(x_0)$, $l_2 \in o\text{-}\underline{AD}^+F(x_0)$ and $l_1 < l_2$. Let $l = \frac{1}{2}(l_1+l_2)$. Then $l_1 < l < l_2$. By (a-S1$_R$) for l_1 there exists $l_1'' \in L^+_{sup}(F,x_0)$ such that $l_1 \le l_1'' < l$. By (a-I1$_R$) for l_2 there exists $l_2'' \in L^+_{inf}(F,x_0)$ such that $l_2 \ge l_2'' > l$. By Lemma 3.1 l is belonging to both $L^+_{sup}(F,x_0)$ and $L^+_{inf}(F,x_0)$, however it is a contradiction to Lemma 3.2. Therefore $o\text{-}\overline{AD}^+F(x_0) \not< o\text{-}\underline{AD}^+F(x_0)$. It can be proved similarly that $o\text{-}\overline{AD}^-F(x_0) \not< o\text{-}\underline{AD}^-F(x_0)$. □

Lemma 3.3. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$, F, F_1, F_2 mappings from D into Y and $\alpha \in \mathbb{R}$. Suppose that X satisfies* (M).
Then

(1)

$$L^\pm_{sup}(\alpha F,x_0) = \begin{cases} \alpha L^\pm_{sup}(F,x_0) & \text{if } \alpha \ge 0, \\ \alpha L^\pm_{inf}(F,x_0) & \text{if } \alpha < 0. \end{cases}$$
$$L^\pm_{inf}(\alpha F,x_0) = \begin{cases} \alpha L^\pm_{inf}(F,x_0) & \text{if } \alpha \ge 0, \\ \alpha L^\pm_{sup}(F,x_0) & \text{if } \alpha < 0. \end{cases}$$

(2)

$$L^\pm_{sup}(F_1,x_0) + L^\pm_{sup}(F_2,x_0) \subset L^\pm_{sup}(F_1+F_2,x_0),$$
$$L^\pm_{inf}(F_1,x_0) + L^\pm_{inf}(F_2,x_0) \subset L^\pm_{inf}(F_1+F_2,x_0).$$

Proof. (1) is clear by definition. We show (2). Let $l_1 \in L^+_{sup}(F_1,x_0)$ and $l_2 \in L^+_{sup}(F_2,x_0)$. If

$$F_1(x) - F_1(x_0) + F_2(x) - F_2(x_0) \not< l_1(x-x_0) + l_2(x-x_0),$$

then

$$F_1(x) - F_1(x_0) \not< l_1(x-x_0) \text{ or } F_2(x) - F_2(x_0) \not< l_2(x-x_0).$$

Therefore

$$E^+_{sup}(l_1+l_2;F_1+F_2,x_0)\subset E^+_{sup}(l_1;F_1,x_0)\cup E^+_{sup}(l_2;F_2,x_0).$$

If x_0 is a right dispersion point of $E^+_{sup}(l_1;F_1,x_0)$ and of $E^+_{sup}(l_2;F_2,x_0)$, then it is right dispersion point of $E^+_{sup}(l_1+l_2;F_1+F_2,x_0)$. Therefore $l_1+l_2\in L^+_{sup}(F_1+F_2,x_0)$. The rest can be proved similarly. □

Theorem 3.3. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0\in D\in\mathcal{O}_X$, F,F_1,F_2 mappings from D into Y and $\alpha\in\mathbb{R}$ with $\alpha>0$. Suppose that X satisfies* (M)*.*

(1) *If F is approximately right upper differentiable at right density point x_0 of $\{x\mid x\in D, x-x_0\in\mathcal{K}_X\}$, then αF is also so and $o\text{-}\overline{AD}^+(\alpha F)(x_0)=\alpha o\text{-}\overline{AD}^+F(x_0)$ and $-\alpha F$ is approximately right lower differentiable at x_0 and $o\text{-}\overline{AD}^+(-\alpha F)(x_0)=-\alpha o\text{-}\underline{AD}^+F(x_0)$. If F is approximately right lower differentiable at x_0, then αF is also so and $o\text{-}\underline{AD}^+(\alpha F)(x_0)=\alpha o\text{-}\underline{AD}^+F(x_0)$ and $-\alpha F$ is approximately right upper differentiable at x_0 and $o\text{-}\underline{AD}^+(-\alpha F)(x_0)=-\alpha o\text{-}\overline{AD}^+F(x_0)$.*

(2) *If F_1,F_2,F_1+F_2 are approximately right upper differentiable at right density point x_0 of $\{x\mid x\in D, x-x_0\in\mathcal{K}_X\}$, then*

$$o\text{-}\overline{AD}^+F_1(x_0)+o\text{-}\overline{AD}^+F_2(x_0)\not< o\text{-}\overline{AD}^+(F_1+F_2)(x_0).$$

If F_1,F_2,F_1+F_2 are approximately right lower differentiable at x_0, then

$$o\text{-}\underline{AD}^+F_1(x_0)+o\text{-}\underline{AD}^+F_2(x_0)\not> o\text{-}\underline{AD}^+(F_1+F_2)(x_0).$$

(3) *If F is approximately left upper differentiable at left density point x_0 of $\{x\mid x\in D, x_0-x\in\mathcal{K}_X\}$, then αF is also so and*

$$o\text{-}\overline{AD}^-(\alpha F)(x_0)=\alpha o\text{-}\overline{AD}^-F(x_0)$$

and $-\alpha F$ is approximately left lower differentiable at x_0 and

$$o\text{-}\overline{AD}^-(-\alpha F)(x_0)=-\alpha o\text{-}\underline{AD}^-F(x_0).$$

If F is approximately left lower differentiable at x_0, then αF is also so and

$$o\text{-}\underline{AD}^-(\alpha F)(x_0)=\alpha o\text{-}\underline{AD}^-F(x_0)$$

and $-\alpha F$ is approximately left upper differentiable at x_0 and

$$o\text{-}\underline{AD}^-(-\alpha F)(x_0)=-\alpha o\text{-}\overline{AD}^-F(x_0).$$

(4) *If F_1,F_2,F_1+F_2 are approximately left upper differentiable at left density point x_0 of $\{x\mid x\in D, x_0-x\in\mathcal{K}_X\}$, then*

$$o\text{-}\overline{AD}^-F_1(x_0)+o\text{-}\overline{AD}^-F_2(x_0)\not< o\text{-}\overline{AD}^-(F_1+F_2)(x_0).$$

If F_1,F_2,F_1+F_2 are approximately left lower differentiable at x_0, then

$$o\text{-}\underline{AD}^-F_1(x_0)+o\text{-}\underline{AD}^-F_2(x_0)\not> o\text{-}\underline{AD}^-(F_1+F_2)(x_0).$$

Proof. (1) and (3) are clear by definition. We show (2) and (4). Let $l_1 \in o\text{-}\overline{AD}^+F_1(x_0)$ and $l_2 \in o\text{-}\overline{AD}^+F_2(x_0)$. By (a-S1$_R$) for any $l' \in \mathcal{L}(X,Y)$ with $l' > 0$ there exist $l_1'' \in L^+_{sup}(F_1, X_0)$ and $l_2'' \in L^+_{sup}(F_2, X_0)$ such that $l_1 \leq l_1'' < l_1 + l'$ and $l_2 \leq l_2'' < l_2 + l'$. Since $F_1 + F_2$ is also approximately right upper differentiable at x_0, by (a-S2$_R$) it holds that $l'' \not< l$ for any $l \in o\text{-}\overline{AD}^+(F_1+F_2)(x_0)$ and for any $l'' \in L^+_{sup}(F_1+F_2, X_0)$. Since by Lemma 3.3 $l_1'' + l_2'' \in L^+_{sup}(F_1+F_2, x_0)$, it holds that $l_1 + l_2 \leq l_1'' + l_2'' \not< l$. Note that l_1'' and l_2'' can take near l_1 and l_2 enough. Therefore $l_1 + l_2 \not< l$. Actually assume that $l_1 + l_2 < l$. Then $l_1'' + l_2'' < l_1 + l_2 + 2l' < l$ for any $l' < \frac{1}{2}(l - l_1 - l_2)$. It is a contradiction. Therefore

$$o\text{-}\overline{AD}^+F_1(x_0) + o\text{-}\overline{AD}^+F_2(x_0) \not< o\text{-}\overline{AD}^+(F_1+F_2)(x_0).$$

The rest can be proved similarly. □

Lemma 3.4. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y. Suppose that X satisfies* (M).

(1) *$l \in o\text{-}AD^+F(x_0)$ if and only if l satisfies* (a-S1$_R$) *and* (a-I1$_R$).

(2) *$l \in o\text{-}AD^-F(x_0)$ if and only if l satisfies* (a-S1$_L$) *and* (a-I1$_L$).

Proof. The necessity is clear. We show the sufficiency. We show that if $l \in \mathcal{L}(X,Y)$ satisfies (a-S1$_R$), then it satisfies (a-I2$_R$). Assume that l does not satisfy (a-I2$_R$). Then there exists $l'' \in L^+_{inf}(F, x_0)$ such that $l'' > l$. By (a-S1$_R$) there exists $l''' \in L^+_{sup}(F, x_0)$ such that $l \leq l''' < l''$. It is a contradiction to Lemma 3.2. Therefore l satisfies (a-I2$_R$). The rest can be proved similarly. □

Theorem 3.4. *Let X be a vector lattice with unit, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and F_1, F_2 mappings from D into Y. Suppose that X satisfies* (M).

(1) *If F_1 and F_2 are approximately right differentiable at right density point x_0 of $\{x \mid x \in D, x - x_0 \in \mathcal{K}_X\}$, then $F_1 + F_2$ is also so and*

$$o\text{-}AD^+F_1(x_0) + o\text{-}AD^+F_2(x_0) = o\text{-}AD^+(F_1+F_2)(x_0).$$

(2) *If F_1 and F_2 are approximately left differentiable at left density point x_0 of $\{x \mid x \in D, x_0 - x \in \mathcal{K}_X\}$, then $F_1 + F_2$ is also so and*

$$o\text{-}AD^-F_1(x_0) + o\text{-}AD^-F_2(x_0) = o\text{-}AD^-(F_1+F_2)(x_0).$$

Proof. Let $l_1 \in o\text{-}AD^+F_1(x_0)$ and $l_2 \in o\text{-}AD^+F_2(x_0)$. For any $l' \in \mathcal{L}(X,Y)$ with $l' > 0$ there exist $l_1'' \in L^+_{sup}(F_1, x_0)$ and $l_2'' \in L^+_{sup}(F_2, x_0)$ such that $l_1 \leq l_1'' < l_1 + l'$ and $l_2 \leq l_2'' < l_2 + l'$. Since by Lemma 3.3 $l_1'' + l_2'' \in L^+_{sup}(F_1+F_2, x_0)$, $l_1 + l_2$ satisfies (a-S1$_R$) for $F_1 + F_2$. Similarly $l_1 + l_2$ satisfies (a-I1$_R$) for $F_1 + F_2$. Therefore by Lemma 3.4 $F_1 + F_2$ is approximately right differentiable and

$$o\text{-}AD^+F_1(x_0) + o\text{-}AD^+F_2(x_0) \subset o\text{-}AD^+(F_1+F_2)(x_0).$$

In the above formula we put $-F_1$ into F_1 and $F_1 + F_2$ into F_2. Then we get

$$o\text{-}AD^+(-F_1)(x_0) + o\text{-}AD^+(F_1+F_2)(x_0) \subset o\text{-}AD^+F_2(x_0).$$

By Theorem 3.3

$$\begin{aligned} o\text{-}AD^+(F_1+F_2)(x_0) &\subset o\text{-}AD^+F_2(x_0) - o\text{-}AD^+(-F_1)(x_0) \\ &= o\text{-}AD^+F_2(x_0) + o\text{-}AD^+F_1(x_0). \end{aligned}$$

Therefore

$$o\text{-}AD^+F_1(x_0) + o\text{-}AD^+F_2(x_0) = o\text{-}AD^+(F_1+F_2)(x_0).$$

The rest can be proved similarly. □

4 In the case of $X = \mathbb{R}^d$ Approximately derivative becomes a subset of bounded linear mappings generally. The problem that it consists of a single point is not solved. However it is true to show the following in the case where X is finite dimensional; see [14].

Lemma 4.1. *Let X and Y be vector lattices and $l \in \mathcal{L}(X,Y)$.*
If $\{x_n\}$ is relatively uniformly convergent to 0 in X, then $\{l(x_n)\}$ is also so in Y.

Proof. Since $\{x_n\}$ is relatively uniformly convergent to 0 in X, there exist $\{\varepsilon_n\} \in \mathcal{U}_{\mathbb{R}}(\mathbb{N})$ and $u \in X$ with $u > 0$ such that $|x_n| \le \varepsilon_n u$ for any natural number n. Then there exists a monotone sequnce $\{r_n\}$ of real numbers such that it is divergent to infinity and $\{r_n x_n\}$ is relatively uniformly convergent to 0. Actually there exists a monotone sequence $\{N(m)\}$ of natural numbers such that $|x_n| \le \frac{1}{m^2}u$ if $n > N(m)$. Let

$$r_n = \begin{cases} 1 & \text{if } n \le N(1), \\ m & \text{if } N(m) < n \le N(m+1) \ (m = 1, 2, \ldots). \end{cases}$$

Since

$$|r_n x_n| = \begin{cases} |x_n| & \text{if } n \le N(1), \\ m|x_n| & \text{if } N(m) < n \le N(m+1) \ (m = 1, 2, \ldots), \end{cases}$$

and $m|x_n| \le \frac{1}{m}u$, $\{r_n x_n\}$ is relatively uniformly convergent to 0 and $\{r_n\}$ is divergent to infinity. Since $\{r_n x_n\}$ is relatively uniformly convergent to 0, it is bounded. Therefore $\{r_n l(x_n)\}$ is also so, that is, there exists $v \in Y$ with $v > 0$ such that $r_n|l(x_n)| \le v$. For m select N such that $r_{N+1} \ge m$. Then $|l(x_n)| \le \frac{1}{r_n}v \le \frac{1}{m}v$ for any natural number $n > N$. It means that $l(x_n)$ is relatively uniformly convergent to 0. □

Lemma 4.2. *Let $X = \mathbb{R}^d$, Y a complete vector lattice, $x_0 \in X$ and $l \in \mathcal{L}(X,Y)$.*
Then

$$o\text{-}\overline{AD}^+ l(x_0) = o\text{-}\underline{AD}^+ l(x_0) = o\text{-}\overline{AD}^- l(x_0) = o\text{-}\underline{AD}^- l(x_0) = \{l\}.$$

Proof. We show that $o\text{-}\overline{AD}^+ l(x_0) = \{l\}$. The rest can be proved similarly. Since $l \in o\text{-}\overline{AD}^+ l(x_0)$ is clear, we show that for any element of $o\text{-}\overline{AD}^+ l(x_0)$ it is equals to l. First we consider a necessary and sufficient condition for $l'' \in L^+_{sup}(l, x_0)$. Note that $\mathcal{K}_X = \{(e_1, \ldots, e_d) \mid e_i > 0 \text{ for any } i\}$ and $\mathcal{L}(X,Y) \cong Y^d$.
In the case of $l'' > l$:
Since

$$E^+_{sup}(l''; l, x_0) = \{x \mid x - x_0 \in \mathcal{K}_X, l(x-x_0) \not< l''(x-x_0)\} = \emptyset,$$

it holds that $l'' \in L^+_{sup}(l, x_0)$.

In the case of $l'' = l$:

Since

$$E^+_{sup}(l''; l, x_0) = \{x \mid x - x_0 \in \mathcal{K}_X\},$$

it holds that for any $h \in \mathcal{K}_X$

$$E^+_{sup}(l''; l, x_0) \cap [x_0, x_0 + h] = \{x \mid x - x_0 \in \mathcal{K}_X\} \cap [x_0, x_0 + h].$$

Therefore x_0 is never a right dispersion point of $E^+_{sup}(l''; l, x_0)$. Then $l'' \notin L^+_{sup}(l, x_0)$.
In the case of $l'' \not\geq l$:

Note that for any $x \in X$ with $x > 0$ there exist $r > 0$ and $0 \leq \theta_i \leq \frac{\pi}{2}$ $(i = 1, \dots, d-1)$ such that

$$\begin{aligned} x &= f(r, \theta_1, \dots, \theta_{d-1}) \\ &= r(\cos\theta_1 \cdots \cos\theta_{d-1}, \cos\theta_1 \cdots \sin\theta_{d-1}, \dots, \sin\theta_1). \end{aligned}$$

Therefore there exists $f(r_0, \theta_{1,0}, \dots, \theta_{d-1,0})$ with $r_0 > 0$, $0 \leq \theta_{i,0} \leq \frac{\pi}{2}$ $(i = 1, \dots, d-1)$ such that

$$l''(f(r_0, \theta_{1,0}, \dots, \theta_{d-1,0})) \not\geq l(f(r_0, \theta_{1,0}, \dots, \theta_{d-1,0})).$$

Then there exists α_i with $0 < \alpha_i + \theta_{i,0} < \frac{\pi}{2}$, $\alpha_i \neq 0$ such that for any θ_i with $|\theta_i - \theta_{i,0}| \leq |\alpha_i|$, $0 \leq \theta_i \leq \frac{\pi}{2}$ it holds that

$$l''(f(r_0, \theta_1, \dots, \theta_{d-1})) \not\geq l(f(r_0, \theta_1, \dots, \theta_{d-1})).$$

If not, then for $\alpha_{i,1}$ with $0 < \alpha_{i,1} + \theta_{i,0} < \frac{\pi}{2}$, $\alpha_{i,1} \neq 0$ there exists $\theta_{i,1}$ with $0 < |\theta_{i,1} - \theta_{i,0}| \leq |\alpha_{i,1}|$, $0 \leq \theta_{i,1} \leq \frac{\pi}{2}$ such that

$$l''(f(r_0, \theta_{1,1}, \dots, \theta_{d-1,1})) \geq l(f(r_0, \theta_{1,1}, \dots, \theta_{d-1,1})).$$

Moreover for $\alpha_{i,2}$ with $0 < \alpha_{i,2} + \theta_{i,0} < \frac{\pi}{2}$, $0 \neq \alpha_{i,2} \leq \frac{1}{2}|\alpha_{i,1}|$ there exists $\theta_{i,2}$ with $0 < |\theta_{i,2} - \theta_{i,0}| \leq |\alpha_{i,2}|$, $0 \leq \theta_{i,2} \leq \frac{\pi}{2}$ such that

$$l''(f(r_0, \theta_{1,2}, \dots, \theta_{d-1,2})) \geq l(f(r_0, \theta_{1,2}, \dots, \theta_{d-1,2})).$$

Repeat this way, then we get a sequence $\{f(r_0, \theta_{1,k}, \dots, \theta_{d-1,k})\}$ such that it is relatively uniformly convergent to $f(r_0, \theta_{1,0}, \dots, \theta_{d-1,0})$ and

$$l''(f(r_0, \theta_{1,k}, \dots, \theta_{d-1,k})) \geq l(f(r_0, \theta_{1,k}, \dots, \theta_{d-1,k})).$$

It is a contradiction to Lemma 4.1. Therefore there exists α_i with $0 < \alpha_i + \theta_{i,0} < \frac{\pi}{2}$, $\alpha_i \neq 0$ such that for any θ_i with $|\theta_i - \theta_{i,0}| \leq |\alpha_i|$, $0 \leq \theta_i \leq \frac{\pi}{2}$ it holds that

$$l''(f(r_0, \theta_1, \dots, \theta_{d-1})) \not\geq l(f(r_0, \theta_1, \dots, \theta_{d-1})).$$

Since l'' and l are linear, the above inequality is true for any $r > 0$. Let

$$W = \left\{ f(r, \theta_1, \dots, \theta_{d-1}) \;\middle|\; r > 0, |\theta_i - \theta_{i,0}| \leq |\alpha_i|, 0 \leq \theta_i \leq \frac{\pi}{2} \ (i = 1, \dots, d-1) \right\}.$$

Then $\{x \mid x - x_0 \in \mathcal{K}_X\} \cap (x_0 + W) \subset E^+_{sup}(l''; l, x_0)$. Let $El(h) = El(h_1, \dots, h_d)$ be the intersection of an ellipsoid, which radii are $h_1, \dots, h_d$, and $\{(x_1, \dots, x_d) \mid x_i > 0 \text{ for any } i\}$. Then $x_0 + El(h) \subset [x_0, x_0 + h]$. Since x_0 is a right density point of $\{x \mid x - x_0 \in \mathcal{K}_X\}$, that is, for any $e \in \mathcal{K}_X$ and for any $\varepsilon \in \mathcal{K}_{\mathbb{R}}$ there exists $e_1 \in \mathcal{K}_X$ such that for any $h \in \mathcal{K}_X$ with $0 < h \leq e_1$ there exists $\{[a_k, b_k] \mid k = 1, 2, \ldots\}$ which satisfies

$$\begin{aligned} \{x \mid x - x_0 \in \mathcal{K}_X\}^C \cap [x_0, x_0 + h] &\subset \bigcup_{k=1}^{\infty} [a_k, b_k]^e, \\ \sum_{k=1}^{\infty} q([a_k, b_k]) &\leq \varepsilon q([x_0, x_0 + h]), \end{aligned}$$

if x_0 is a right dispersion point of $E^+_{sup}(l''; l, x_0)$, that is,

$$\begin{aligned} E^+_{sup}(l''; l, x_0) \cap [x_0, x_0 + h] &\subset \bigcup_{k=1}^{\infty} [c_k, d_k]^e, \\ \sum_{k=1}^{\infty} q([c_k, d_k]) &\leq \varepsilon q([x_0, x_0 + h]), \end{aligned}$$

then

$$\begin{aligned} (x_0 + W) \cap (x_0 + El(h)) &\subset \left(\bigcup_{k=1}^{\infty} [a_k, b_k]^e \right) \cup \left(\bigcup_{k=1}^{\infty} [c_k, d_k]^e \right), \\ \sum_{k=1}^{\infty} q([a_k, b_k]) + \sum_{k=1}^{\infty} q([c_k, d_k]) &\leq 2\varepsilon q([x_0, x_0 + h]) \end{aligned}$$

proving that $(x_0 + W) \cap (x_0 + El(h))$ is a null set. On the other hand

$$\begin{aligned} q((x_0 + W) \cap (x_0 + El(h))) &\geq \frac{|\alpha_1 \cdots \alpha_{d-1}|}{2\pi^{d-1}} \times \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2} + 1)} \times h_1 \cdots h_d \\ &= \frac{|\alpha_1 \cdots \alpha_{d-1}|}{2\pi^{\frac{d}{2}-1} \Gamma(\frac{d}{2} + 1)} h_1 \cdots h_d \\ &= \frac{|\alpha_1 \cdots \alpha_{d-1}|}{2\pi^{\frac{d}{2}-1} \Gamma(\frac{d}{2} + 1)} q([x_0, x_0 + h]), \end{aligned}$$

where Γ is Γ-function. It is a contradiction. Therefore x_0 is never a right dispersion point of $E^+_{sup}(l''; l, x_0)$. Then $l'' \notin L^+_{sup}(l, x_0)$.

Therefore $l'' \in L^+_{sup}(l, x_0)$ if and only if $l'' > l$. Let $l_1 \in o\text{-}\overline{AD}^+ l(x_0)$. For any $l' > 0$ there exists $l'' \in L^+_{sup}(l, x_0)$ such that $l_1 \leq l'' < l_1 + l'$. Since l' is arbitrary, it holds that $l \leq l_1$, moreover by Theorem 3.1 it hold that $l_1 = l$. □

Theorem 4.1. *Let $X = \mathbb{R}^d$, Y a complete vector lattice, $x_0 \in D \in \mathcal{O}_X$ and $l \in \mathcal{L}(X, Y)$.*

(1) *If F is approximately right differentiable at right density point x_0 of $\{x \mid x \in D, x - x_0 \in \mathcal{K}_X\}$, then o-$AD^+F(x_0)$ consists of a single point.*

(2) *If F is approximately left differentiable at left density pont x_0 of $\{x \mid x \in D, x_0 - x \in \mathcal{K}_X\}$, then o-$AD^+F(x_0)$ consists of a single point.*

Proof. In Theorem 3.4 (1) put $F_1 = F$ and $F_2 = -F$ and by Lemma 4.2

$$o\text{-}AD^+F(x_0) - o\text{-}AD^+F(x_0) = o\text{-}AD^+0(x_0) = \{0\}.$$

Therefore o-$AD^+F(x_0)$ consists of a single point. Similarly it can be proved that o-$AD^-F(x_0)$ consists of a single point. □

5 Relation We consider a relation between the approximately derivative and the derivative. However it is not known any desirable relation. In this section we consider the case where $X = \mathbb{R}$ and Y is totally ordered.

Theorem 5.1. *Let $X = \mathbb{R}$, Y a complete vector lattice with total ordering, $x_0 \in D \in \mathcal{O}_X$ and F a mapping from D into Y.*

(1) *If F is right differentiable at x_0, then it is approximately right differentiable and $o\text{-}D^+F(x_0) = o\text{-}AD^+F(x_0)$.*

(2) *If F is left differentiable at x_0, then it is approximately right differentiable and $o\text{-}D^-F(x_0) = o\text{-}AD^-F(x_0)$.*

Proof. Let $l = o\text{-}D^+F(x_0)$. Then there exists $\{w_{x_0,e}\} \in \mathcal{U}^s_{\mathcal{L}(X,Y)}(\mathcal{K}_X, \geq)$ such that for any $e \in \mathcal{K}_X$ there exists $\delta_{x_0} \in \mathcal{K}_\mathbb{R}$ such that $|F(x_0+h) - F(x_0) - l(h)| \leq w_{x_0,e}(h)$ for any $h \in X$ with $0 < h \leq \delta_{x_0}e$. Let $l' \in \mathcal{L}(X, Y)$ with $l' > 0$. Since $\mathcal{U}^s_{\mathcal{L}(X,Y)}(\mathcal{K}_X, \geq)$ is totally ordered, there exists $e \in \mathcal{K}_X$ such that $w_{x_0,e}(h) < \frac{1}{2}l'(h)$ for any $h \in \mathcal{K}_X$ with $0 < h \leq \delta_{x_0}e$. Let $l'' = l + \frac{1}{2}l'$. Then $l \leq l'' < l + l'$ and $l'' \in L^+_{sup}(F, x_0)$. Actually since for any $h \in \mathcal{K}_X$ with $0 < h \leq \delta_{x_0}e$, $x_0 + h \in D$

$$F(x_0 + h) - F(x_0) \leq (l + w_{x_0,e})(h) < l''(h),$$

it holds that $E^+_{sup}(l''; F, x_0) \cap [x_0, x_0 + h] = \emptyset$. Therefore x_0 is a right dispersion point of $E^+_{sup}(l''; F, x_0)$. Then $l'' \in L^+_{sup}(F, x_0)$. Therefore l satisfies (a-S1$_R$). Similarly it can be proved that l satisfies (a-I1$_R$). By Lemma 3.4 F is approximately right differentiable at x_0. By Theorem 4.1 we obtain that $o\text{-}D^+F(x_0) = o\text{-}AD^+F(x_0)$. The rest can be proved similarly. □

References

[1] A. Alexiewicz, *On Denjoy integrals of abstract functions*, Towarzysewo Naukowe Warszawskie (Soc. Sci. Lett. Varsovie C. R. Cl. III. Sci. Math. Phys.) **41** (1948), 97–129.

[2] G. Birkhoff, *Lattice Theory*, American Mathematical Society, Providence, 1940.

[3] A. Boccuto, *Differential and integral calculus in Riesz spaces*, Tatra Mountains Mathematical Publications **14** (1998), 293–323.

[4] A. Boccuto, B. Riečan, and M. Vrábelová, *Kurzweil-Henstock Integral in Riesz Spaces*, Bentham Science Publishers, Sharjah, 2009.

[5] R. Cristescu, *Topological Vector Spaces*, Noordhoff International Publishing, Leyden, 1977.

[6] S. Izumi, *An abstract integral (X)*, Proceedings of the Imperial Academy of Japan **18** (1942), no. 9, 543–547.

[7] S. Izumi, G. Sunouchi, M. Orihara, and M. Kasahara, *Denjoy integrals, I*, Proceedings of the Physico-Mathematical Society of Japan **17** (1943), 102–120 (in Japanese).

[8] ———, *Denjoy integrals, II*, Proceedings of the Physico-Mathematical Society of Japan **17** (1943), 329–353 (in Japanese).

[9] T. Kawasaki, *Order derivative of operators in vector lattices*, Mathematica Japonica **46** (1997), no. 1, 79–84.

[10] ———, *On Newton integration in vector spaces*, Mathematica Japonica **46** (1997), no. 1, 85–90.

[11] ———, *Order Lebesgue integration in vector lattices*, Mathematica Japonica **48** (1998), no. 1, 13–17.

[12] ———, *Approximately order derivatives in vector lattices*, Mathematica Japonica **49** (1999), no. 2, 229–239.

[13] ———, *Order derivative and order Newton integral of operators in vector lattices*, Far East Journal of Mathematical Sciences **1** (1999), no. 6, 903–926.

[14] ———, *Uniquely determinedness of the approximately order derivative*, Scientiae Mathematicae Japonicae **57** (2003), no. 2, 373–376.

[15] ———, *Denjoy integral and Henstock-Kurzweil integral in vector lattices, I*, Czechoslovak Mathematical Journal **59** (2009), no. 2, 381–399.

[16] ———, *Denjoy integral and Henstock-Kurzweil integral in vector lattices, II*, Czechoslovak Mathematical Journal **59** (2009), no. 2, 401–417.

[17] ———, *Convergence theorems for the Henstock-Kurzweil integral taking values in a vector lattice*, Scientiae Mathematicae Japonicae **75** (2012), no. 2, 223–233.

[18] Y. Kubota, *Theory of the Integral*, Maki, Tokyo, 1977 (in Japanese).

[19] P.-Y. Lee, *Lanzhou Lectures on Henstock Integration*, World Scientific, Singapore, 1989.

[20] W. A. J. Luxemburg and A. C. Zaanen, *Riesz Spaces*, North-Holland, Amsterdam, 1971.

[21] P. McGill, *Integration in vector lattices*, Journal of the London Mathematical Societyy. Second Series **11** (1975), 347–360.

[22] B. Riečan and T. Neubrunn, *Integral, Measure, and Ordering*, Kluwer Academic Publishers, Dordrecht, 1997.

[23] P. Romanovski, *Intégrale de Denjoy dans les espaces abstraits*, Recueil Mathématique (Matematicheskii Sbornik) N. S. **9** (1941), no. 1, 67–120.

[24] H. H. Schaefer, *Banach Lattices and Positive Operators*, Springer-Verlag, Berlin / Heidelberg / New York, 1974.

[25] B. Z. Vulikh, *Introduction to the Theory of Partially Orderd Spaces*, Wolters-Noordhoff, Groningen, 1967.

Communicated by *Jun Kawabe*

T. Kawasaki
College of Engineering, Nihon University, Fukushima 963–8642, Japan
E-mail: toshiharu.kawasaki@nifty.ne.jp

Scientiae Mathematicae Japonicae, **78**, No. 1 (2015), 81–86

REPRODUCING PROPERTY FOR INTERPOLATIONAL PATH OF OPERATOR MEANS

JUN ICHI FUJII

Received February 24, 2014 ; revised March 4, 2014

ABSTRACT. We show that the solution of the 2-variable Karcher equation for the derivative solidarity coincides with the original interporational path of operator means, where the derivative solidarity for an interpolational path of operator means $A\,\mathrm{m}_t B$ is defined as $A\,\mathbf{s}_{\mathrm{m}}B = \frac{\partial A\,\mathrm{m}_t B}{\partial t}\Big|_{t=0}$.

Let m be *an operator mean* in the sense of Kubo-Ando [7] which is defined by a positive operator monotone function f_{m} on the half interval $(0,\infty)$ with $f_{\mathrm{m}}(1) = 1$;

$$A\,\mathrm{m}\,B = A^{\frac{1}{2}} f_{\mathrm{m}}\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) A^{\frac{1}{2}}$$

for positive invertible operators A and B on a Hilbert space. Thus the operator mean can be constructed by a numerical function $f_{\mathrm{m}}(x) = 1\,\mathrm{m}\,x$ which is called the representing function of m. For a symmetric operator mean m, i.e., $A\,\mathrm{m}\,B = B\,\mathrm{m}\,A$, the initial conditions

$$A\,\mathrm{m}_0 B = A, \quad A\,\mathrm{m}_{\frac{1}{2}} B = A\,\mathrm{m}\,B, \quad A\,\mathrm{m}_1 B = B$$

and the following inductive relation

$$A\,\mathrm{m}_{\frac{2k+1}{2^{n+1}}} B = (A\,\mathrm{m}_{\frac{k}{2^n}} B)\,\mathrm{m}\,(A\,\mathrm{m}_{\frac{k+1}{2^n}} B) = (A\,\mathrm{m}_{\frac{k+1}{2^n}} B)\,\mathrm{m}\,(A\,\mathrm{m}_{\frac{k}{2^n}} B) \tag{2}$$

for nonnegative numbers n and k with $2k+1 < 2^{n+1}$ determine the continuous path $A\,\mathrm{m}_t B$ from A to B of operator means. In particular,

$$A\,\mathrm{m}_{\frac{1}{2^n}} B = A\,\mathrm{m}\,(A\,\mathrm{m}_{\frac{1}{2^{n-1}}} B) = A(A\,\mathrm{m}\,(A\,\mathrm{m}_{\frac{1}{2^{n-2}}} B)) = \cdots = \overbrace{A(A(\cdots(A}^{n \text{ times}}\,\mathrm{m}\,B\overbrace{)\cdots))}^{n \text{ times}}. \tag{3}$$

Then, if the limit

$$A\,\mathbf{s}_{\mathrm{m}} B = \lim_{n\to\infty} 2^n (A\,\mathrm{m}_{\frac{1}{2^n}} B - A)$$

exists, it defines the *solidarity* whose representing function $F_{\mathbf{s}}(x) = 1\,\mathbf{s}\,x$ is a strictly increasing operator monotone function. The solidarity $\mathbf{s}$ in ([4]) is defined as a binary operation $A\,\mathbf{s}\,B$ for positive (invertible) operators A and B by

$$A\,\mathbf{s}\,B = A^{\frac{1}{2}} F\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) A^{\frac{1}{2}}$$

for some operator monotone function F on $(0,\infty)$. It has also typical properties of operator means except the monotonicity on the left-term. In particular, note that the *transformer equality*

$$T(A\,\mathbf{s}\,B)T^* = (TAT^*)\,\mathbf{s}\,(TBT^*)$$

2000 *Mathematics Subject Classification.* 47A64, 47A63, 26E60.

Key words and phrases. operator mean, operator monotone function, path, geodesic, quasi-arithmetic mean.

holds for invertible operators T.

If this path m_t is differentiable for t, then

$$A\,\mathbf{s}_{\mathrm{m}}B = \lim_{t\to 0}\frac{A\,\mathrm{m}_t B - A}{t} = \frac{\partial A\,\mathrm{m}_t B}{\partial t}\Big|_{t=0}.$$

So it is called the *derivative solidarity* for m. Its representing function $F_{\mathbf{s}}$ satisfies $F_{\mathbf{s}}(1) = 0$ and $F'_{\mathbf{s}}(1) = 1$ ([6]).

If a path satisfies

$$(A\,\mathrm{m}_r B)\,\mathrm{m}_t(A\,\mathrm{m}_s B) = A\,\mathrm{m}_{(1-t)r+ts}\,B$$

for all weights $r,\ s,\ t \in [0,1]$, then we call it an *interporational path* and also call the original mean an *interpolational* one as in [5, 6]. In the preceding paper [3], we showed that m_t is interpolational if and only if it satisfies the *mixing property*:

$$(a\,\mathrm{m}\,b)\,\mathrm{m}\,(c\,\mathrm{m}\,d) = (a\,\mathrm{m}\,c)\,\mathrm{m}\,(b\,\mathrm{m}\,d)$$

for all positive numbers a, b, c and d. This shows that the *logarithmic operator mean*

$$A\,\mathbf{L}\,B = A^{\frac{1}{2}}\ell\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)A^{\frac{1}{2}}$$

for the function $\ell(x) = (x-1)/\log x$ is not interpolational. We also showed in [6] that every interpolational path is convex by the maximality of the arithmetic mean in the symmetric operator means $\mathrm{m} = \mathrm{m}_{1/2}$;

$$A\,\mathrm{m}_{\frac{t+s}{2}}B = (A\,\mathrm{m}_t B)\,\mathrm{m}\,(A\,\mathrm{m}_s B) \le \frac{A\,\mathrm{m}_t B + A\,\mathrm{m}_s B}{2}.$$

Moreover it is differentiable and hence has always the derivative solidarity. This construction is similar to Uhlmann's one [9] that defines the relative entropy from interpolations.

For $r \in [-1,1]$, the following parametrized operator means $\#_t^{(r)}$, which are also called the *quasi-arithmetic* ones (cf. [2]),

$$A\#_t^{(r)}B = A^{\frac{1}{2}}\left((1-t)I + t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right)^{\frac{1}{r}}A^{\frac{1}{2}}$$

are interpolational. The path $\#_t = \#_t^{(0)} = \lim_{\varepsilon\downarrow 0}\#_t^{(\varepsilon)}$;

$$A\#_t B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}}$$

is that of the geometric operator mean and it is also the geodesic of the Finsler manifold of the positive invertible operators by Corach-Porta-Recht [1].

In [5, 6], we considered a map $\mathrm{m} \mapsto \mathbf{s}_{\mathrm{m}}$ from the interporational means m to the solidarities, say *Uhlmann's transform* by the above reason, but we could not discuss the inverse map then. In this paper, we will show that the solution X of the (2-variable) Karcher equation

$$(1-t)\,(X\,\mathbf{s}_{\mathrm{m}}A) + t\,(X\,\mathbf{s}_{\mathrm{m}}B) = 0 \tag{4}$$

is the original path $A\,\mathrm{m}_t B$ as M.Pálfia suggested as we see later. This Karcher equation is equivalent to

$$(1-t)F\left(X^{-\frac{1}{2}}AX^{-\frac{1}{2}}\right) + tF\left(X^{-\frac{1}{2}}BX^{-\frac{1}{2}}\right) = 0 \tag{4'}$$

for the representing function $F_{\mathbf{s}}(x) = 1\,\mathbf{s}_{\mathrm{m}}x$.

To see this, we make some preparations. By the interpolationality of m_t, its representing function f_t has the following property:

Lemma 1. $f_s(f_t(x)) = f_{st}(x)$.

Proof. By the interpolationality, we have

$$f_s(f_t(x)) = 1 \,\mathrm{m}_s (1 \,\mathrm{m}_t \, x) = (1 \,\mathrm{m}_0 \, x) \,\mathrm{m}_s (1 \,\mathrm{m}_t \, x) = 1 \,\mathrm{m}_{0(1-s)+ts} \, x = f_{st}(x). \qquad \square$$

Since m is symmetric and m_t is homogeneous, we have:

Lemma 2. $B \,\mathrm{m}_{1-t} A = A \,\mathrm{m}_t B$ *and* $x f_{1-t}\left(\frac{1}{x}\right) = f_t(x)$.

Proof. The former follows from the construction (2). So we have

$$x f_{1-t}\left(\frac{1}{x}\right) = x\left(1 \,\mathrm{m}_{1-t}\left(\frac{1}{x}\right)\right) = x \,\mathrm{m}_{1-t}\, 1 = 1 \,\mathrm{m}_t \, x = f_t(x). \qquad \square$$

Consider the derivative function $F_t(x) = \frac{\partial f_t(x)}{\partial t}$ (where $F_{\mathbf{s}} = F_0$). In [6], we showed $F_0(f_t(x)) = tF_t(x)$. Moreover we have:

Lemma 3. $F_s(f_t(x)) = tF_{ts}(x)$ *and* $F_{1-t}\left(\frac{1}{x}\right) = -\frac{1}{x}F_t(x)$.

Proof. By the definition of F_s and Lemma 1, we obtain

$$F_s(f_t(x)) = \lim_{r\to s} \frac{f_r(f_t(x)) - f_s(f_t(x))}{r-s} = t \lim_{r\to s} \frac{f_{tr}(x) - f_{ts}(x))}{tr - ts} = tF_{ts}(x).$$

Also the formula

$$-\frac{1}{x}F_t(x) = \lim_{s\to t} \frac{f_s(x)/x - f_t(x)/x}{-(s-t)} = \lim_{s\to t} \frac{f_{1-s}(1/x) - f_{1-t}(1/x)}{(1-s)-(1-t)} = F_{1-t}\left(\frac{1}{x}\right)$$

follows from the property $f_{1-t}(1/x) = \frac{f_t(x)}{x}$ in Lemma 2. $\square$

Let $\mathbf{s}_t$ be the solidarity defined by the derivative F_t. Then the above Lemma shows the formulae for tangent vectors and the transpose relation:

Theorem 4. *For the above solidarity* $\mathbf{s}_t$ *for an interpolational path* m_t,

$$A \,\mathbf{s}_s (A \,\mathrm{m}_t B) = tA \,\mathbf{s}_{ts}\, B \quad \textit{and} \quad -A \,\mathbf{s}_t B = B \,\mathbf{s}_{1-t}\, A$$

for parameters $s, t \in [0,1]$.

Proof. Lemma 2 shows the first formula by

$$\begin{aligned} A \,\mathbf{s}_s (A \,\mathrm{m}_t B) &= A^{\frac{1}{2}} F_s \left(f_t \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)\right) A^{\frac{1}{2}} \\ &= tA^{\frac{1}{2}} F_{st} \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) A^{\frac{1}{2}} = tA \,\mathbf{s}_{ts}\, B. \end{aligned}$$

Also, since the transformer equality and Lemma 2 imply

$$X \,\mathbf{s}_t I = X(I \,\mathbf{s}_t X^{-1}) = XF_t(X^{-1}) = -F_{1-t}(X),$$

we have

$$-A \,\mathbf{s}_t B = B^{\frac{1}{2}} \left[\left(B^{-\frac{1}{2}} A B^{-\frac{1}{2}}\right) \mathbf{s}_t I\right] B^{\frac{1}{2}} = -B^{\frac{1}{2}} F_{1-t} \left(B^{-\frac{1}{2}} A B^{-\frac{1}{2}}\right) B^{\frac{1}{2}} = B \,\mathbf{s}_{1-t}\, A. \qquad \square$$

Now we show the reproducing property:

Theorem 5. *Let F be the representing function for a derivative solidarity $\mathbf{s}_{\mathrm{m}}$ for an interpolational mean* m*, the solution of the Karcher equation (4) is the interpolational path $A\,\mathrm{m}_t B$.*

Proof. The homogeneity for operator means shows that we have only to show that the solution y of the numerical equation

$$(1-t)F\left(\frac{1}{y}\right)+tF\left(\frac{x}{y}\right)=0 \tag{4''}$$

is given by $y=f_t(x)=1\,\mathrm{m}_t x$. In fact, by the above lemma, we have

$$\begin{aligned}
&(1-t)F_0\left(\frac{1}{f_t(x)}\right)+tF_0\left(\frac{x}{f_t(x)}\right)\\
&\qquad=-\frac{1-t}{f_t(x)}F_1(f_t(x))-\frac{t}{f_{1-t}(1/x)}F_1(f_{1-t}(1/x))\\
&\qquad=-\frac{(1-t)t}{f_t(x)}F_t(x)-\frac{t(1-t)}{f_{1-t}(1/x)}F_{1-t}(1/x)\\
&\qquad=-\frac{t(1-t)}{f_t(x)}F_t(x)+\frac{t(1-t)}{(1/x)f_t(x)}\frac{1}{x}F_t(x)=0.
\end{aligned}$$

Thus we obtain $y=f_t(x)$. □

In a RIMS Workshop held at November 6–8, 2013 in Kyoto, M.Pálfia [8] posed an interesting problem when the solution of the Karcher equation (4) is a path of operator means for an operator monotone function F with $F(1)=0$ and $F'(1)=1$. So we observe the above theorem from this constructive viewpoint: In (4″), it follows from the monotonicity of F that

$$\frac{x}{y}=F^{-1}\left(-\frac{1-t}{t}F\left(\frac{1}{y}\right)\right).$$

So, putting

$$g_t(y)=yF^{-1}\left(-\frac{1-t}{t}F\left(\frac{1}{y}\right)\right),$$

we have $x=g_t(y)$. Thus if we find an interpolational path f_t with $y=f_t(g_t(y))=f_t(x)$, the solution X coincides with $A\,\mathrm{m}_t B$ for the corresponding path $1\,\mathrm{m}_t x=f_t(x)$. But we notice that g_t is not monotone:

Remark. Let m_t be an interpolational path defined by a function

$$f_t(x)=\left(1-t+t\sqrt{x}\right)^2.$$

Then it must be a solution of (4′). The derivative solidarity is determined by $F(x)=2(\sqrt{x}-1)$. Since

$$F^{-1}(z)=\left(\frac{z}{2}+1\right)^2,$$

the above function is

$$g_t(y)=y\left(1-\frac{1-t}{t}\left(\sqrt{\frac{1}{y}}-1\right)\right)^2=\frac{\left(\sqrt{y}-(1-t)\right)^2}{t^2},$$

which is convex with the minimum 0 at $y=(1-t)^2$. It is not monotone on $(0,\infty)$, but in the region $y>(1-t)^2$, the function $g_t(y)$ is monotone and its inverse function is $(1-t+t\sqrt{x})^2$, which is operator monotone.

Finally we give an example of operator monotone function F satisfying $F(1) = 0$ and $F'(1) = 1$ that does not induce an operator mean:

Example. Let $F(x) = 2(\sqrt{2(x+1)} - 2)$. Then we have $F(1) = 0$, $F'(1) = 0$ and

$$F^{-1}(z) = \frac{(z+4)^2}{8} - 1.$$

It follows that

$$g_t(y) = \frac{\left(2\sqrt{y} - \frac{1-t}{t}\left(\sqrt{2(1+y)} - 2\sqrt{y}\right)\right)^2}{2} - y$$

and hence

$$t^2 g_t(y) = (3-2t)y + (1-t)^2 - 2(1-t)\sqrt{2(1+y)y}.$$

Thus we have

$$g_{\frac{1}{2}}(y) = 8y + 1 - 4\sqrt{2(1+y)y}.$$

This function is convex and $g_{\frac{1}{2}}(y) = 0$ for $y = \frac{2\pm\sqrt{2}}{8}$. But, even in the region $y > \frac{2+\sqrt{2}}{8}$ where g_t is monotone and positive, the inverse function

$$f_{\frac{1}{2}}(x) = \frac{2(x+1) + \sqrt{2x^2 + 12x + 2}}{8}$$

is not operator monotone. In fact, a concave function $h(x) = \sqrt{x^2 + 6x + 1}$ is not operator monotone (and hence not operator concave) since $\operatorname{Arg} z < \operatorname{Arg} h(z)$ for some z in the upper half complex plane: Consider $z = re^{it}$ for $t = \frac{3\pi}{4}$. Then

$$h(z)^2 = 1 + 6r\left(-\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) - ir^2,$$

$\mathrm{Re} z^2 = 0$ and $\mathrm{Re}\, h(z)^2 = 1 - 3\sqrt{2}r$. It follows that $\operatorname{Arg} f(z) > \frac{3\pi}{4}$ for $r < \frac{1}{3\sqrt{2}}$, which shows h is not operator monotone.

Acknowledgement. This study is partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 23540200, 2011.

References

[1] G.Corach, H.Porta and L.Recht, *Geodesics and operator means in the space of positive operators*, Internat. J. Math., **4** (1993), 193–202.

[2] J.I.Fujii, Structure of Hiai-Petz parametrized geometry for positive definite matrices, Linear Algebra Appl. **432** (2010), 318–326.

[3] J.I.Fujii, *Interpolationality for symmetric operator means*, Sci. Math. Japon., **75** (2012), 345–352.

[4] J.I.Fujii, M.Fujii and Y.Seo, *An extension of the Kubo-Ando theory: Solidarities*, Math. Japon., **35** (1990), 387–396.

[5] J.I.Fujii and E.Kamei, *Uhlmann's interpolational method for operator means*, Math. Japon., **34** (1989), 541–547.

[6] J.I.Fujii and E.Kamei, *Interpolational paths and their derivatives*, Math. Japon., **39** (1994), 557–560.

[7] F.Kubo and T.Ando, *Means of positive linear operators*, Math. Ann., **248** (1980) 205–224.

[8] M.Pálfia, *Means of positive definite operators as minimizers of convex functions*, to appear in Sūrikaisekikenkyūsho Kōkyūroku.

[9] A.Uhlmann, *Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory*, Commun. Math. Phys., **54**(1977), 22–32.

Communicated by *Masatoshi Fujii*

Department of Art and Sciences (Information Science), Osaka Kyoiku University, Asahigaoka, Kashiwara, Osaka 582-8582, Japan.

E-mail address, fujii@cc.osaka-kyoiku.ac.jp

OPTIMAL SERVICE HOURS WITH SPECIAL OFFERS

Hyewon Kim AND Hiroaki Sandoh

Received January 7, 2014; revised April 4, 2014

Abstract. In managing service provider businesses, it is important to decentralize consumers at peak time and increase sales not at peak time as well. Shy and Stanbacka[5] have dealt with this problem to explore optimal service hours under a specific ideal time distribution, and discussed the existence of optimal opening and closing times. In the actual environments, however, service providers strategically introduce a wide variety of special offers such as discounted price to collect more consumers.

In this study, we deal with optimal service hours with a special offer of price discount immediately after the opening time and just before the closing time with the view to attracting extra consumers whose ideal and convenient service times are before the opening time and after the closing time. Under the ideal service time distribution by Shy and Stanbacka[5], the provider's profit is first formulated as an objective function to be maximized and then clarified is the condition under which the service provider can earn more profit by special offers than without special offers. An optimal business hours is also explored to clarify the conditions where there exist optimal opening and closing times. Numerical examples are also presented to illustrate the proposed model formulation.

1 Introduction

Business hours have been traditionally regulated particularly in many European countries although the liberalization of business hours generated debates in these three decades(see, e.g., De Meza[4], Ferris[2], Clemenz [1], Inderst and Irmen[3]).

On the other hand, service providers as well as retailers are eager to make more profits by strategic managements. It is especially important for service industries to decentralize consumers in peak time and increase sales not in peak time. Shy and Stanbacka[5] have dealt with this problem to explore optimal service hours. Under a specific ideal service time distribution of consumers, they discussed the existence of optimal service hours.

A special offer such as early birds specials and/or closing time discount/sale is one of the effective strategies for service industries as well as retailers since they can possibly increase the sales not in peak time. We can observe, in the real circumstances, special offers in a wide variety of service and retailing industries, e.g., morning perm at a beauty salon, happy hour at a hotel, midnight discount of a telecommunications industry, special time discount in business logistics and so forth.

This study confines itself to a service provider having a special offer of price discount immediately after the opening time and just before the closing time. This type of special offer is effective since they can attract extra consumers whose ideal or convenient service times are before the opening time or after the closing time.

First, we formulate the provider's profit as an objective function under the ideal service time distribution introduced by Shy and Stanbacka[5]. Clarified are the conditions under which the service provider can increase his profit by the special offer.

2010 *Mathematics Subject Classification.* 90B05, 90B50 .
Key words and phrases. expected profit, special offer, service hours.

Second, we explore optimal business hours maximizing the provider's profit to show the conditions where there exist unique optimal opening and closing times. Numerical examples are also presented to illustrate the characteristics of the proposed model.

2 Model Formulation

2.1 Assumptions and notations The assumptions and the relevant notations of this study are as follows;

(a) Each individual consumer has her own ideal service time to visit the provider or to receive his service.

(b) The utility of each consumer is given by U_t when she purchase a service product at time t.

(c) Each consumer obtains utility, u_0, by purchasing a service product.

(d) The regular selling price of service is p.

(e) The provider sells his service product at price αp as his special offer, where $0 < \alpha < 1$.

(f) The time during which a special offer is provided is denoted by $\tau(> 0)$.

(g) A consumer owes ω per unit time in visiting the provider or receiving his service earlier or later than her ideal time.

(h) The opening and closing times are, respectively, t_o and t_c, where we have $0 \leq t_o \leq t_c \leq 1$.

(i) The raw price per service product is given by c_1, while the operation cost of the service provider per unit of time is c_2.

2.2 Ideal time distribution In this study, we assume that a demand quantity, q_t, at ideal time t is given by

$$q_t = \begin{cases} n[\mu + 4(1-\mu)t], & 0 \leq t < \frac{1}{2} \\ n[4 - 3\mu - 4(1-\mu)t], & \frac{1}{2} \leq t \leq 1 \end{cases}, \tag{21}$$

where n repsresents the population size and μ ($0 \leq \mu \leq 1$) measures the degree of uniformity. Figure 1 shows the ideal time distribution given by Eq. (21) for $n = 1$ against various values of μ.

Shy and Stanbacka[5] have assumed the above ideal time distribution on the unit circle with the view to formalizing the idea that there are spillovers between time periods. In this study, however, we assume the same structure of the ideal time distribution on the unit time interval $[0, 1]$. This is because spillovers are an important factor only when the service provider sells his products for almost whole unit time period, and in such a situation the strategic determination of service hours might not be necessary.

We here introduce an additional assumption as follows:

(j) When the selling price is reduced to αp at t, demand quantity q_t increases to $[1+\beta(\alpha)]q_t$ with $\beta(\alpha) > 0$ for $0 < \alpha < 1$.

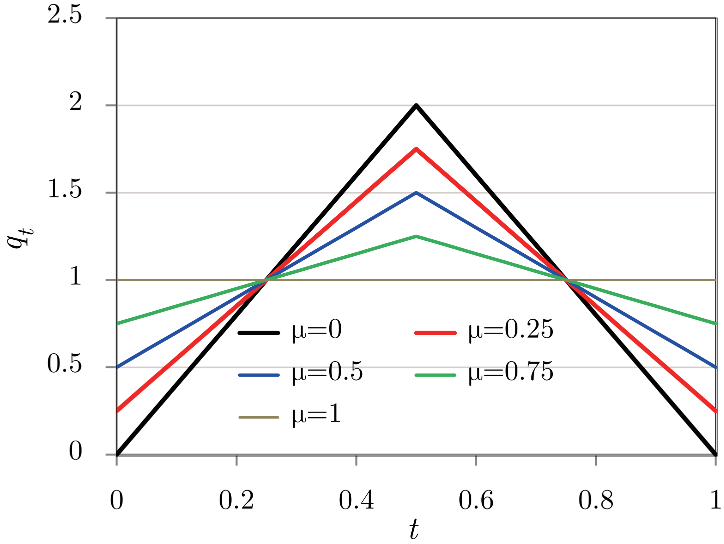

Figure 1: Ideal time distribution($n = 1$).

Assumption (j) signifies the price elasticity η of demand is given by

$$\eta = -\frac{\frac{\beta(\alpha)q_t}{q_t}}{\frac{(\alpha-1)p}{p}} = \frac{\beta(\alpha)}{1-\alpha},\ 0 < \alpha < 1,$$

where $\lim_{\alpha \to 1-0} \beta(\alpha) = 0$.

In the following, consumers represented by the demand quantity q_t are called type $\mathscr{A}$, while those expressed by $\beta(\alpha)q_t$ are referred to type $\mathscr{B}$. Moreover, we concentrate upon the case where values of α and $\beta(\alpha)$ are both specified to specific values, and $\beta(\alpha)$ is written as β for simplicity.

3 Consumers' Behavior

3.1 Best response Since the ideal time distribution by Eq. (21) reveals a symmetrical shape, the opening time, t_o, and the closing time, t_c, are also symmetrical with respect to $t = \frac{1}{2}$, accordingly we have

$$t_c = 1 - t_o.$$

Hence, we focus on the former half of period $[0, \frac{1}{2}]$ to discuss the opening time, t_o, hereafter.
(1) Type $\mathscr{A}$ consumers' response When the provider introduces early birds specials and/or closing time discount/sale, the best response of type $\mathscr{A}$ consumers with ideal time t becomes as follows:

i) If $t \in [0, t_o^{(1a)}]$, type $\mathscr{A}$ consumers are reluctant to wait until t_o, and purchase no service product, where

$$t_o^{(1a)} = t_o - \frac{u_0 - \alpha p}{\omega}. \tag{31}$$

Consequently, their net utility is given by

$$U_t = 0.$$

ii) If $t \in (t_o^{(1a)}, t_o]$, type $\mathscr{A}$ consumers purchase a service product at the discounted price, αp, by waiting until t_o. In this case, their net utility becomes

$$U_t = u_0 - \alpha p - \omega(t_o - t).$$

iii) Type $\mathscr{A}$ consumers with $t \in (t_o, t_o + \tau]$ purchase a service product at their ideal service time at αp, and hence we have

$$U_t = u_0 - \alpha p.$$

iv) In the case of $t \in (t_o + \tau, t_o^{(2)}]$, the consumers purchase a service product earlier than their ideal time at αp, and thereby

$$U_t = u_0 - \alpha p - \omega[t - (t_o + \tau)],$$

where

$$t_o^{(2)} = t_o + \tau + \frac{(1-\alpha)p}{\omega}. \tag{32}$$

It should be noted in Eq. (32) that $t_o^{(2)} \neq t_o + \tau + \frac{u_0 - \alpha p}{\omega}$ since consumers with ideal time t can obtain positive utility $u_0 - p$ even at t, and therefore $t_o^{(2)}$ should be derived from the condition in reference to t; $u_0 - \alpha p - \omega[t - (t_o + \tau)] \geq u_0 - p$.

v) When $t \in (t_o^{(2)}, \frac{1}{2}]$, type $\mathscr{A}$ consumers will purchase a service product at p, at their ideal time t, and hence

$$U_t = u_0 - p.$$

(2) Type $\mathscr{B}$ consumers' response The best response of type $\mathscr{B}$ consumers with ideal time t is described as follows:

i) If $t \in [0, t_o^{(1b)}]$, type $\mathscr{B}$ consumers would not wait until t_o and purchase no service product at αp, where

$$t_o^{(1b)} = t_o - \frac{(1-\alpha)p}{\omega}. \tag{33}$$

Consequently, their net utility becomes

$$U_t = 0.$$

ii) If $t \in (t_o^{(1b)}, t_o]$, type $\mathscr{B}$ consumers purchases a service product at αp, by waiting until t_o. In this case, the maximum value of their net utility can be represented by

$$U_t = (1-\alpha)p - \omega(t_o - t).$$

iii) Consumers with $t \in (t_o, t_o + \tau]$ purchase a service product at discounted price, αp, at their ideal service time, and hence their maximum net utility can be expressed as

$$U_t = (1-\alpha)p.$$

iv) In the case of $t \in (t_o + \tau, t_o^{(2)}]$, type $\mathscr{B}$ consumers purchase a product earlier than their ideal time at αp, and their maximum net utility becomes

$$U_t = (1-\alpha)p - \omega[t - (t_o + \tau)].$$

v) When $t \in (t_o^{(2)}, \frac{1}{2}]$, type $\mathscr{B}$ consumers would purchase no service product yielding

$$U_t = 0.$$

3.2 Domain of opening time It is neither reasonable nor proper for a consumer with ideal time $t < 0$ to wait until t_o, and we assume

$$\min\left(t_o^{(1a)}, t_o^{(1b)}\right) = t_o^{(1a)} = t_o - \frac{u_0 - \alpha p}{\omega} \geq 0,$$

which constrains the opening time to satisfy

$$t_o \geq \frac{u_0 - \alpha p}{\omega}. \tag{34}$$

The right-hand-side of this equation is denoted by t_L in the following.

Likewise, it is reasonable to assume

$$t_o^{(2)} \leq \frac{1}{2},$$

which is equivalent to

$$t_o \leq \frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega}. \tag{35}$$

The right-hand-side of the above equation is denoted by t_U.

It should be noted here that Eqs.(34) and (35) yield,

$$\frac{u_0 - \alpha p}{\omega} \leq \frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega},$$

which singifies, at the same time, that ω should satisfy

$$\omega \geq \frac{2[u_0 + (1-2\alpha)p]}{1 - 2\tau}. \tag{36}$$

From Eqs. (34) and (35), the domain of t_o is, as a result, given by

$$t_L \equiv \frac{u_0 - \alpha p}{\omega} \leq t_o \leq \frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega} \equiv t_U. \tag{37}$$

4 Provider's Profit

Let $Q_{1A}(t_o)$ express the number of type $\mathscr{A}$ consumers who purchase a service product at αp, then

$$\begin{aligned} Q_{1A}(t_o) &= \int_{t_o^{(1a)}}^{t_o^{(2)}} q_t dt \\ &= 2n\left[\tau + \frac{u_0 + (1-2\alpha)p}{\omega}\right]\left[\mu + 2(1-\mu)\left(2t_o + \tau - \frac{u_0 - p}{\omega}\right)\right]. \end{aligned} \tag{41}$$

By letting $Q_{1B}(t_o)$ signify the number of type $\mathscr{B}$ consumers who purchase a service product at αp, we have

$$\begin{aligned} Q_{1B}(t_o) &= 2\int_{to^{(1b)}}^{t_o^{(2)}} \beta q_t dt \\ &= 2n\beta\mu\left[\tau + \frac{2(1-\alpha)p}{\omega}\right] + 4n\beta(1-\mu)\left[\tau + \frac{2(1-\alpha)p}{\omega}\right](2t_o + \tau). \end{aligned} \tag{42}$$

On the other hand, let us denote, by $Q_2(t_o)$, the number of consumers who purchase a service product at the regular price, p, then we have

$$
\begin{aligned}
Q_2(t_o) &= 2\int_{t_o^{(2)}}^{\frac{1}{2}} q_t dt \\
&= 2n\left[\frac{1}{2} - t_o - \tau - \frac{(1-\alpha)p}{\omega}\right]\left\{\mu + 2(1-\mu)\left[\frac{1}{2} + t_o + \tau + \frac{(1-\alpha)p}{\omega}\right]\right\}. \qquad (43)
\end{aligned}
$$

Hence, the provider's profit becomes

$$\Pi(t_o) = (\alpha p - c_1)\left[Q_{1A}(t_o) + Q_{1B}(t_o)\right] + (p - c_1)Q_2(t_o) - c_2(1 - 2t_o). \qquad (44)$$

We here introduce the following additional constraints so that the provider's profit can take on a positive value at its demand peak and a negative value at its demand off-peak;

$$
\begin{aligned}
n(2-\mu)(p-c_1) &> c_2, \qquad (45) \\
n\mu(p-c_1) &< c_2. \qquad (46)
\end{aligned}
$$

Further, we also assume

$$c_1 < \alpha p,$$

not to lose profit by the special offer. This provides a lower bound for α and consequently the domain of α is given by

$$\frac{c_1}{p} < \alpha < 1. \qquad (47)$$

The above observations yield the following proposition:

Proposition 1 *If $u_0 - p \le (1-\alpha)p$ and $\beta(\alpha p - c_1) > (1-\alpha)p$, the service provider can increase his profit by introducing a special offer.*

Proof. When the service provider should not offer the discounted price, αp, type $\mathscr{A}$ consumers with ideal service time t satisfying $t_o - \frac{u_0 - p}{\omega} \le t \le t_o + \tau + \frac{(1-\alpha)p}{\omega}$ would purchase a service product at its regular price, p ($> \alpha p$). This indicates that the service provider prepares himself for decrease in profit due to the special offer given by

$$\Pi_1 = (1-\alpha)p\int_{t_o - \frac{u_0 - p}{\omega}}^{t_o + \tau + \frac{(1-\alpha)p}{\omega}} q_t dt. \qquad (48)$$

At the same time, however, the special offer will induce type $\mathscr{A}$ consumers with ideal time t satisfying $t_o - \frac{u_0 - \alpha p}{\omega} \le t < t_o - \frac{u_0 - p}{\omega}$ to enjoy the special offer by shifting their ideal time, and thereby the provider can increase his profit by

$$\Pi_2 = (\alpha p - c_1)\int_{t_o - \frac{u_0 - \alpha p}{\omega}}^{t_o - \frac{u_0 - p}{\omega}} q_t dt. \qquad (49)$$

In addition, type $\mathscr{B}$ consumers with ideal time t satisfying $t_o - \frac{(1-\alpha)p}{\omega} \le t \le t_o + \tau + \frac{(1-\alpha)p}{\omega}$ would purchase a service product at αp, the provider can further increase his profit by

$$\Pi_3 = (\alpha p - c_1)\int_{t_o - \frac{(1-\alpha)p}{\omega}}^{t_o + \tau + \frac{(1-\alpha)p}{\omega}} \beta q_t dt. \qquad (410)$$

Let Π_0 be defined by $\Pi_0 \equiv \Pi_2 + \Pi_3 - \Pi_1$, then we have

$$\begin{aligned} \Pi_0 &= [\beta(\alpha p - c_1) - (1-\alpha)p] \int_{t_o - \frac{u_0 - p}{\omega}}^{t_o + \tau + \frac{(1-\alpha)p}{\omega}} q_t dt \\ &\quad + (\alpha p - c_1)\left[\beta \int_{t_o - \frac{(1-\alpha)p}{\omega}}^{t_o - \frac{u_0-p}{\omega}} q_t dt + \int_{t_o - \frac{u_0 - \alpha p}{\omega}}^{t_o - \frac{u_0-p}{\omega}} q_t dt\right]. \end{aligned} \tag{411}$$

If $\beta(\alpha p - c_1) > (1-\alpha)p$, the first term in the right-hand-side of Eq. (411) takes on a positive value. In addition, if $u_0 - p \le (1-\alpha)p$, the second term in the right-hand-side of Eq. (411) is also positive, and consequently, $\Pi_0 > 0$. ■

5 Optimal Strategy

This section seeks for an optimal opening time, t_o^*, and thereby an optimal closing time, t_c^*, can also be obtained by the symmetric structure of the ideal time distribution. Numerical examples are also presented to illustrate the proposed model formulation.

5.1 Analysis From Eq. (44), we have

$$\frac{d\Pi(t_o)}{dt_o} = (\alpha p - c_1)\left[\frac{dQ_{1A}(t_o)}{dt_o} + \frac{Q_{1B}(t_o)}{dt_o}\right] + (p - c_1)\frac{dQ_2(t_o)}{dt_o} + 2c_2. \tag{51}$$

By letting us denote, by $\pi(t_o)$, the right-hand-side of Eq. (51), we have

$$\begin{aligned} \pi(t_o) &= 8n(1-\mu)(\alpha p - c_1)\left\{\tau + \frac{u_0 + (1-2\alpha)p}{\omega} + \beta\left[\tau + \frac{2(1-\alpha)p}{\omega}\right]\right\} \\ &\quad -8n(1-\mu)(p-c_1)\left[t_o + \tau + \frac{(1-\alpha)p}{\omega}\right] - 2n\mu(p - c_1) + 2c_2, \end{aligned} \tag{52}$$

which indicates $\pi(t_o)$ is strictly decreasing in t_o.

In addition, we have

$$\begin{aligned} \pi\left(\frac{u_0 - \alpha p}{w}\right) &= -8n(1-\mu)(1-\alpha)p\left[\tau + \frac{u_0 + (1-2\alpha)p}{\omega}\right] \\ &\quad +8n\beta(1-\mu)(\alpha p - c_1)\left[\tau + \frac{2(1-\alpha)p}{\omega}\right] \\ &\quad -2n\mu(p-c_1) + 2c_2, \end{aligned} \tag{53}$$

$$\begin{aligned} \pi\left(\frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega}\right) &= 8n(1-\mu)(\alpha p - c_1)\left\{(\beta+1)\left[\tau + \frac{2(1-\alpha)p}{\omega}\right] + \frac{u_0 - p}{\omega}\right\} \\ &\quad -4n(1-\mu)(p-c_1) - 2n\mu(p-c_1) + 2c_2. \end{aligned} \tag{54}$$

Now, let A and B be defined by

$$\begin{aligned} A &\equiv -4n(1-\mu)(1-\alpha)p\left[\tau + \frac{u_0 + (1-2\alpha)p}{\omega}\right] \\ &\quad +4n\beta(1-\mu)(\alpha p - c_1)\left[\tau + \frac{2(1-\alpha)p}{\omega}\right] - n\mu(p-c_1) + c_2, \end{aligned} \tag{55}$$

$$\begin{aligned} B &\equiv 4n(1-\mu)(\alpha p - c_1)\left\{(\beta+1)\left[\tau + \frac{2(1-\alpha)p}{\omega}\right] + \frac{u_0 - p}{\omega}\right\} - 2n(1-\mu)(p-c_1) \\ &\quad -n\mu(p-c_1) + c_2, \end{aligned} \tag{56}$$

and then the optimal opening time, t_o^*, can be discussed under the following classification:

(a) If we have $A > 0$, further classification is necessary.

i) In the case of $B \geq 0$, t_o^* is given by

$$t_o^* = \frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega} = t_U.$$

ii) On the contrary, in case we have $B < 0$, t_o^* is given by

$$\begin{aligned} t_o^* &= \frac{\alpha p - c_1}{p - c_1}\left\{\tau + \frac{u_0 + (1-2\alpha)p}{\omega} + \beta\left[\tau + \frac{2(1-\alpha)p}{\omega}\right]\right\} \\ &\quad - \frac{\mu}{4(1-\mu)} + \frac{c_2}{4n(1-\mu)(p-c_1)} - \tau - \frac{(1-\alpha)p}{\omega}. \end{aligned}$$

(b) If we have $A \leq 0$, then $\pi(t_o) \leq 0$ and hence

$$t_o^* = \frac{u_0 - \alpha p}{\omega} = t_L.$$

As for the optimal opening time, t_o^*, we have the following proposition:

Proposition 2 *For the ideal time distribution with $\mu = 1$, if $p - c_1 \geq \frac{c_2}{n}$, the optimal opening time becomes $t_o^* = \frac{u_0 - \alpha p}{\omega} = t_L$, otherwise we have $t_o^* = \frac{1}{2} - \tau - \frac{(1-\alpha)p}{\omega} = t_U$.*

Proof. In the case of $\mu = 1$, the relationship, $p - c_1 \geq \frac{c_2}{n}$, reveals $A \leq 0$ from Eq. (55) along with $\pi(t_o) \leq 0$. On the contrary, $p - c_1 < \frac{c_2}{n}$ signifies $B > 0$ and hence $\pi(t_o) > 0$. ■

5.2 Numerical examples This subsection presents numerical examples to illustrate the proposed model. Table 1 shows the optimal opening time, t_o^*, and its corresponding profit, $\Pi(t_o^*)$, together with t_L and t_U against various values of μ and α when $(n, \tau, u_0, p, \omega, c_1, c_2, \beta) = (1, 0.05, 10, 9, 40, 4, 3, 0.35)$. It is observed in Table 1 that the optimal opening time, t_o^*, satisfies $t_L < t_o^* < t_U$ in the case of $\mu = 0.25$. In the case of $\mu = 0.5$ as well, $\alpha = 0.80$ and 0.85 indicate $t_L < t_o^* < t_U$. In the other cases, we have $t_o^* = t_L$. This is because the ideal time distribution shows a faltter shape with a smaller value of q_t at its demand peak when μ increases.

Table 1: Optimal strategies

μ	0.25			0.5			0.75		
α	0.75	0.80	0.85	0.75	0.80	0.85	0.75	0.80	0.85
t_L	0.081	0.07	0.059	0.081	0.07	0.059	0.081	0.07	0.059
t_U	0.398	0.405	0.416	0.398	0.405	0.416	0.398	0.405	0.416
t_o^*	0.145	0.159	0.167	0.081	0.092	0.100	0.081	0.07	0.059
$\Pi(t_o^*)$	2.278	2.410	2.500	2.130	2.246	2.323	2.043	2.191	2.287

Figure 2 shows the shape of the profit function, $\Pi(t_o)$, for $\mu = 0.25$, 0.50 and 0.75 against $\alpha = 0.8$ with the other parameter values set to the same for Table 1. It is observed in Fig. 1 that $\Pi(t_o)$ apparently has its maximum when $\mu = 0.25$. In the case of $\mu = 0.5$, $\Pi(t_o)$ has its maximum at $t_o = 0.092$ as shown in Table 1, while $\Pi(t_o)$ decreases with increasing t_o for $\mu = 0.75$.

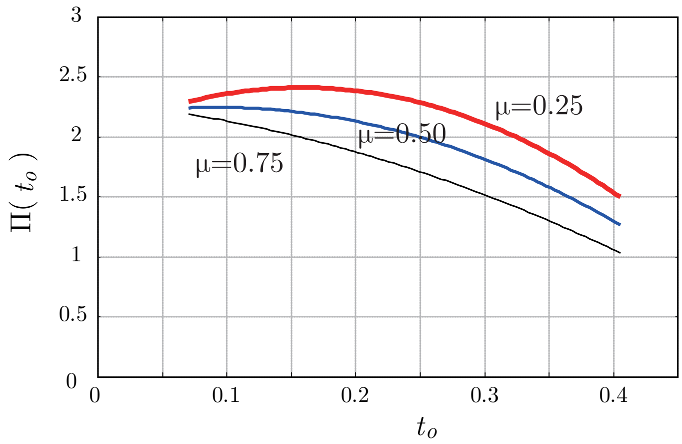

Figure 2: Behavior of profit function

6 Concluding Remarks

This study proposed a mathematical model of an optimal number of service hours for service providers that offer early birds specials and closing time discounts for a service product. By introducing the ideal service time distribution considered by Shy and Stancbaka, clarified were the conditions under which service proviers can earn more profit by special offers. The conditions were also shown that there exist optimal opening and closing times. Numerical examples were presented to illustrate the theoretical underpinnings of the proposed model formulation, and to show the effectiveness of introducing special offers for service providers.

Under the proposed model, however, optimality of the discounted price has not been discussed. One of useful extensions of our work is to explicitly introduce the price elasticity of demand to explore an optimal strategy with regard to the opening (and closing) time as well as the discounted price.

References

[1] Clemenz, G., Competition via shopping hours: a case for regulation? *Journal of Institutional and Theoretical Economics*, **150** (1994) pp. 625-641.

[2] Ferris, J.S., Time, space, and shopping: the regulation of shopping hours, *Journal of Law, Economics, and Organization*, **6** (1990) pp. 55-72.

[3] Inderst, R. and Irmen, A., Shopping hours and price competition, *European Economic Review*, **49** (2005) pp. 1105-1124.

[4] De Meza, D., The fourth commandment: is it Pareto efficient? *Economic Journal*, **94** (1984) pp. 379-383.

[5] Shy, O. and Stanbacka, R., Service hours with asymmetric distributions of ideal service time, *International Journal of Industrial Organization*, **24** (2006), pp. 763-771.

Communicated by *Koyu Uematsu*

Graduate School of Economics, OSAKA UNIVERSITY
1-7, Machikaneyama, Toyonaka, Osaka 560-0043 JAPAN.
sandoh@econ.osaka-u.ac.jp

A COMMENT ON A REGULARITY CONDITION IN A CURVED EXPONENTIAL FAMILY

Etsuo Kumagai

Received February 2, 2014; revised March 28, 2014

Abstract. Based on a curved exponential family, there is a regularity condition that the score function with random variables is the linear independence, which is commonly used in the information geometry. An equivalence relation to the regularity condition is that the Fisher information is positive definite under the curved exponential family. We investigate a key condition for two regularity conditions and we recognize it as the linear independence for the first derivative of natural parameter with respect to the parameter.

1 Introduction [3] introduced the ideas of the statistical curvature with respect to the asymptotic information loss. [1] introduced the statistical differential manifold and developed α-connection and m-connection in the curved exponential family. A lot of researchers have investigated the information geometry and there are a lot of fruitful and valuable results for the asymptotic.

In the framework of the information geometry, the statistical manifold $\{p(\boldsymbol{x};\boldsymbol{\theta})\}$ is based on the family of distributions with a parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbf{R}^k$. Among the regularity conditions in this framework, a regularity condition which we consider is that the derivatives $\{\partial \ell(\boldsymbol{\theta})/\partial \theta_i\}$ are linearly independent where $\ell(\boldsymbol{\theta})$ is the log-likelihood ([4](page 76), [2](page 29)). It seems that the linear independence on the derivatives is reasonable for constructing a tangent space in the statistical manifold, but we wonder whether this assumption is rational with respect to the underlying distribution of the random variable. Remark that we do not intend to consider singular models that do not satisfy the usual regularity conditions.

Based on a curved exponential family, an equivalence relation to the regularity condition that the score function with random variables is the linear independence is the regularity condition that the Fisher information is positive definite ([2](pages 25–29)). We investigate a key condition for above two regularity conditions and we recognize it as the linear independence for the first derivative of natural parameter with respect to the parameter.

2 The regularity condition [2](pages 25–29) considers a family of probability distribution on $\mathcal{X}$, i.e., $\mathcal{S} = \{p_\theta = p(\boldsymbol{x};\boldsymbol{\theta}) \,|\, \boldsymbol{\theta} \in \Theta \subseteq \mathbf{R}^k\}$ as k-dimensional statistical model on a set $\mathcal{X}$ which is a discrete set or $\mathbf{R}^m$ $(k \leq m)$. Letting p be a probability (density) function on $\mathcal{X}$, the support of p is defined by $\mathrm{supp}(p) \stackrel{\mathrm{def}}{=} \{\boldsymbol{x} \,|\, p(\boldsymbol{x}) > 0\}$ which is assumed to be constant with respect to $\boldsymbol{\theta}$, and $\mathcal{X}$ is redefined as $\mathrm{supp}(p)$, so that the statistical model $\mathcal{S}$ is a subset of

$$\mathcal{P}(\mathcal{X}) \stackrel{\mathrm{def}}{=} \left\{ p : \mathcal{X} \to \mathbf{R} \;\middle|\; p(\boldsymbol{x}) > 0 \ (\forall x \in \mathcal{X}),\ a \int_{\mathcal{X}} p(\boldsymbol{x}) d\boldsymbol{x} = 1 \right\}.$$

Note that the integral is interpreted as the summation if $\mathcal{X}$ is a discrete case. For the statistical model $\mathcal{S} = \{p_\theta\}$, defining the mapping $\varphi : \mathcal{S} \to \mathbf{R}^k$ by $\varphi(p_\theta) = \boldsymbol{\theta}$ implies a

2010 *Mathematics Subject Classification.* 62B10.

Key words and phrases. Curved exponential family, regularity condition, linear independence, information geometry.

coordinate system for $\mathcal{S}$. Furthermore suppose that there is a C^∞ diffeomorphism ψ : $\Theta \to \psi(\Theta)(\subset \mathbf{R}^k)$, so that, if we use $\boldsymbol{\eta} = \psi(\boldsymbol{\theta})$ as another parameter, then it holds that $\mathcal{S} = \{p_\theta \,|\, \boldsymbol{\theta} \in \Theta\} = \{p_{\psi^{-1}(\eta)} \,|\, \boldsymbol{\eta} \in \psi(\Theta)\}$. Thus a parameterization of $\mathcal{S}$ is a coordinate system of $\mathcal{S}$ as a C^∞ differentiable manifold.

Letting $[\theta^i]$ be a coordinate system in the statistical manifold $\mathcal{S}$ implies the vector fields formed by the natural bases $\{\partial_i\} \in T_\theta(\mathcal{S})$ which is the tangent space. Note that ∂_i means $\dfrac{\partial}{\partial\theta^i}$ and $\{\partial_i\}$ are vector fields. The Fisher information matrix at $\boldsymbol{\theta}$ in $\mathcal{S}$ is defined by the $k \times k$ matrix $G(\boldsymbol{\theta}) = (g_{ij}(\boldsymbol{\theta}))$ where the (i,j)-th element of $G(\boldsymbol{\theta})$ is, for $i,j = 1,\ldots,k$,

$$g_{ij}(\boldsymbol{\theta}) \;\stackrel{\text{def}}{=}\; E_\theta\Big[\partial_i\ell \;\partial_j\ell\Big] \;=\; \int_{\mathcal{X}} \partial_i\ell(\boldsymbol{x};\boldsymbol{\theta})\,\partial_j\ell(\boldsymbol{x};\boldsymbol{\theta})\,p(\boldsymbol{x};\boldsymbol{\theta})\,d\boldsymbol{x} \quad (\in \mathbf{R}),$$

where $\ell = \ell(\boldsymbol{x};\boldsymbol{\theta}) = \log p(\boldsymbol{x};\boldsymbol{\theta})$ is the log-likelihood function and E_θ means the expectation with respect to the distribution p_θ. [2] shows the assumptions as follows:

(page 28, line 3 from below to page 29, line 5) *The matrix $G(\boldsymbol{\theta})$ is symmetric $(g_{ij}(\boldsymbol{\theta}) = g_{ji}(\boldsymbol{\theta}))$, and since for any k-dimensional vector $\boldsymbol{c} = (c^1,\ldots,c^k)^t$ (^{t} denotes transpose),*

$$\boldsymbol{c}^t G(\boldsymbol{\theta})\boldsymbol{c} \;=\; \int \left\{\sum_{i=1}^{k} c^i\,\partial_i\ell(\boldsymbol{x};\boldsymbol{\theta})\right\}^2 p(\boldsymbol{x};\boldsymbol{\theta})\,d\boldsymbol{x} \;\geq\; 0, \tag{1}$$

it is also positive semidefinite. We assume further that $G(\boldsymbol{\theta})$ is positive definite. ¿From the equation above, we see that this is equivalent to stating that the elements of $\{\partial_1\ell,\ldots,\partial_k\ell\}$ when viewed as functions on $\mathcal{X}$ are linearly independent, which, in turn, is equivalent to stating that the elements of $\{\partial_1 p_\theta,\ldots,\partial_k p_\theta\}$ are linearly independent.

(page 29, lines 13–16, 18) *Now suppose that the assumption above hold, and define the inner product of the natural basis of the coordinate system $[\theta^i]$ by $g_{ij}(\boldsymbol{\theta}) = \langle\partial_i,\,\partial_j\rangle$. This uniquely determines a Riemannian metric $g(\boldsymbol{\theta}) = \langle,\rangle$. We call this the Fisher metric, or alternatively, the information metric.*

(page 29, line 18) *Indeed we may write $\langle X, Y\rangle_\theta = E_\theta[(X\ell)(Y\ell)]$ for all tangent vectors $X, Y \in T_\theta(\mathcal{S})$.*

The above assumptions seem to be the key to the regularity conditions in the information geometry. For convenience sake, we shall define the following conditions:

Condition A The matrix $G(\boldsymbol{\theta}) = \Big(\big\langle\partial_i,\,\partial_j\big\rangle_\theta\Big) = \Big(E_\theta\big[\partial_i\ell\;\partial_j\ell\big]\Big)$ in (1) is positive definite where ∂_i, ∂_j are the natural bases in $T_\theta(\mathcal{S})$ and ℓ is the log-likelihood function.

Condition B The elements of $\big\{\partial_1\ell,\ldots,\partial_k\ell\big\}$ are linearly independent.

In the curved exponential family with a parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbf{R}^k$, the density of the random variable $\boldsymbol{X} \in \mathbf{R}^m$ is $p(\boldsymbol{x};\boldsymbol{\theta}) = \exp\{\langle\boldsymbol{\alpha}(\boldsymbol{\theta}),\,\boldsymbol{x}\rangle - \psi(\boldsymbol{\alpha}(\boldsymbol{\theta}))\}p_0(\boldsymbol{x})$, where $p_0(\boldsymbol{x})$ is a pivotal probability measure, $\boldsymbol{\alpha}(\boldsymbol{\theta}) \in \mathbf{R}^m$ is a curved natural parameter parametrized by $\boldsymbol{\theta}$, the Euclidean inner product $\langle,\rangle$ in the exponent is the usual product of two vectors, and $\psi(\boldsymbol{\alpha}(\boldsymbol{\theta})) \in \mathbf{R}$ is the cumulant generating function. Note that the natural parameter space $\mathcal{A}$ is defined by $\mathcal{A} = \{\boldsymbol{\alpha} \in \mathbf{R}^m | \int \exp\{\langle\boldsymbol{\alpha},\,\boldsymbol{x}\rangle\}p_0(\boldsymbol{x})d\boldsymbol{x} < \infty\}$ and $\{\boldsymbol{\alpha}(\boldsymbol{\theta})\} \in \mathcal{A}$.

Now we assume that $\mathrm{supp}(p) = \mathbf{R}^m$ which is independent of the parameter $\boldsymbol{\theta}$. The usual derivative of the log-likelihood $\ell = \log p(\boldsymbol{x};\boldsymbol{\theta})$ with respect to the vector $\boldsymbol{\theta}$ is defined by

$$\partial_{\boldsymbol{\theta}}\,\ell \;\stackrel{\text{def}}{=}\; \frac{\partial\,\ell}{\partial\,\boldsymbol{\theta}} \;=\; \big(\partial_1\ell,\ldots,\partial_k\ell\big)^T \;=\; \big\langle\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}),\;\boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta})\big\rangle \;\;(\in\;\mathbf{R}^k),$$

where $\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \partial\boldsymbol{\alpha}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^T$ and $\boldsymbol{\beta}(\boldsymbol{\theta}) = E_\theta[\boldsymbol{X}]$, that is, the i-th element is

$$\partial_i \ell \;=\; \langle \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}),\ \boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta})\rangle \quad (\in \mathbf{R}) \quad (i = 1, \ldots, k),$$

where $\partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{\partial}{\partial\theta^i}\boldsymbol{\alpha}(\boldsymbol{\theta})$. The matrix $G(\boldsymbol{\theta})$ is obtained by

$$G(\boldsymbol{\theta}) \;=\; E_\theta\left[\langle \partial_{\boldsymbol{\theta}}\,\ell, \partial_{\boldsymbol{\theta}}\,\ell\rangle\right] \;=\; \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})^t\ \boldsymbol{\Sigma}(\boldsymbol{\theta})\ \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \quad (\in \mathbf{R}^{k\times k}), \tag{2}$$

where $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the covariance matrix of $\boldsymbol{X}$ under the probability $p(\boldsymbol{x};\boldsymbol{\theta})$.

Since the derivative of ℓ is $\boldsymbol{\partial}\ell = \langle \dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}),\ \boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta})\rangle$ for $k < m$, Condition B means the following relationship:

$$\sum_{i=1}^{k} c_i\,\partial_i \ell \;=\; \left\langle \sum_{i=1}^{k} c_i\,\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta}),\ \boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta})\right\rangle \;=\; 0 \implies \forall\, c_i \;=\; 0. \tag{3}$$

Although the derivatives $\{\partial_i \ell\}$ $(i = 1, \ldots, k)$ are supposed to be linearly independent, they are also random variables based on the distribution p_θ and the matrix $G(\boldsymbol{\theta})$ is assumed to be calculated by both the random variables $\{\partial_i \ell\}$ and their distribution p_θ. Note that, in the exponential family $(k = m)$, since $\boldsymbol{\partial}\ell = \boldsymbol{X} - \boldsymbol{\beta}$, Condition B means the following relationship:

$$\sum_{i=1}^{m} c_i\,\partial_i \ell \;=\; \sum_{i=1}^{m} c_i\,(X_i - \beta_i) \;=\; 0 \implies \forall\, c_i \;=\; 0. \tag{4}$$

3 An underlying condition for those regularity conditions With respect to the two conditions in the previous section, we investigate what are an underlying condition if Condition A is equivalent to Condition B under the curved exponential family.

Lemma 3.1 *In the curved exponential family, assume that the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is positive definite. Then Condition A is equivalent to the linear independence of $\{\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta})\}$, i.e.,*

$$\sum_{i=1}^{k} c_i\,\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta}) \;=\; \mathbf{0} \quad\implies\quad \forall c_i = 0. \tag{5}$$

Proof: Since the covariance matrix is positive definite, we decompose the matrix as follows: $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}$. Then the matrix $G(\boldsymbol{\theta})$ in (2) is decomposed by

$$G(\boldsymbol{\theta}) \;=\; \left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\right)^t \left(\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})\right)$$

and is considered as the Gram matrix, so that, by its property, Condition A is equivalent to that the k components $\{\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta})\}$ of $m \times 1$ vectors in $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ are linearly independent, i.e.,

$$\sum_{i=1}^{k} c_i\,\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta}) \;=\; \boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}\left(\sum_{i=1}^{k} c_i\,\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta})\right) \;=\; \mathbf{0} \quad\implies\quad \forall c_i = 0,$$

so that, since the matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})^{1/2}$ has the inverse, Condition A equals that $\{\partial_i\boldsymbol{\alpha}(\boldsymbol{\theta})\}$ are linearly independent. $\square$

This lemma means that Condition A depends on only the derivatives of the natural parameter $\boldsymbol{\alpha}(\boldsymbol{\theta})$, not the log-likelihood function $\ell(\boldsymbol{\theta})$ with the random variable $\boldsymbol{X}$ directly.

Note that, since $\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\alpha}$ in an exponential family, it holds that $\partial_i \boldsymbol{\alpha} = \boldsymbol{e}_i$ which is the i-th unit vector, so that the equivalent condition in Lemma 3.1 is

$$\sum_{i=1}^{m} c_i\, \boldsymbol{e}_i \;=\; \mathbf{0} \quad \Longrightarrow \quad \forall c_i = 0, \tag{6}$$

which is trivial because of the property of unit vectors. Next we consider the equivalent condition for Condition B.

Lemma 3.2 *In the curved exponential family, Condition B is equivalent to the condition (5), i.e., the linear independence of* $\{\partial_i \boldsymbol{\alpha}(\boldsymbol{\theta})\}$.

Proof: Since Condition B is that the derivatives $\{\partial_i \ell\}$ are linearly independent, i.e., (3), for the equation

$$\left\langle \sum_{i=1}^{k} c_i\, \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}),\ \boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta}) \right\rangle = 0 \tag{7}$$

in Condition B, we consider two cases as follows: Case (i) $\sum_{i=1}^{k} c_i\, \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}) = \mathbf{0}$ and Case (ii) $\sum_{i=1}^{k} c_i\, \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}) \neq \mathbf{0}$. For the Case (i), the equation (7) always holds without reference to the random variable $\boldsymbol{X}$, so that the Condition (5) implies Condition B.

On the other hand, for the Case (ii), it holds under the condition as follows:

$$\boldsymbol{X} - \boldsymbol{\beta}(\boldsymbol{\theta}) \;\in\; \mathcal{N}\left(\sum_{i=1}^{k} c_i\, \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}) \right), \tag{8}$$

which is a normal space against the vector $\sum_{i=1}^{k} c_i\, \partial_i \boldsymbol{\alpha}(\boldsymbol{\theta}) (\neq \mathbf{0})$. If the random variable $\boldsymbol{X}$ satisfies the condition (8) for the Case (ii), then the equation (7) holds, but the necessary condition $\forall c_i = 0$ in Condition B contradicts the Case (ii).

Therefore the equation (7) in Condition B is equivalent to the Case (i), i.e., the sufficient condition in (5) without reference to the random variable $\boldsymbol{X}$ in $\{\partial_i \ell\}$, so that we have the required result. □

Note that, since $\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\alpha}$ in an exponential family, it holds that, in the same fashion before, the equivalent condition in Lemma 3.2 is (4), which is equivalent to the condition (6) because of $\langle \boldsymbol{e}_i, \boldsymbol{X} - \boldsymbol{\beta} \rangle = X_i - \beta_i$ for $i = 1, \ldots, m$. The previous two lemmas imply the following theorem:

Theorem 3.1 *If the first derivative* $\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ *of the natural parameter is full rank in the curved exponential family and the covariance matrix* $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ *is positive definite, then, without reference to the random variable* $\boldsymbol{X}$ *in the derivatives of log-likelihood function, Condition A with respect to the Fisher information matrix* $\boldsymbol{G}(\boldsymbol{\theta})$ *is equivalent to Condition B with respect to the log-likelihood function* $\ell(\boldsymbol{\theta})$. □

Therefore, from Theorem 3.1, what [2](page 29) stated that *"We assume further that* $G(\boldsymbol{\theta})$ *is positive definite. ¿From the equation above, we see that this is equivalent to stating that the elements of* $\{\partial_1 \ell, \ldots, \partial_k \ell\}$ *when viewed as functions on* $\mathcal{X}$ *are linearly independent."*

just means that the derivatives $\{\partial_i \boldsymbol{\alpha}(\boldsymbol{\theta})\}$ are linearly independent under the positive definite covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$.

Because a relationship in the curved exponential family

$$\dot{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})\,\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \tag{9}$$

holds, we have the following corollary:

Corollary **3.1** *If the first derivative $\dot{\boldsymbol{\beta}}(\boldsymbol{\theta})$ of the expectation parameter is full rank in the curved exponential family and the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is positive definite, then, without reference to the random variable $\boldsymbol{X}$ in the derivatives of log-likelihood function, Condition A is equivalent to Condition B.*

Proof: Because of the relationship (9), if the first derivative $\dot{\boldsymbol{\beta}}(\boldsymbol{\theta})$ of the expectation parameter is full rank in a curved exponential family and the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is positive definite, then the first derivative $\dot{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ of the natural parameter is full rank, so that the derivatives $\{\partial_i \boldsymbol{\alpha}(\boldsymbol{\theta})\}$ are linearly independent and we have the required result by Theorem 3.1. □

4 Conclusion Based on the curved exponential family, we investigate the regularity conditions of Condition A and Condition B with respect to the linear independence and we conclude they are equivalent to the linear independence for the first derivative of the natural parameter with respect to the parameter under the condition that the covariance matrix is positive definite.

Acknowledgement: The author would like to express his appreciation to the referee.

References

[1] S.Amari (1985), *Differential-Geometrical Methods in Statistics*, Springer.

[2] S.Amari and H.Nagaoka (2000), *Methods of Information Geometry*, American Mathematical Society.

[3] B.Efron (1975), Defining the Curvature of a Statistical Problem (with Applications to Second Order Efficiency), *Ann. Stat.*, **3**, 1189–1242.

[4] M.K.Murray and J.W.Rice (1993), *Differential Geometry and Statistics*, Chapman & Hall.

Communicated by *Hisao Nagao*

Division of Mathematical Science,
Graduate School of Engineering Science,
Osaka University,
1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.

A MEAN VALUE PROPERTY FOR POLYCALORIC FUNCTIONS

Megumi Sano

Received January 11, 2014 ; revised March 28, 2014

Abstract. In this paper we prove a mean value property for polycaloric functions in one space dimensional case. The proof given here is a slight modification of that of the recent paper by F.Da Lio and L.Rodino [3] and seems more straightforward.

1 Introduction There are many papers that deal with a mean value property for polyharmonic functions (see [1, 2, 4, 6, 7] etc.). Especially, in 2011, G. Lysik ([7]) gave a simple and elegant proof of the following mean value property for polyharmonic functions and its inverse. Let $m \in \mathbf{N}$ and let U be a domain in $\mathbf{R}^N$. If $u \in C^{2m}(U)$ and $\Delta^m u = 0$, then for any ball $B_R(x) \subset U$ it holds

$$\frac{1}{|B_R(x)|}\int_{B_R(x)} u(y)dy = \sum_{k=0}^{m} \frac{\Delta^k u(x)}{4^k(\frac{N}{2}+1)_k k!}R^{2k} \tag{1.1}$$

$$\text{where } (a)_k = a(a+1)\cdots(a+k-1) \text{ for } k \in \mathbf{N}.$$

The main subject of this paper concerns the heat version of the result (1.1). First, we fix some terminologies. Let $U \subset \mathbf{R}^N$ be an open set and $U_T = U \times (0, T]$ denote a parabolic cylinder. We say that a function u defined on U_T is *caloric* if u is a solution of the linear heat equation $(\partial_t - \Delta_x)u(x,t) = 0$, $(x,t) \in U_T$, where $\Delta_x = \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2}$. Also, in this paper, u is called *polycaloric* if u is a solution of the equation $(\partial_t - \Delta_x)^m u(x,t) = 0, (x,t) \in U_T$ for some $m \in \mathbf{N}$. For fixed $x \in \mathbf{R}^N, t \in \mathbf{R}$, and $r > 0$, let

$$E(x,t;r) = \left\{ (y,s) \in \mathbf{R}^N \times \mathbf{R} \,\middle|\, s \le t, \Phi(x-y, t-s) \ge \frac{1}{r^N} \right\}$$

denote a heat ball with a top point (x,t), where

$$\Phi(x,t) = \begin{cases} \dfrac{1}{(4\pi t)^{N/2}} \exp\left(-\dfrac{|x|^2}{4t}\right) & (x \in \mathbf{R}^N, t > 0) \\ 0 & (x \in \mathbf{R}^N, t < 0) \end{cases}$$

is the fundamental solution of the heat equation. Note that a heat ball is symmetric with respect to y_i-axis $(i = 1, \cdots, N)$ and

$$E(0,0;1) = \left\{ (y,s) \in \mathbf{R}^N \times \mathbf{R} \mid -\frac{1}{4\pi} \le s < 0, |y| \le \sqrt{2Ns\log(-4\pi s)} \right\}.$$

It is well known that caloric functions possess the mean value property. Namely, if u is caloric on U_T, then for each heat ball $E(x,t;r) \subset U_T$ it holds

$$u(x,t) = \frac{1}{4r^N}\iint_{E(x,t;r)} u(y,s)\frac{|x-y|^2}{(t-s)^2}dyds \tag{1.2}$$

2010 *Mathematics Subject Classification.* Primary 35B05, Secondary 35K05, 35K25.
Key words and phrases. mean value, polycaloric function, Pizzetti's formula.

(see [5]: p.p 53-54 Theorem 3, or [10]). There is also an inverse mean value property of caloric functions under certain conditions ([9]).

Heat version of the result (1.1) is also known. Namely, in 2006, F. Da Lio and L. Rodino [3] proved the following asymptotic expansion formula for the heat integral mean (1.2) as a power series with respect to the radius of the heat ball:

Let $u \in C^\infty(\mathbf{R}^{N+1})$ and $(x,t) \in \mathbf{R}^{N+1}$, then it holds

$$\begin{aligned}\frac{1}{4r^N}\iint_{E(x,t;r)} u(y,s)\frac{|x-y|^2}{(t-s)^2}dyds \\ = u(x,t) + \sum_{k=1}^{M} r^{2k} H_k u(x,t) + O\left(r^{2M+2}\right) \quad \text{as } r \to 0,\end{aligned} \tag{1.3}$$

where H_k is given by

$$H_k u = \beta_{k,N}\left(\partial_t - \frac{N}{2k+N}\Delta_x\right)^{k-1}(\partial_t - \Delta_x)\, u \tag{1.4}$$

and

$$\beta_{k,N} = (-1)^k \frac{N}{k!}\frac{1}{(2k+N)}\left(\frac{N}{2k+N}\right)^{\frac{N}{2}+1}\left(\frac{1}{4\pi}\right)^k.$$

One of the key ideas in [3] is to introduce the differential operator H_k which is the k-th power of different heat operators whose diffusion coefficients depend on the iteration number k, though the exact meaning of H_k is less clear.

In this paper, we prove the formula (1.3) in [3] by another method, when the space dimension $N = 1$. We do not need to introduce the weighted power H_k and, in the author's opinion, the method seems more straightforward.

Theorem 1. *Let $N = 1$, $u \in C^\infty(U_T)$, $r > 0$ and $M \in \mathbf{N}$. Then we have*

$$\begin{aligned}&\frac{1}{4r}\iint_{E(x,t;r)} u(y,s)\frac{|x-y|^2}{(t-s)^2}dyds \\ &= u(x,t) + \sum_{k=1}^{M}\frac{r^{2k}}{k!}\sum_{l=0}^{k-1}(\partial_t - \partial_x^2)^{k-l}(\partial_t)^l u(x,t) \times C_{l,k} + O(r^{2M+2}) \quad \textit{as } r \to 0,\\ &\textit{where } C_{l,k} = \frac{(-1)^k}{(4\pi)^k(2k+1)^{k+\frac{3}{2}}}\begin{pmatrix} k-1 \\ l \end{pmatrix}(2k)^l.\end{aligned}$$

Theorem 1 is the formula (1.3) in one space dimensional case.

Remark 2. *Theorem 1 was proved in [3]. Indeed, by using the binomial theorem, we get*

$$H_k u = \beta_{k,N}\sum_{l=0}^{k-1}\begin{pmatrix} k-1 \\ l \end{pmatrix}\left(\frac{2k}{2k+N}\right)^l\left(\frac{N}{2k+N}\right)^{k-1-l}(\partial_t - \Delta_x)^{k-l}(\partial_t)^l u.$$

Therefore we obtain Theorem 1 for general dimensional case. However we do not need to introduce the differential operator H_k (1.4) in one space dimensional case. An assumption of one space dimension is a technical problem due to obtain representations (2.7, 2.8) in Lemma 4 by factorizing $v^{(2k)}(0)$ $(k = 1, 2, \cdots)$ concretely (see §2). It seems to be difficult in higher dimension case.

We finally give mean value properties for the polycaloric equation (see Corollary 10 in §3) and the higher order heat equation (see Proposition 12 in §3). The author hopes that mean value properties are useful for getting qualitative properties of solutions for the polycaloric equation and the higher order heat equation.

2 Proof of theorem 1 In this section, we prove Theorem 1. We set $(x,t)=(0,0)$ to simplify the description. Let $u:\mathbf{R}^N\times\mathbf{R}\to\mathbf{R}$ be a smooth function. Set $E(r)=E(0,0,r)$ and put

$$\phi(r)=\frac{1}{r^N}\iint_{E(r)}u(x,t)\frac{|x|^2}{t^2}dxdt=\iint_{E(1)}u(ry,r^2s)\frac{|y|^2}{s^2}dyds. \tag{2.1}$$

In the following, we will carry out the Maclaurin expansion of $\phi(r)$ with respect to $r\in\mathbf{R}$. Set $v(r)=u(x,t)=u(ry,r^2s)$ for $(y,s)\in\mathbf{R}^N\times\mathbf{R}$. By differentiating $\phi(r)$ directly, we have

$$\phi^{(n)}(0)=\iint_{E(1)}v^{(n)}(0)\frac{|y|^2}{s^2}dyds. \tag{2.2}$$

We use standard notations of multi-indices; for $y=(y_1,\cdots,y_N)\in\mathbf{R}^N$ and a multi-index $\alpha=(\alpha_1,\cdots,\alpha_N)\in\mathbf{N}_0^N$, we write $y^\alpha=y_1^{\alpha_1}\cdots y_N^{\alpha_N}$ and $|\alpha|=\alpha_1+\cdots+\alpha_N$. Next lemma concerns the evaluation of $v^{(n)}(0)$ and is valid for general dimension $N\in\mathbf{N}$.

Lemma 3 ($v^{(n)}(0)$)**.** *For $k\in\mathbf{N}_0$, we obtain*

$$\phi^{(2k-1)}(0)=0, \tag{2.3}$$

$$v^{(2k)}(0)=\sum_{j=0}^{k}\sum_{|\beta|=k-j}(\partial_x^2)^\beta(\partial_t)^j u(0,0)\times A_{\beta,k}(y,s) \tag{2.4}$$

where

$$A_{\beta,k}(y,s)=\frac{(2k)!}{(2\beta)!j!}y^{2\beta}s^j.$$

Proof. Since $v(r)$ is a C^∞ function of r, for all $M\geq 1$ we have

$$v(r)=\sum_{n=0}^{2M+1}\frac{v^{(n)}(0)}{n!}r^n+O(r^{2M+2})\ \text{ as }\ r\to 0. \tag{2.5}$$

On the other hand, since $v(r)$ is a composed function of $u(x,t)$ and $x=ry, t=r^2s$, we have

$$\begin{aligned} v(r)&=\sum_{m=0}^{2M+1}\frac{1}{m!}\left((ry_1)\frac{\partial}{\partial x_1}+\cdots+(ry_N)\frac{\partial}{\partial x_N}+(r^2s)\frac{\partial}{\partial t}\right)^m u(0,0)+O(r^{2M+2})\\ &=\sum_{m=0}^{2M+1}\frac{1}{m!}\sum_{|\alpha|+j=m}\frac{m!}{\alpha_1!\cdots\alpha_N!j!}(ry)^\alpha(r^2s)^j(\partial_x^\alpha\partial_t^j)u(0,0)+O(r^{2M+2})\\ &=\sum_{m=0}^{2M+1}\sum_{|\alpha|+j=m}\frac{y^\alpha s^j}{\alpha!j!}\ (\partial_x^\alpha\partial_t^j)u(0,0)\times r^{|\alpha|+2j}+O(r^{2M+2}). \end{aligned} \tag{2.6}$$

By comparing the coefficients of r^n in the both expressions of (2.5) and (2.6), we obtain

$$\frac{v^{(n)}(0)}{n!}=\sum_{|\alpha|+2j=n}\frac{y^\alpha s^j}{\alpha!j!}\ (\partial_x^\alpha\partial_t^j)u(0,0).$$

Thus,

$$\begin{aligned}\phi^{(n)}(0) &= \iint_{E(1)} v^{(n)}(0)\frac{|y|^2}{s^2}dyds \\ &= \sum_{|\alpha|+2j=n} \frac{n!}{\alpha!j!}\,(\partial_x^\alpha \partial_t^j)u(0,0) \times \iint_{E(1)} y^\alpha s^j \frac{|y|^2}{s^2}dyds\,.\end{aligned}$$

Since $E(1)$ is symmetric about y_i-axis($i=1,\cdots,N$), $\iint_{E(1)} y^\alpha s^j \frac{|y|^2}{s^2}dyds$ vanishes when at least one α_i of $\alpha=(\alpha_1,\cdots,\alpha_N)$ is odd (i.e. when n is odd because $|\alpha|+2j=n$). This proves (2.3). Next, we consider the case $\alpha=2\beta$ for some $\beta\in\mathbf{N}_0^N$ and let $n=2k$ $(k\in\mathbf{N})$. Then we obtain

$$\begin{aligned}v^{(2k)}(0) &= \sum_{2|\beta|+2j=2k} (\partial_x^2)^\beta(\partial_t)^j u(0,0) \times \frac{(2k)!}{(2\beta)!j!}y^{2\beta}s^j \\ &= \sum_{j=0}^{k}\sum_{|\beta|=k-j} (\partial_x^2)^\beta(\partial_t)^j u(0,0) \times \frac{(2k)!}{(2\beta)!j!}y^{2\beta}s^j,\end{aligned}$$

which implies (2.4).

□

Lemma 4 (Factorization). *Let $N=1$. Then*

$$v^{(2k)}(0) = \sum_{l=0}^{k}(\partial_t-\partial_x^2)^{k-l}(\partial_t)^l u(0,0) \times B_{l,k}(y,s) \tag{2.7}$$

where

$$B_{l,k}(y,s) = (-1)^{k+l}\sum_{m=0}^{l}\begin{pmatrix} k-l+m \\ m\end{pmatrix} \times A_{k-l+m,k}(y,s) \tag{2.8}$$

for $0\le l\le k$.

Proof. By the assumption $N=1$ and (2.4), it is enough to prove that

$$\sum_{j=0}^{k}(\partial_x^2)^{k-j}(\partial_t)^j u(0,0)\times A_{k-j,k} = \sum_{l=0}^{k}(\partial_t-\partial_x^2)^{k-l}(\partial_t)^l u(0,0)\times B_{l,k}. \tag{2.9}$$

We prove (2.9) by comparing the coefficients of $(\partial_x^2)^{k-j}(\partial_t)^j u(0,0)$ in both sides.

Since

$$\begin{aligned}&\sum_{l=0}^{k}(\partial_t-\partial_x^2)^{k-l}(\partial_t)^l u(0,0)B_{l,k} \\ &= (\partial_t-\partial_x^2)^k u(0,0)B_{0,k} + (\partial_t-\partial_x^2)^{k-1}(\partial_t)u(0,0)B_{1,k} + \cdots + (\partial_t)^k u(0,0)B_{k,k},\end{aligned}$$

the coefficient of $(\partial_x^2)^{k-j}(\partial_t)^j u(0,0)$ on the right hand side of (2.9) is given by

$$\begin{aligned}
&(-1)^{k-j}\left[\binom{k}{k-j}B_{0,k}+\binom{k-1}{k-j}B_{1,k}+\binom{k-2}{k-j}B_{2,k}+\cdots\right.\\
&\quad\left.+\binom{k-j+1}{k-j}B_{j-1,k}+\binom{k-j}{k-j}B_{j,k}\right]\\
&=(-1)^{k-j}\sum_{l=0}^{j}\binom{k-l}{k-j}B_{l,k}.
\end{aligned}$$

Inserting the definition of $B_{l,k}$ in (2.8) into this expression, we assure that the coefficient of $(\partial_x^2)^{k-j}(\partial_t)^j u(0,0)$ on the right hand side of (2.9) is given by

$$(-1)^{k-j}\sum_{l=0}^{j}\binom{k-l}{k-j}(-1)^{k+l}\sum_{m=0}^{l}\binom{k-l+m}{m}\times A_{k-l+m,k}. \tag{2.10}$$

Since

$$\begin{aligned}
&\sum_{l=0}^{j}\binom{k-l}{k-j}(-1)^{k+l}\sum_{m=0}^{l}\binom{k-l+m}{m}A_{k-l+m,k}\\
&=\binom{k}{k-j}(-1)^k\binom{k}{0}A_{k,k}\\
&+\binom{k-1}{k-j}(-1)^{k+1}\left[\binom{k-1}{0}A_{k-1,k}+\binom{k}{1}A_{k,k}\right]\\
&+\binom{k-2}{k-j}(-1)^{k+2}\left[\binom{k-2}{0}A_{k-2,k}+\binom{k-1}{1}A_{k-1,k}+\binom{k}{2}A_{k,k}\right]\\
&+\cdots\\
&+\binom{k-j}{k-j}(-1)^{k+j}\left[\binom{k-j}{0}A_{k-j,k}+\cdots+\binom{k-1}{j-1}A_{k-1,k}+\binom{k}{j}A_{k,k}\right],
\end{aligned}$$

coefficients of $A_{k-i,k}$ for all $0\le i\le j-1$ in (2.10) is given by

$$\begin{aligned}
&(-1)^{k-j}(-1)^{k+i}\sum_{n=0}^{j-i}(-1)^n\binom{k-i-n}{k-j}\binom{k-i}{n}\\
&=(-1)^{i-j}\sum_{n=0}^{j-i}(-1)^n\binom{k-i}{k-j}\binom{j-i}{n}=0,
\end{aligned}$$

where the last equality comes from $\sum_{n=0}^{p}(-1)^n\binom{p}{n}=(-1+1)^p=0$.

Then we prove that

$$\sum_{l=0}^{j}\binom{k-l}{k-j}(-1)^{k+l}\sum_{m=0}^{l}\binom{k-l+m}{m}A_{k-l+m,k}=\binom{k-j}{k-j}(-1)^{k+j}A_{k-j,k}. \tag{2.11}$$

Therefore, by (2.10) and (2.11), the coefficient of $(\partial_x^2)^{k-j}(\partial_t)^j u(0,0)$ on the right hand side of (2.9) is $A_{k-j,k}$. We have thus proved Lemma 4. □

From (2.2) and (2.7), we deduce

$$\phi^{(2k)}(0) = \sum_{l=0}^{k} (\partial_t - \partial_x^2)^{k-l} (\partial_t)^l u(0,0) \times \iint_{E(1)} B_{l,k}(y,s) dyds \tag{2.12}$$

Note that, on the right hand side of (2.12), the heat operator $(\partial_t - \partial_x^2)$ acts on u except for $l = k$.

Lemma 5. *We put*

$$\tilde{C}_{l,k} = \iint_{E(1)} B_{l,k}(y,s) \frac{y^2}{s^2} dyds.$$

Then we get

$$\tilde{C}_{l,k} = \frac{(2k)!(-1)^k 4}{k!(4\pi)^k (2k+1)^{k+\frac{3}{2}}} \begin{pmatrix} k-1 \\ l \end{pmatrix} (2k)^l \tag{2.13}$$

for $0 \le l \le k-1$ *and* $\tilde{C}_{k,k} = 0$.

Proof. We prove Lemma 5 by simple calculations. First, by the definition of $B_{l,k}$ in (2.8)

$$B_{l,k} = (-1)^{k+l} \sum_{m=0}^{l} \begin{pmatrix} k-l+m \\ m \end{pmatrix} \frac{(2k)!}{(2k-2l+2m)!(l-m)!} y^{2k-2l+2m} s^{l-m}$$

for $0 \le l \le k$, we have

$$\tilde{C}_{l,k} = (-1)^{k+l} \sum_{m=0}^{l} \begin{pmatrix} k-l+m \\ m \end{pmatrix} \frac{(2k)!}{(2k-2l+2m)!(l-m)!} \iint_{E(1)} y^{2k-2l+2m+2} s^{l-m-2} dyds.$$

Direct calculation shows that

$$\begin{aligned}
\iint_{E(1)} y^{2k-2l+2m+2} s^{l-m-2} dyds &= \int_{s=-1/4\pi}^{s=0} s^{l-m-2} \int_{|y| \le \sqrt{2s \log(-4\pi s)}} y^{2k-2l+2m+2} dyds \\
&= \frac{2}{(2k-2l+2m+3)} \int_{-1/4\pi}^{0} s^{l-m-2} \left\{ 2s \log(-4\pi s) \right\}^{k-l+m+\frac{3}{2}} ds \\
&= \frac{(-1)^{l-m} 2^{k-l+m+\frac{3}{2}}}{(k-l+m+\frac{3}{2})(4\pi)^{k+\frac{1}{2}}} \int_0^\infty t^{k-l+m+\frac{3}{2}} \exp\left(-\left(k+\frac{1}{2}\right) t \right) dt \\
&= \frac{(-1)^{l-m} 4^{k-l+m} 2^3 \Gamma(k-l+m+\frac{3}{2})}{(4\pi)^k \sqrt{\pi} (2k+1)^{k-l+m+\frac{5}{2}}}
\end{aligned}$$

where $\Gamma(\cdot)$ is the Gamma function. Thus, we get

$$\begin{aligned}
\tilde{C}_{l,k} &= \frac{(-1)^k (2k)! 4^{k-l} 8}{(4\pi)^k \sqrt{\pi} (2k+1)^{k-l+\frac{5}{2}} (k-l)!} \sum_{m=0}^{l} \frac{(-1)^m (k-l+m)! 4^m \Gamma(k-l+m+\frac{3}{2})}{m!(2k-2l+2m)!(l-m)!(2k+1)^m}. \\
&= \frac{(-1)^k (2k)! 4}{k!(4\pi)^k (2k+1)^{k-l+\frac{5}{2}}} \begin{pmatrix} k \\ l \end{pmatrix} \sum_{m=0}^{l} (-1)^m \begin{pmatrix} l \\ m \end{pmatrix} \frac{2k-2l+2m+1}{(2k+1)^m},
\end{aligned}$$

where the last equality comes from the fact $\Gamma(s+1) = s\Gamma(s)$.

Since we have the following equation

$$
\begin{aligned}
&(2k+1)^l \sum_{m=0}^{l} (-1)^m \binom{l}{m} \frac{2k-2l+2m+1}{(2k+1)^m} \\
&= (2k+1) \sum_{m=0}^{l} \binom{l}{m} (-1)^m (2k+1)^{l-m} - 2 \sum_{m=0}^{l-1} (-1)^m \binom{l}{m} (l-m)(2k+1)^{l-m} \\
&= (2k+1)(2k)^l - 2l(2k+1) \sum_{m=0}^{l-1} (-1)^m \binom{l-1}{m} (2k+1)^{l-m-1} \\
&= 2(k-l)(2k)^{l-1}(2k+1)
\end{aligned}
$$

Therefore we obtain $\tilde{C}_{k,k} = 0$ and (2.13). □

From all Lemmas, we obtain

$$\phi^{(2k)}(0) = \sum_{l=0}^{k-1} (\partial_t - \partial_x^2)^{k-l} (\partial_t)^l u(0,0) \times \tilde{C}_{l,k} \quad (k = 1, 2, \ldots),$$

which proves Theorem 1.

3 A mean value property for polycaloric functions In this section, first we recall the well-known regularity property of (poly-) caloric functions.

Proposition 6 (caloric function is smooth)**.** *If $u : U_T \to \mathbf{R}$ is caloric, then $u \in C^\infty(U_T)$.*

Proof. See [5]: p.p 59-61 Theorem 8. □

Proposition 7 (polycaloric function is smooth)**.** *If $u : U_T \to \mathbf{R}$ is polycaloric, then $u \in C^\infty(U_T)$.*

Proof. Assume that there exists $m \in \mathbf{N}$ such that $(\partial_t - \Delta_x)^m u = 0$ in U_T. Then we find caloric functions $u_0, u_1, \cdots, u_{m-1} : U_T \to \mathbf{R}$ such that

$$u(x,t) = u_0(x,t) + t u_1(x,t) + \cdots + t^{m-1} u_{m-1}(x,t) \tag{3.1}$$

holds true, by proposition 1 in [8]. Indeed, for $j = 1, 2, \cdots, m$, we may choose

$$u_{m-j}(x,t) = \frac{1}{(m-j)!} \sum_{k=0}^{j-1} \frac{(-t)^k}{k!} (\partial_t - \Delta_x)^{m-j+k} u(x,t).$$

Therefore $u_0, u_1, \cdots, u_{m-1}$ are caloric and satisfy the equation (3.1). By proposition 6 and (3.1), we obtain $u \in C^\infty(U_T)$. □

By proposition 6 and proposition 7, we obtain several corollaries which are proved by F.Da Lio and L.Rodino [3] as follows. We do not need the additional assumption that u is smooth, after assuming that u is caloric or polycaloric.

corollary 8 (A mean value property for analytic functions. [3] Proposition 2.2)**.** *Let $N = 1$ and $u \in C^\infty(U_T)$. Assume that $(\partial_t - \partial_x^2)u(x,t)$ is an analytic function in U_T. Then $\phi(r)$ given in (2.1) is an analytic function of $r \in \mathbf{R}$ in a neighborhood of $r = 0$, and it holds*

$$\begin{aligned}
&\frac{1}{4r}\iint_{E(x,t;r)} u(y,s)\frac{(x-y)^2}{(t-s)^2}dyds \\
&= u(x,t) + \sum_{k=1}^{\infty}\frac{r^{2k}}{k!}\sum_{l=0}^{k-1}(\partial_t - \partial_x^2)^{k-l}(\partial_t)^l u(x,t) \times C_{l,k} \\
&\textit{where } C_{l,k} = \frac{(-1)^k}{(4\pi)^k(2k+1)^{k+\frac{3}{2}}}\begin{pmatrix} k-1 \\ l \end{pmatrix}(2k)^l.
\end{aligned}$$

Remark 9. *If u is caloric on U_T, then $u \in C^\infty(U_T)$ and $(\partial_t - \partial_x^2)u(x,t)$ is obviously analytic in U_T and for each heat ball $E(x,t;r) \subset U_T$ the following equation holds:*

$$\frac{1}{4r}\iint_{E(x,t;r)} u(y,s)\frac{(x-y)^2}{(t-s)^2}dyds = u(x,t).$$

Corollary 10 can be considered as the generalization of (1.1) to the polycaloric case.

corollary 10 (A mean value property for polycaloric functions)**.** *Let $N = 1$ and $(\partial_t - \partial_x^2)u(x,t)$ be an analytic function in U_T. If u is polycaloric on U_T (i.e.$(\partial_s - \partial_y^2)^m u(y,s) = 0$, $(y,s) \in U_T$, $m \in \mathbf{N}$), then for each heat ball $E(x,t;r) \subset U_T$ the following equality holds:*

$$\begin{aligned}
&\frac{1}{4r}\iint_{E(x,t;r)} u(y,s)\frac{(x-y)^2}{(t-s)^2}dyds \\
&= u(x,t) + \sum_{k=1}^{m-1}\frac{r^{2k}}{k!}\sum_{l=0}^{k-1}(\partial_t - \partial_x^2)^{k-l}(\partial_t)^l u(x,t) \times C_{l,k} \\
&+ \sum_{k=m}^{\infty}\frac{r^{2k}}{k!}\sum_{l=k-m+1}^{k-1}(\partial_t - \partial_x^2)^{k-l}(\partial_t)^l u(x,t) \times C_{l,k}, \\
&\textit{where } C_{l,k} = \frac{(-1)^k}{(4\pi)^k(2k+1)^{k+\frac{3}{2}}}\begin{pmatrix} k-1 \\ l \end{pmatrix}(2k)^l.
\end{aligned}$$

Proof. This is a direct consequence of Theorem 1 and Proposition 7. □

corollary 11 ([3] Corollary 2.1)**.** *Let $N = 1$. Suppose that there exist $n_1 \geq 0$ and $n_2 \geq 1$ such that*

$$(\partial_t - \partial_x^2)(\partial_t)^{n_1}u = 0 \ \textit{ and } \ (\partial_t - \partial_x^2)^{n_2}u = 0 \ \textit{ in } U_T.$$

Then for all $r > 0$ we have

$$\begin{aligned}
&\frac{1}{4r}\iint_{E(x,t;r)} u(y,s)\frac{(x-y)^2}{(t-s)^2}dyds \\
&= u(x,t) + \sum_{k=1}^{M}\frac{r^{2k}}{k!}\sum_{l=0}^{k-1}(\partial_t - \partial_x^2)^{k-l}(\partial_t)^l u(x,t) \times C_{l,k},
\end{aligned} \tag{3.2}$$

with $M = n_1 + n_2 - 1$ (when $n_1 = 0$ or $n_2 = 1$ the sum in the right-hand side of (3.2) does not appear).

Proof. Note that we get $u \in C^\infty(U_T)$, since u is polycaloric in U_T. See the proof of corollary 2.1 in [3]. $\square$

We finally give a mean value property for the higher order heat equation $\partial_t u + (-1)^m \Delta^m u = 0$ $(m \in \mathbf{N})$ for general dimension. In the proof, we use proposition 2.2 and a result in the proof of proposition 2.1 in [3].

Proposition 12 (A mean value property for the higher order heat equation)**.** *Let $u \in C^\infty(U_T)$ and $(\partial_t - \Delta_x)u(x,t)$ be an analytic function in U_T. Assume that u is a solution of the higher order heat equation $\partial_t u + (-1)^m \Delta^m u = 0$. Then for each heat ball $E(x,t;r) \subset U_T$ the following equality holds:*

$$\frac{1}{4r^N} \iint_{E(x,t;r)} u(y,s) \frac{|x-y|^2}{(t-s)^2} dyds = u(x,t) + \sum_{k=1}^{\infty} r^{2k} H_k u(x,t), \tag{3.3}$$

where H_k is given by

$$H_k u = \begin{cases} \dfrac{\rho_{k,N}}{k!} \displaystyle\sum_{h=0}^{k} (-1)^{k-h} \begin{pmatrix} k \\ h \end{pmatrix} (N+2h) \left(\frac{N}{2k+N} \right)^h \Delta^{mk+(1-m)h} u, & (m : odd) \\ \dfrac{\rho_{k,N}}{k!} \displaystyle\sum_{h=0}^{k} \begin{pmatrix} k \\ h \end{pmatrix} (N+2h) \left(\frac{N}{2k+N} \right)^h \Delta^{mk+(1-m)h} u, & (m : even) \end{cases}$$

$$\text{where} \quad \rho_{k,N} = \frac{1}{2k+N} \left(\frac{N}{2k+N} \right)^{\frac{N}{2}+1} \left(\frac{1}{4\pi} \right)^k .$$

Proof. Let $p \in \mathbf{N}$. Note that u satisfies

$$\partial_t^p u = \begin{cases} \Delta^{pm} u, & (m : \text{odd}) \\ (-1)^p \Delta^{pm} u, & (m : \text{even}) \end{cases} \tag{3.4}$$

since u is a smooth solution of the higher order heat equation $\partial_t u + (-1)^m \Delta^m u = 0$. On the other hand, (3.3) holds by proposition 2.2 in [3], and according to a result in [3] (p,268, line 2 and 9), H_k is given by

$$H_k u = \frac{\rho_{k,N}}{k!} \sum_{h=0}^{k} (-1)^{k-h} \begin{pmatrix} k \\ h \end{pmatrix} (N+2h) \left(\frac{N}{2k+N} \right)^h \Delta^h (\partial_t)^{k-h} u. \tag{3.5}$$

Finally, combining (3.4) and (3.5), we get the proposition 12. $\square$

4 Acknowledgment The author expresses a deep appreciation to Professor Futoshi Takahashi (Osaka City University) for suggesting this topic to me and for his invaluable advice and encouragement. Moreover part of this work is supported by the JASSO Student Exchange Support Program 2013 (Scholarship for Short Term Visit) "Graduate Student Oversea Short Visits Research Program by International Exchange Network in Sciences". The author also thanks all people related to this program, in particular Professor Yoshihiro Ohnita (a director of OCAM I).

References

[1] B.Bojanov, *An extension of the Pizzetti formula for polyharmonic functions*, Acta Math. Hunger. 91 (2001), 99-113.

[2] J.H.Bramble and L.E.Payne, *Mean value theorems for polyharmonic functions*, Amer. Math. Monthly 73 (1966), 124-127.

[3] F.Da Lio and L.Rodino, *A Pizzetti-type formula for the heat operator*, Arch. Math. (Basel) 87 (2006), 261-271.

[4] Dupaigne, Louis, *Stable Solutions of Elliptic Partial Differential Equations*, Chapman Hall/CRC Monogr. Surv. Pure Appl. Math. (2011).

[5] Lawrence C.Evans, *Partial Differential Equations (Second Edition)*, Amer. Math. Soc., Graduate Studies in Mathematics (2010).

[6] A.Friedman, *Mean-values and polyharmonic polynomials*, Michigan Math. J., **4** (1957), 67-74.

[7] G.Lysik, *On the mean value property for polyharmonic functions*, Acta Math. Hungar. , **133** (1-2) (2011), 133-139.

[8] M.Nishio, K.Shimomura, N.Suzuki, *A mean value property of poly-temperatures on a strip domain*, J. London Math. Soc. (2) 58 (1998), no. 2, 381-393.

[9] N.Suzuki and N.A.Watson, *A characterization of heat balls by a mean value property for temperatures*, Proc. Amer. Math. Soc., Vol. 129, No. 9 (2001), 2709-2713.

[10] N.A.Watson, *A Theory of subtemperatures in several variables*, Proc. London Math. Society **26** (1973), 385-417.

Communicated by *Yoshitsugu Kabeya*

Department of Mathematics, Graduate School of Science,
Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-8585 JAPAN
Email address ： m13saM0311@st.osaka-cu.ac.jp

Submission to the SCMJ

In September 2012, the way of submission to Scientiae Mathematicae Japonicae (SCMJ) was changed. Submissions should be sent electronically (in PDF file) to the editorial office of International Society for Mathematical Sciences (ISMS).

(1) Preparation of files and Submission

a. Authors who would like to submit their papers to the SCMJ should make source files of their papers in LaTeX2e using the ISMS style file (scmjlt2e.sty) Submissions should be in PDF file compiled from the source files. Send the PDF file to s1bmt@jams.jp .

b. Prepare a Submission Form and send it to the ISMS. The required items to be contained in the form are:

1. Editor's name whom the author chooses from the Editorial Board (http://www.jams.or.jp/hp/submission_f.html)and would like to take in charge of the paper for refereeing.
2. Title of the paper.
3. Authors' names.
4. Corresponding author's name, e-mail address and postal address (affiliation).
5. Membership number in case the author is an ISMS member.

Japanese authors should write 3 and 4 both in English and in Japanese.

At http://www.jams.or.jp/hp/submission_f.html, the author can find the Submission Form. Fulfill the Form and sent it to the editorial office by pushing the button "transmission". Or, without using the Form, the author may send an e-mail containing the items 1-5 to s1bmt@jams.jp

(2) Registration of Papers

When the editorial office receives both a PDF file of a submitted paper and a Submission Form, we register the paper. We inform the author of the registration number and the received date. At the same time, we send the PDF file to the editor whom the author chooses in the Submission Form and request him/her to begin the process of refereeing. (Authors need not send their papers to the editor they choose.)

(3) Reviewing Process

a. The editor who receives, from the editorial office, the PDF file and the request of starting the reviewing process, he/she will find an appropriate referee for the paper.

b. The referee sends a report to the editor. When revision of the paper is necessary, the editor informs the author of the referee's opinion.

c. Based on the referee report, the editor sends his/her decision (acceptance of rejection) to the editorial office.

(4) a. Managing Editor of the SCMJ makes the final decision to the paper valuing the editor's decision, and informs it to the author.

b. When the paper is accepted, we ask the author to send us a source file and a PDF file of the final manuscript.

c. The publication charges for the ISMS members are free if the membership dues have been paid without delay. If the authors of the accepted papers are not the ISMS members, they should become ISMS members and pay ¥6,000 (US$75, Euro55) as the membership dues for a year, or should just pay the same amount without becoming the members.

Items required in Submission Form

1. Editor's name who the authors wish will take in charge of the paper
2. Title of the paper
3. Authors' names

3'. 3. in Japanese for Japanese authors

4. Corresponding author's name and postal address (affiliation)

4'. 4. in Japanese for Japanese authors

5. ISMS membership number
6. E-mail address

Call for ISMS Members

Call for Academic and Institutional Members

Discounted subscription price: When organizations become the Academic and Institutional Members of the ISMS, they can subscribe our journal Scientiae Mathematicae Japonicae at the yearly price of US$225. At this price, they can add the subscription of the online version upon their request.

Invitation of two associate members: We would like to invite two persons from the organizations to the associate members with no membership fees. The two persons will enjoy almost the same privileges as the individual members. Although the associate members cannot have their own ID Name and Password to read the online version of SCMJ, they can read the online version of SCMJ at their organization.

To apply for the Academic and Institutional Member of the ISMS, please use the following application form.

Application for Academic and Institutional Member of ISMS

Subscription of SCMJ Check one of the two.	☐Print (US$225) ☐Print ＋ Online (US$225)
University (Institution)	
Department	
Postal Address where SCMJ should be sent	
E-mail address	
Person in charge	Name: Signature:
Payment Check one of the two.	☐Bank transfer ☐Credit Card (Visa, Master)
Name of Associate Membership	1.
	2.

Call for Individual Members

We call for individual members. The privileges to them and the membership dues are shown in "**Join ISMS !**" on the inside of the back cover.

Items required in Membership Application Form

1. Name
2. Birth date
3. Academic background
4. Affiliation
5. 4's address
6. Doctorate
7. Contact address
8. E-mail address
9. Special fields
10. Membership category (See **Table 1** in "**Join ISMS !**")

Individual Membership Application Form

1. **Name**	
2. **Birth date**	
3. **Academic background**	
4. **Affiliation**	
5. **4's address**	
6. **Doctorate**	
7. **Contact address**	
8. **E-mail address**	
9. **Special fields**	
10. **Membership category**	

Contributions (Gift to the ISMS)

We deeply appreciate your generous contributions to support the activities of our society.

The donation are used (1) to make medals for the new prizes (Kitagawa Prize, Kunugi Prize, and ISMS Prize), (2) to support the IVMS at Osaka University Nakanoshima Center, and (3) for a special fund designated by the contributors.

Your remittance to the following accounts of ours will be very much appreciated.

(1) Through a post office, remit to our giro account (in Yen only):
No. 00930-1-11872, Japanese Association of Mathematical Sciences (JAMS)
or send International Postal Money Order (in US Dollar or in Yen) to our address:
International Society for Mathematical Sciences
2-1-18 Minami Hanadaguchi, Sakai-ku, Sakai, Osaka 590-0075, Japan

(2) A/C 94103518, ISMS
CITIBANK, Japan Ltd., Shinsaibashi Branch
Midosuji Diamond Building
2-1-2 Nishi Shinsaibashi, Chuo-ku, Osaka 542-0086, Japan

****

Payment Instructions:

Payment can be made through a post office or a bank, or by credit card. Members may choose the most convenient way of remittance. Please note that we do not accept payment by bank drafts (checks). For more information, please refer to an invoice.

Methods of Overseas Payment:

Payment can be made through (1) a post office, (2) a bank, (3) by credit card, or (4) UNESCO Coupons.

Authors or members may choose the most convenient way of remittance as are shown below. Please note that **we do not accept payment by bank drafts (checks)**.

(1) Remittance through a post office to our giro account No. 00930-1-11872 or send International Postal Money Order to our postal address (2) Remittance through a bank to our account No. 94103518 at Shinsaibashi Branch of CITIBANK (3) **Payment by credit cards** (AMEX, VISA, MASTER or NICOS), or (4) Payment by UNESCO Coupons.

Methods of Domestic Payment:

Make remittance to:

(1) Post Office Transfer Account - 00930-3-73982 or

(2) Account No.7726251 at Sakai Branch, SUMITOMO MITSUI BANKING CORPORATION, Sakai, Osaka, Japan.

All of the correspondences concerning subscriptions, back numbers, individual and institutional memberships, should be addressed to the Publications Department, International Society for Mathematical Sciences.

Join ISMS !

ISMS Publications: We published **Mathematica Japonica (M.J.)** in print, which was first published in 1948 and has gained an international reputation in about sixty years, and its offshoot **Scientiae Mathematicae (SCM)** both online and in print. In January 2001, the two publications were unified and changed to **Scientiae Mathematicae Japonicae (SCMJ)**, which is the "21st Century New Unified Series of Mathematica Japonica and Scientiae Mathematicae" and published both online and in print. Ahead of this, the online version of SCMJ was first published in September 2000. The whole number of SCMJ exceeds 270, which is the largest amount in the publications of mathematical sciences in Japan. The features of SCMJ are:

1) About 80 eminent professors and researchers of not only Japan but also 20 foreign countries join the Editorial Board. The accepted papers are published both online and in print. SCMJ is reviewed by Mathematical Review and Zentralblatt from cover to cover.
2) SCMJ is distributed to many libraries of the world. The papers in SCMJ are introduced to the relevant research groups for the positive exchanges between researchers.
3) **ISMS Annual Meeting**: Many researchers of ISMS members and non-members gather and take time to make presentations and discussions in their research groups every year.

The privileges to the individual ISMS Members:

(1) No publication charges
(2) Free access (**including printing out**) to the online version of SCMJ
(3) Free copy of each printed issue

The privileges to the Institutional Members:

Two associate members can be registered, free of charge, from an institution.

Table 1: Membership Dues for 2015

Categories	Domestic	Overseas	Developing countries
1-year Regular member	¥8,000	US$80 , Euro75	US$50, Euro47
1-year Students member	¥4,000	US$50 , Euro47	US$30 , Euro28
Life member*	Calculated as below*	US$750 , Euro710	US$440, Euro416
Honorary member	Free	Free	Free

(Regarding submitted papers,we apply above presented new fee after April 15 in 2015 on registoration date.) * Regular member between 63 - 73 years old can apply the category.

(73 − age) × ¥3,000

Regular member over 73 years old can maintain the qualification and the privileges of the ISMS members, if they wish.

Categories of 3-year members were abolished.

CONTENTS

Notices from the ISMS