Publ. Math. Debrecen
Manuscript (January 31, 2022)

Mercer type inequalities for normalised isotonic linear functionals with applications

By László Horváth

Abstract. In this paper we give new Mercer type inequalities for normalised isotonic linear functionals which contain Niezgoda's inequality as a very special case. We deal with some particular forms of the obtained inequalities and study some refinements of them. The results are applied to means generated by normalised isotonic linear functionals. As another application we extend Mercer's inequality to an operator inequality for convex (not operator convex) functions. An unusual feature of this result is to use closed normal subalgebras instead of a single operator.

1. Introduction

A function $f : C \to \mathbb{R}$ defined on an interval $C \subset \mathbb{R}$ is said to be convex if the inequality

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

holds for any $x, y \in C$ and for every $t \in [0,1]$.

Let the set I denote either $\{1, \ldots, n\}$ for some $n \geq 1$ or $\mathbb{N}_+ := \{1, 2, \ldots\}$. We say that the numbers $(p_i)_{i \in I}$ represent a discrete probability distribution if $p_i \geq 0$ $(i \in I)$ and $\sum\limits_{i \in I} p_i = 1$.

Perhaps the most useful inequalities for convex functions are the different types of Jensen's inequalities. The functional form of Jensen's inequality was

---

Mathematics Subject Classification: 39B62, 26A51, 26D15.
Key words and phrases: Jensen-Mercer's inequality, normalised isotonic linear functional, mean, operator inequality.
The research was supported by Hungarian Scientific Research Fund (OTKA) Grant No. K139346.

given by Jessen [6]. Jessen's inequality can be formulated by linear functionals satisfying the following properties:

(C_1) Let E be a nonempty set, and let L be a subspace of the vector space of real valued functions defined on E. It is also assumed that $1_E \in L$ (for every $c \in \mathbb{R}$ the constant function $c_E : E \to \mathbb{R}$ is defined by $c_E(x) = c$).

(C_2) Let $A : L \to \mathbb{R}$ be a linear functional.

(C_3) Assume A is nonnegative that is $A(\varphi) \geq 0$ for all nonnegative $\varphi \in L$.

(C_4) Assume $A(1_E) = 1$.

Linear functionals satisfying (C_3) are often called isotonic (or monotonic). It is said that a linear functional is normalised (or unital) if (C_4) holds.

The closure of a subset H of $\mathbb{R}$ is denoted by $\overline{H}$.

Theorem 1. *(Jessen's inequality, see [5]) Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Let f be a convex function on the interval $C \subset \mathbb{R}$, and let $\varphi \in L$ such that $\varphi(x) \in C$ for all $x \in E$. Then*

(a) $A(\varphi) \in \overline{C}$.

(b) If $A(\varphi) \in C$, $f \circ \varphi \in L$, and f is continuous at $A(\varphi)$, then

$$f(A(\varphi)) \leq A(f \circ \varphi).$$

The previous result is often referred to as Jensen's inequality for isotonic linear functionals.

Mercer [9] established an interesting variant of the discrete Jensen inequality, namely:

Theorem 2. *(Jensen-Mercer's inequality) If C is an interval, $f : C \to \mathbb{R}$ is a convex function, $p_1, \ldots, p_n$ represent a discrete probability distribution, and $x_1, \ldots, x_n \in [a, b] \subset C$, then*

$$f\left(a + b - \sum_{i=1}^{n} p_i x_i\right) \leq f(a) + f(b) - \sum_{i=1}^{n} p_i f(x_i).$$

Niezgoda [13] extended Theorem 2, and the principal tool in his treatment is majorization. His result is the next:

Theorem 3. *(Niezgoda's inequality) Let $f : C \to \mathbb{R}$ be a continuous convex function on interval $C \subset \mathbb{R}$. Suppose $\mathbf{a} = (a_1, \ldots, a_m)$ with $a_j \in C$, and $\mathbf{X} = (x_{ij})$ is a real $n \times m$ matrix such that $x_{ij} \in C$ for all i, j. If $\mathbf{a}$ majorizes each row of $\mathbf{X}$, that is*

$$\mathbf{x}_i = (x_{i1}, \ldots, x_{im}) \prec (a_1, \ldots, a_m) = \mathbf{a} \text{ for each } i = 1, \ldots, n,$$

then we have the inequality

$$f\left(\sum_{j=1}^{m}a_j-\sum_{j=1}^{m-1}\sum_{i=1}^{n}p_ix_{ij}\right)\leq\sum_{j=1}^{m}f\left(a_j\right)-\sum_{j=1}^{m-1}\sum_{i=1}^{n}p_if\left(x_{ij}\right),$$

where $p_1,\ldots,p_n$ represent a discrete probability distribution.

The following functional version of Jensen-Mercer's inequality is due to Cheung, Matković and Pečarić [2].

Theorem 4. *Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Let f be a continuous convex function on the interval $C\subset\mathbb{R}$. If $\varphi\in L$ such that $\varphi(x)\in[a,b]\subset C$ for all $x\in E$, $f\circ\varphi\in L$ and $f\circ(a+b-\varphi)\in L$, then*

$$f\left(a+b-A\left(\varphi\right)\right)\leq A\left(f\left(a+b-\varphi\right)\right)$$

$$\leq\frac{b-A\left(\varphi\right)}{b-a}f\left(b\right)+\frac{A\left(\varphi\right)-a}{b-a}f\left(a\right)\leq f\left(a\right)+f\left(b\right)-A\left(f\circ\varphi\right).$$

The main goal of this paper to give a generalization of Theorem 4 which contains Niezgoda's inequality as a very special case. We deal with some particular forms of the obtained inequalities and study some refinements of them. We first apply the results to means generated by normalised isotonic linear functionals. As another application we extend Mercer's inequality to an operator inequality for convex (not operator convex) functions. An unusual feature of this result is to use closed normal subalgebras instead of a single operator.

2. Preliminary results

We introduce a majorization relation for finite sequences of real numbers (see Marshall and Olkin [7]).

Definition 5. *Let $C\subset\mathbb{R}$ be an interval. We say that $\mathbf{y}:=(y_1,\ldots,y_n)\in C^n$ majorizes $\mathbf{x}:=(x_1,\ldots,x_n)\in C^n$, written $\mathbf{y}\succ\mathbf{x}$, if*

$$\sum_{i=1}^{k}y_{[i]}\geq\sum_{i=1}^{k}x_{[i]},\quad k=1,\ldots,n-1 \text{ and } \sum_{i=1}^{n}y_{[i]}=\sum_{i=1}^{n}x_{[i]},$$

where $x_{[1]}\geq x_{[2]}\geq\ldots\geq x_{[n]}$ and $y_{[1]}\geq y_{[2]}\geq\ldots\geq y_{[n]}$ are the entries of $\mathbf{x}$ and $\mathbf{y}$, respectively, in decreasing order.

The following two classical results are associated to majorization theory.

Theorem 6. *(weighted Hardy-Littlewod-Pólya inequality, see* [12]*) Let* $C \subset \mathbb{R}$ *be an interval, and let* $f : C \to \mathbb{R}$ *be a convex function. If* $(x_1, \ldots, x_n) \in C^n$, $(y_1, \ldots, y_n) \in C^n$ *and* $q_1, \ldots, q_n$ *are nonnegative numbers such that*

(a) $x_1 \geq \ldots \geq x_n$,

(b) $\sum\limits_{k=1}^{r} q_k x_k \leq \sum\limits_{k=1}^{r} q_k y_k$ $(r = 1, \ldots, n-1)$,

(c) $\sum\limits_{k=1}^{n} q_k x_k = \sum\limits_{k=1}^{n} q_k y_k$,

then

$$\sum_{i=1}^{n} q_i f(x_i) \leq \sum_{i=1}^{n} q_i f(y_i).$$

We stress that the nonnegativity of the numbers $q_1, \ldots, q_n$ cannot be omitted in Theorem 6. The so-called majorisation inequality is the special case of the previous result where $\mathbf{y} \succ \mathbf{x}$ and $q_1 = \ldots = q_r = 1$.

Theorem 7. *(Fuchs' inequality, see* [3]*) Let* $C \subset \mathbb{R}$ *be an interval, and let* $f : C \to \mathbb{R}$ *be a convex function. If* $(x_1, \ldots, x_n) \in C^n$, $(y_1, \ldots, y_n) \in C^n$ *and* $q_1, \ldots, q_n$ *are real numbers such that*

(a) $x_1 \geq \ldots \geq x_n$ and $y_1 \geq \ldots \geq y_n$,

(b) $\sum\limits_{k=1}^{r} q_k x_k \leq \sum\limits_{k=1}^{r} q_k y_k$ $(r = 1, \ldots, n-1)$,

(c) $\sum\limits_{k=1}^{n} q_k x_k = \sum\limits_{k=1}^{n} q_k y_k$,

then

$$\sum_{i=1}^{n} q_i f(x_i) \leq \sum_{i=1}^{n} q_i f(y_i).$$

A refinement of the Jessen's inequality will be used from the paper Horváth [5]. To formulate this we need the following hypotheses:

(H_1) Assume a normalised isotonic linear functional $A : L \to \mathbb{R}$ is given, that is (C_1-C_4) are satisfied. Further, it is assumed that for all $\varphi \in L$ the function $|\varphi|$ also belongs to L (in this case L is a Stone vector lattice).

(H_2) Let the index set I denote either $\{1, \ldots, n\}$ for some $n \geq 1$ or $\mathbb{N}_+$. Let the index set J denote either $\{1, \ldots, k\}$ for some $k \geq 1$ or $\mathbb{N}_+$.

(H_3) Let $(\lambda_j)_{j \in J}$ represent a positive probability distribution which means that $\lambda_j > 0$ $(j \in J)$, and $\sum\limits_{j \in J} \lambda_j = 1$. For each $j \in J$ let π_j be a permutation of the set I (a permutation π of I refers to a bijection from I onto itself).

(H_4) Suppose we are given a sequence $\mathfrak{A}_I = (A_i)_{i\in I}$ of isotonic linear functionals $A_i : L \to \mathbb{R}$ with $A_i(1_E) > 0$ for all $i \in I$ and $\sum\limits_{i\in I} A_i = A$.

The following result, in a more general form, is Theorem 4.1 in [5]. The phrase closed interval means an interval in $\mathbb{R}$ which is a closed (not necessarily compact) set.

Theorem 8. *Assume (H_1-H_4). Let $C \subset \mathbb{R}$ be a closed interval, and $f : C \to \mathbb{R}$ be a continuous convex function. Let $\varphi \in L$ taking values in C such that $f \circ \varphi \in L$. Then*

$$f(A(\varphi)) \le C_{funct} = C_{funct}(\varphi, f, \lambda, \pi, \mathfrak{A}_I)$$

$$:= \sum_{i\in I}\left(\sum_{j\in J}\lambda_j A_{\pi_j(i)}(1_E)\right) f\left(\frac{\sum\limits_{j\in J}\lambda_j A_{\pi_j(i)}(\varphi)}{\sum\limits_{j\in J}\lambda_j A_{\pi_j(i)}(1_E)}\right) \le A(f\circ\varphi).$$

3. Main results

We first establish an extension of Theorem 4. The result is also a generalization of Theorem 3 to normalised isotonic linear functionals.

Theorem 9. *Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Let $C \subset \mathbb{R}$ be an interval, and let $f : C \to \mathbb{R}$ be a convex function. Assume further that $m \ge 2$ is an integer,*

$$\varphi_k, \psi_k \in L, \quad k = 1, \ldots, m$$

such that $\varphi_k(x)$, $\psi_k(x) \in C$ for all $x \in E$ ($k = 1, \ldots, m$), $A(\varphi_m) \in C$, f is continuous at $A(\varphi_m)$,

$$f\circ\varphi_k \in L, \quad f\circ\psi_k \in L, \quad k = 1, \ldots, m, \tag{1}$$

and
either

(a_1) $\varphi_1 \ge \varphi_2 \ge \ldots \ge \varphi_m$ and $\psi_1 \ge \psi_2 \ge \ldots \ge \psi_m$,
(b_1) $q_k \in \mathbb{R}$ ($k = 1, \ldots, m-1$), $q_m > 0$,
(c_1) $\sum\limits_{k=1}^{r} q_k\varphi_k \le \sum\limits_{k=1}^{r} q_k\psi_k$ ($r = 1, \ldots, m-1$),
(d_1) $\sum\limits_{k=1}^{m} q_k\varphi_k = \sum\limits_{k=1}^{m} q_k\psi_k$,

or

(a$_2$) $\varphi_1 \geq \varphi_2 \geq \ldots \geq \varphi_m$,
(b$_2$) $q_k \geq 0$ $(k = 1, \ldots, m-1)$, $q_m > 0$,
(c$_2$) $\sum\limits_{k=1}^{r} q_k \varphi_k \leq \sum\limits_{k=1}^{r} q_k \psi_k$ $(r = 1, \ldots, m-1)$,
(d$_2$) $\sum\limits_{k=1}^{m} q_k \varphi_k = \sum\limits_{k=1}^{m} q_k \psi_k$,

or

(a$_3$) $q_k = 1$ $(k = 1, \ldots, m)$,
(b$_3$) $(\varphi_1(x), \ldots, \varphi_m(x)) \prec (\psi_1(x), \ldots, \psi_m(x))$ $(x \in E)$.
Then

$$f\left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k A(\psi_k) - \sum_{k=1}^{m-1} q_k A(\varphi_k)\right)\right)$$

$$\leq A\left(f \circ \left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k \psi_k - \sum_{k=1}^{m-1} q_k \varphi_k\right)\right)\right) \tag{2}$$

$$\leq \frac{1}{q_m}\left(\sum_{k=1}^{m} q_k A(f \circ \psi_k) - \sum_{k=1}^{m-1} q_k A(f \circ \varphi_k)\right). \tag{3}$$

Proof. Suppose conditions (a$_1$-d$_1$) are satisfied.
By (d$_1$),

$$\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k \psi_k - \sum_{k=1}^{m-1} q_k \varphi_k\right) = \varphi_m, \tag{4}$$

and hence the linearity of A implies that

$$\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k A(\psi_k) - \sum_{k=1}^{m-1} q_k A(\varphi_k)\right)$$

$$= \frac{1}{q_m} A\left(\sum_{k=1}^{m} q_k \psi_k - \sum_{k=1}^{m-1} q_k \varphi_k\right) = A(\varphi_m).$$

Since A is a normalised isotonic linear functional, $\varphi_m \in L$ such that $\varphi_m(x) \in C$ for all $x \in E$, $A(\varphi_m) \in C$, $f \circ \varphi_m \in L$, and f is continuous at $A(\varphi_m)$, inequality (2) is an immediate consequence of Jessen's inequality.

To prove (3), first we use (a$_1$-d$_1$), together with $\psi_k(x)$, $\varphi_k(x) \in C$ for all $x \in E$ $(k = 1, \ldots, m)$ and the convexity of f, to apply Fuchs' inequality: it gives that

$$\sum_{k=1}^{m} q_k f(\varphi_k(x)) \leq \sum_{k=1}^{m} q_k f(\psi_k(x)), \quad x \in E.$$

By (4), it follows from this that

$$\sum_{k=1}^{m-1} q_k f\circ\varphi_k + q_m f\circ\left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k\psi_k - \sum_{k=1}^{m-1} q_k\varphi_k\right)\right) \le \sum_{k=1}^{m} q_k f\circ\psi_k. \quad (5)$$

Since (1) is satisfied and A is an isotonic linear functional, we have from (5) that

$$q_m A\left(f\circ\left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k\psi_k - \sum_{k=1}^{m-1} q_k\varphi_k\right)\right)\right)$$
$$\le \sum_{k=1}^{m} q_k A\left(f\circ\psi_k\right) - \sum_{k=1}^{m-1} q_k A\left(f\circ\varphi_k\right),$$

and therefore inequality (3) follows by using $q_m > 0$.

Under the other two groups of conditions we can prove similarly by using either Hardy-Littlewod-Pólya inequality or majorization inequality instead of Fuchs' inequality.

The proof is complete. □

The following simple result shows that Theorem 9 extends inequalities

$$f\left(a+b-A\left(\varphi\right)\right) \le A\left(f\left(a+b-\varphi\right)\right) \le f\left(a\right)+f\left(b\right)-A\left(f\circ\varphi\right)$$

in Theorem 4.

Corollary 10. *Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Let $C\subset\mathbb{R}$ be an interval, and let ψ_1, ψ_2, $\varphi\in L$ such that $\psi_1(x)$, $\psi_2(x)\in C$ for all $x\in E$, $\psi_1\le\varphi\le\psi_2$ and $A(\psi_1+\psi_2-\varphi)\in C$. If $f: C\to\mathbb{R}$ is a convex function for which f is continuous at $A(\psi_1+\psi_2-\varphi)$ and $f\circ\psi_1$, $f\circ\psi_2$, $f\circ\varphi$ and $f\circ(\psi_1+\psi_2-\varphi)$ belong to L, then*

$$f\left(A\left(\psi_1\right)+A\left(\psi_2\right)-A\left(\varphi\right)\right) \le A\left(f\circ\left(\psi_1+\psi_2-\varphi\right)\right)$$
$$\le A\left(f\circ\psi_1\right)+A\left(f\circ\psi_2\right)-A\left(f\circ\varphi\right). \quad (6)$$

Proof. The condition $\psi_1\le\varphi\le\psi_2$ implies

$$\left(\varphi\left(x\right),\left(\psi_1+\psi_2-\varphi\right)\left(x\right)\right) \prec \left(\psi_1\left(x\right),\psi_2\left(x\right)\right), \quad x\in E,$$

and therefore we can apply Theorem 9 by choosing $\varphi_1 := \varphi$ and $\varphi_2 := \psi_1+\psi_2-\varphi$ (conditions (a_3-b_3) are satisfied).

The proof is complete. □

Remark 11. *Niezgoda's inequality is contained in Theorem 9 as a very special case: Really, let $f : C \to \mathbb{R}$ be a continuous convex function on interval $C \subset \mathbb{R}$. Suppose $n \geq 1$ and $m \geq 2$ are integers, $\mathbf{a} = (a_1, \dots, a_m)$ with $a_k \in C$, and $\mathbf{X} = (x_{ik})$ is a real $n \times m$ matrix such that $x_{ik} \in C$ for all i, k. Let $p_1, \dots, p_n$ represent a discrete probability distribution. Define the vector space L by*

$$L := \left\{ (y_i)_{i=1}^n \mid y_i \in \mathbb{R}, \quad i = 1, \dots, n \right\},$$

the functions φ_k, $\psi_k \in L$ $(k = 1, \dots, m)$ by

$$\psi_k(i) := a_k, \quad i = 1, \dots, n, \quad k = 1, \dots, m$$

and

$$\varphi_k(i) := x_{ik}, \quad i = 1, \dots, n, \quad k = 1, \dots, m,$$

and the normalised isotonic linear functional $A : L \to \mathbb{R}$ by

$$A\left((y_i)_{i=1}^n\right) := \sum_{i=1}^n p_i y_i.$$

If $\mathbf{a}$ *majorizes each row of* $\mathbf{X}$*, then Theorem 9 gives Niezgoda's inequality.*

It is worth mentioning the form of Corollary 10 when the normalised isotonic linear functional A is defined by integral. In this case weaker conditions are sufficient. An integral version of Theorem 9 can be found in Horváth [4].

Theorem 12. *Let $(E, \mathcal{A}, \mu)$ be a probability space, $C \subset \mathbb{R}$ be an interval, ψ_1, $\psi_2 : E \to C$ be μ-integrable functions, and let $\varphi : E \to \mathbb{R}$ be a measurable function such that $\psi_1 \leq \varphi \leq \psi_2$. If $f : C \to \mathbb{R}$ is a convex function for which $f \circ \psi_1$ and $f \circ \psi_2$ are μ-integrable, then*

$$f\left(\int_E (\psi_1 + \psi_2 - \varphi)\, d\mu\right) \leq \int_E f \circ (\psi_1 + \psi_2 - \varphi)\, d\mu \tag{7}$$

$$\leq \int_E f \circ \psi_1 d\mu + \int_E f \circ \psi_2 d\mu - \int_E f \circ \varphi d\mu. \tag{8}$$

Proof. If L means the vector space of real integrable functions on E and the linear functional A is defined on L by

$$A(\chi) := \int_E \chi d\mu,$$

then A is obviously isotonic and normalised.

It is obvious that $\varphi(x) \in C$ for all $x \in E$ and φ is also μ-integrable.

If f is either increasing or decreasing, then either

$$f \circ \psi_1 \leq f \circ \varphi \leq f \circ \psi_2$$

or

$$f \circ \psi_1 \geq f \circ \varphi \geq f \circ \psi_2.$$

Since $f \circ \varphi$ is measurable and $f \circ \psi_1$ and $f \circ \psi_2$ are μ-integrable, it follows from the previous inequalities that $f \circ \varphi$ is μ-integrable.

If f is not monotonic, then there exists an inner point $c \in C$ such that f is decreasing on $]-\infty, c] \cap C$ and increasing on $[c, \infty[\cap C$. In this case

$$f(c) \leq f \circ \varphi \leq \max(f \circ \psi_1, f \circ \psi_2),$$

and therefore $f \circ \varphi$ is μ-integrable too.

Since

$$\psi_1 \leq \psi_1 + \psi_2 - \varphi \leq \psi_2,$$

the roles of φ and $\psi_1 + \psi_2 - \varphi$ can be interchanged in the previous argument and hence $f \circ (\psi_1 + \psi_2 - \varphi)$ is also μ-integrable.

Inequality (7) now follows immediately from the integral Jensen's inequality.

Inequality (8) can be obtained from (6).

The proof is complete. □

An interesting special case of Theorem 12 is the following. We emphasize that the index set I can also be a countably infinite set.

Corollary 13. *Let the index set I denote either $\{1, \ldots, n\}$ for some $n \geq 1$ or $\mathbb{N}_+$, and let $(p_i)_{i \in I}$ represent a discrete probability distribution. Let $C \subset \mathbb{R}$ be an interval, and let $(a_i)_{i \in I}$, $(b_i)_{i \in I}$, $(x_i)_{i \in I}$ be sequences from C such that $a_i \leq x_i \leq b_i$ for all $i \in I$ and the series $\sum_{i \in I} p_i a_i$ and $\sum_{i \in I} p_i b_i$ are absolutely convergent. If $f : C \to \mathbb{R}$ is a convex function for which the series $\sum_{i \in I} p_i f(a_i)$ and $\sum_{i \in I} p_i f(b_i)$ are absolutely convergent, then*

$$f\left(\sum_{i \in I} p_i (a_i + b_i - x_i)\right) \leq \sum_{i \in I} p_i f(a_i + b_i - x_i)$$

$$\leq \sum_{i \in I} p_i f(a_i) + \sum_{i \in I} p_i f(b_i) - \sum_{i \in I} p_i f(x_i).$$

Proof. Let $E := I$, $\mathcal{A}$ be the set of all subsets of E, and let $\mu := \sum_{i\in I} p_i \varepsilon_i$ where $\varepsilon_i : \mathcal{A} \to \mathbb{R}$ is the unit mass at i $(i \in I)$. Define the functions ψ_1, ψ_2, $\varphi : E \to C$ by

$$\psi_1(i) := a_i, \quad \psi_2(i) := b_i, \quad \varphi(i) := x_i.$$

Now the result is an immediate consequence of Theorem 12 by applying it to the probability space $(E, \mathcal{A}, \mu)$ and to the functions ψ_1, ψ_2, and φ.

The proof is complete. □

Jensen-Mercer's inequality has a lot of different refinements, see e.g. the recent papers Horváth [4] and Moradi and Furuichi [11]. Now, we give some refinements of Theorem 9. There are two ways to achieve this: one is to refine inequality 2, while the other is to refine inequality 3. The proof of inequality 2 shows that in the first case any refinement of the Jessen's inequality can be applied, but in the second case, specific methods are needed. For a more detailed analysis of the problem, see the paper Horváth [4]. Both cases are illustrated below.

Theorem 14. *Assume (H_1-H_4). Let $C \subset \mathbb{R}$ be a closed interval, and $f : C \to \mathbb{R}$ be a continuous convex function. Assume further that $m \geq 2$ is an integer,*

$$\varphi_k, \psi_k \in L, \quad k = 1, \ldots, m$$

such that $\psi_k(x)$, $\varphi_k(x) \in C$ for all $x \in E$ $(k = 1, \ldots, m)$,

$$f \circ \varphi_k \in L, \quad f \circ \psi_k \in L, \quad k = 1, \ldots, m,$$

and either (a_1-d_1) or (a_2-d_2) or (a_3-b_3) is satisfied. Then

$$f\left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k A(\psi_k) - \sum_{k=1}^{m-1} q_k A(\varphi_k)\right)\right) \leq \sum_{i\in I}\left(\sum_{j\in J} \lambda_j A_{\pi_j(i)}(1_E)\right)$$

$$\cdot f\left(\frac{\frac{1}{q_m}\sum_{j\in J}\lambda_j A_{\pi_j(i)}\left(\sum_{k=1}^{m} q_k A(\psi_k) - \sum_{k=1}^{m-1} q_k A(\varphi_k)\right)}{\sum_{j\in J}\lambda_j A_{\pi_j(i)}(1_E)}\right)$$

$$\leq A\left(f \circ \left(\frac{1}{q_m}\left(\sum_{k=1}^{m} q_k \psi_k - \sum_{k=1}^{m-1} q_k \varphi_k\right)\right)\right)$$

$$\leq \frac{1}{q_m}\left(\sum_{k=1}^{m} q_k A(f \circ \psi_k) - \sum_{k=1}^{m-1} q_k A(f \circ \varphi_k)\right).$$

Proof. In the proof of Theorem 9, we have seen that inequality (2) can be obtained by applying the Jessen's inequality, so the result is a simple consequence of the Theorem 8.

The proof is complete. □

Theorem 15. *Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Assume further that L is a Stone vector lattice. Let $C \subset \mathbb{R}$ be an interval, and let ψ_1, ψ_2, $\varphi \in L$ such that $\psi_1(x)$, $\psi_2(x) \in C$ for all $x \in E$, $\psi_1 \leq \varphi \leq \psi_2$ and $A(\psi_1 + \psi_2 - \varphi) \in C$. Define the functions χ, $\omega : E \to \mathbb{R}$ by*

$$\chi(x) := \min(\varphi(x), \psi_1(x) + \psi_2(x) - \varphi(x)), \quad x \in E$$

$$\omega(x) := \max(\varphi(x), \psi_1(x) + \psi_2(x) - \varphi(x)), \quad x \in E.$$

If $f : C \to \mathbb{R}$ is a convex function for which f is continuous at $A(\psi_1 + \psi_2 - \varphi)$ and $f \circ (t\psi_1 + (1-t)\chi)$, $f \circ (t\psi_2 + (1-t)\omega)$, $f \circ \varphi$ and $f \circ (\psi_1 + \psi_2 - \varphi)$ belong to L for all $t \in [0,1]$, then

$$f(A(\psi_1) + A(\psi_2) - A(\varphi)) \leq A(f \circ (\psi_1 + \psi_2 - \varphi))$$

$$\leq A(f \circ (t\psi_1 + (1-t)\chi)) + A(f \circ (t\psi_2 + (1-t)\omega)) - A(f \circ \varphi)$$

$$\leq A(f \circ \psi_1) + A(f \circ \psi_2) - A(f \circ \varphi), \quad t \in [0,1].$$

Proof. Since L is a Stone vector lattice, χ, $\omega \in L$.

It is easy to check that

$$(\varphi(x), \psi_1(x) + \psi_2(x) - \varphi(x)) \prec (t\psi_1 + (1-t)\chi, t\psi_2 + (1-t)\omega)$$

$$\prec (\psi_1(x), \psi_2(x)), \quad x \in E, \quad t \in [0,1],$$

and hence the majorization inequality implies the second and the third inequalities.

The first inequality comes from Corollary 10.

The proof is complete. □

4. Application

In this section we first study means generated by normalised isotonic linear functionals.

The range of a function f is denoted by R_f.

Let $C\subset\mathbb{R}$ be an interval, and let $q:C\to\mathbb{R}$ be a continuous and strictly monotone function. If $(X,\mathcal{A},\mu)$ is a probability space, and $\varphi:X\to C$ is a function such that $q\circ\varphi$ is μ-integrable on X, then

$$M_q(\varphi,\mu):=q^{-1}\left(\int_X q\circ\varphi d\mu\right)$$

is called the quasi-arithmetic mean (integral q-mean) of φ.

Based on the previous notion of quasi-arithmetic mean, a new mean generated by a normalised isotonic linear functional is introduced in Beesack and Pečarić [1] and a more generalized form in Horváth [5].

Definition 16. *(see [5]) Assume (C_1-C_4) that is a normalised isotonic linear functional $A:L\to\mathbb{R}$ is given. Let $C\subset\mathbb{R}$ be an interval, $q:C\to\mathbb{R}$ be a continuous and strictly monotone function, and let $\varphi:E\to\mathbb{R}$ taking values in C such that $\overline{q(]\inf\varphi,\sup\varphi[)}\subset R_q$ and $q\circ\varphi\in L$. Define*

$$M_q(\varphi,A):=q^{-1}(A(q\circ\varphi)).$$

To simplify the next statement, we assume that the interval C is closed. The result contains Theorem 3.3 in the paper Cheung, Matković and Pečarić [2] as a special case.

Theorem 17. *Assume (C_1-C_4) that is a normalised isotonic linear functional A is given. Let $C\subset\mathbb{R}$ be a closed interval, and let q, $r:C\to\mathbb{R}$ be continuous and strictly monotone functions. Assume further that ψ_1, ψ_2, $\varphi:E\to\mathbb{R}$ such that $\psi_1(x)$, $\psi_2(x)\in C$ for all $x\in E$, $\psi_1\le\varphi\le\psi_2$, and*

$$q\circ\varphi,\in L\quad q\circ\psi_k\in L,\quad r\circ\varphi\in L,\quad r\circ\psi_k\in L,\quad k=1,2.$$

If either $q\circ r^{-1}$ is convex and q is strictly increasing, or $q\circ r^{-1}$ is concave and q is strictly decreasing, then

$$\begin{aligned}&r^{-1}(r(M_r(\psi_1,A))+r(M_r(\psi_2,A))-r(M_r(\varphi,A)))\\&\le M_q\left(r^{-1}\circ(r\circ\psi_1+r\circ\psi_2-r\circ\varphi),A\right)\\&\le q^{-1}(q(M_q(\psi_1,A))+q(M_q(\psi_2,A))-q(M_q(\varphi,A))),\end{aligned}$$

while if either $r\circ q^{-1}$ is convex and r is strictly decreasing, or $r\circ q^{-1}$ is concave and r is strictly increasing, then

$$\begin{aligned}&q^{-1}(q(M_q(\psi_1,A))+q(M_q(\psi_2,A))-q(M_q(\varphi,A)))\\&\le M_r\left(q^{-1}\circ(q\circ\psi_1+q\circ\psi_2-q\circ\varphi),A\right)\\&\le r^{-1}(r(M_r(\psi_1,A))+r(M_r(\psi_2,A))-r(M_r(\varphi,A))).\end{aligned}$$

Proof. Assume $q \circ r^{-1}$ is convex and q is strictly increasing.

The function r is strictly monotone, and therefore either $r\circ\psi_1 \le r\circ\varphi \le r\circ\psi_2$ or $r \circ \psi_2 \le r \circ \varphi \le r \circ \psi_1$. In both cases we can apply Corollary 10 which implies

$$q \circ r^{-1}\left(A\left(r \circ \psi_1\right) + A\left(r \circ \psi_2\right) - A\left(r \circ \varphi\right)\right)$$

$$\le A\left(q \circ r^{-1} \circ \left(r \circ \psi_1 + r \circ \psi_2 - r \circ \varphi\right)\right)$$

$$\le A\left(q \circ r^{-1} \circ r \circ \psi_1\right) + A\left(q \circ r^{-1} \circ r \circ \psi_2\right) - A\left(q \circ r^{-1} \circ r \circ \varphi\right)$$

$$= A\left(q \circ \psi_1\right) + A\left(q \circ \psi_2\right) - A\left(q \circ \varphi\right).$$

Since q is strictly increasing, we have that

$$r^{-1}\left(A\left(r \circ \psi_1\right) + A\left(r \circ \psi_2\right) - A\left(r \circ \varphi\right)\right)$$

$$\le q^{-1} A\left(q \circ r^{-1} \circ \left(r \circ \psi_1 + r \circ \psi_2 - r \circ \varphi\right)\right)$$

$$\le q^{-1}\left(A\left(q \circ \psi_1\right) + A\left(q \circ \psi_2\right) - A\left(q \circ \varphi\right)\right),$$

and the result follows from this.

The other cases can be investigated in a similar way.

The proof is complete. □

In the second part of this section $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ always denotes a complex Hilbert space. The Banach algebra of all bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})$. The operator I means the identity operator on $\mathcal{H}$. The spectrum of an operator T is denoted by $\sigma(T)$. If $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint and $\lambda \in \sigma(T)$, then $\lambda \in \mathbb{R}$. For an interval $C \subset \mathbb{R}$, $S(C)$ means the class of all self-adjoint operators from $\mathcal{B}(\mathcal{H})$ whose spectra are contained in C.

Now, following Rudin [14] mainly, we briefly summarize the spectral theory and the symbolic calculus for normal operators in $\mathcal{B}(\mathcal{H})$.

A subset N of $\mathcal{B}(\mathcal{H})$ is called normal if N commutes and $T^* \in N$ whenever $T \in N$.

Theorem 18. *(Spectral theorem for closed normal subalgebras of $\mathcal{B}(\mathcal{H})$, see [14]) Assume N is a closed normal subalgebra of $\mathcal{B}(\mathcal{H})$ which contains I, and let $\triangle$ be the maximal ideal space of N. Then*

(a) There exists a unique resolution E of the identity on the Borel subsets of $\triangle$ which satisfies

$$T = \int_{\triangle} \widehat{T}\, dE$$

for every $T \in N$, where $\widehat{T}$ is the Gelfand transform of T.

(b) The inverse of the Gelfand transform extends to an isometric *-isomorphism Φ of the algebra $L^\infty(E)$ onto a closed subalgebra M of $\mathcal{B}(\mathcal{H})$, $N \subset M$, given by

$$\Phi(f) = \int_\triangle f dE, \quad f \in L^\infty(E). \tag{9}$$

Explicitly, Φ is linear and multiplicative and satisfies

$$\Phi\left(\overline{f}\right) = \Phi(f)^*, \quad \|\Phi(f)\| = \|f\|_\infty .$$

The integral (9) is the abbreviation for

$$\langle \Phi(f)\, x, y\rangle = \int_\triangle f dE_{x,y}, \quad x, y \in \mathcal{H},$$

where $E_{x,y}$ denotes the complex measure

$$E_{x,y}(\omega) := \langle E(\omega)\, x, y\rangle$$

on the Borel subsets of $\triangle$. If $x \in \mathcal{H}$ and $\|x\| = 1$, then $E_{x,x}$ is a probability measure.

It follows from the previous theorem that for every normal operator $T \in \mathcal{B}(\mathcal{H})$ there exists a unique resolution E^T of the identity (called the spectral decomposition of T) on the Borel subsets of $\sigma(T)$ which satisfies

$$T = \int_{\sigma(T)} \lambda dE^T(\lambda).$$

By using E^T, for every bounded Borel function $f : \sigma(T) \to \mathbb{C}$ we can define the operator

$$\int_{\sigma(T)} f dE^T$$

which is denoted by $f(T)$ as usual.

Assume N is a closed normal subalgebra of $\mathcal{B}(\mathcal{H})$ which contains I, and let E be the corresponding resolution of the identity.

If $T \in N$, then T has the same spectra with respect to N and $\mathcal{B}(\mathcal{H})$. It is easy to check that for every $T \in N$ the spectral decomposition E^T of T is the image of E under the mapping $\widehat{T}$, that is

$$E^T(\varpi) = E\left(\widehat{T}^{-1}(\varpi)\right)$$

for every Borel subsets ϖ of $\sigma(T)$. It follows that for every bounded Borel function $f : \sigma(T) \to \mathbb{C}$

$$f(T) = \int\limits_{\triangle} f \circ \widehat{T} dE.$$

Let $T \in N$ be a positive operator. Since the range of $\widehat{T}$ is $\sigma(T)$, $\widehat{T} \geq 0$. It follows that if T_1, $T_2 \in N$ are self-adjoint operators such that $T_1 \leq T_2$, then $\widehat{T_1} \leq \widehat{T_2}$.

Now we give an operator version of Mercer's inequality for convex functions.

Theorem 19. *Let a, $b \in \mathbb{R}$ with $a < b$, and let T_1, T_2, $T \in \mathcal{B}(\mathcal{H})$ such that they commute with each other, T_1, $T_2 \in S([a,b])$ and $T_1 \leq T \leq T_2$. If $f : [a,b] \to \mathbb{R}$ is a convex function and $x \in \mathcal{H}$ with $\|x\| = 1$, then*

$$f(\langle T_1 x, x\rangle + \langle T_2 x, x\rangle - \langle Tx, x\rangle) \leq \langle f(T_1 + T_2 - T)x, x\rangle \tag{10}$$
$$\leq \langle f(T_1)x, x\rangle + \langle f(T_2)x, x\rangle - \langle f(T)x, x\rangle .$$

Proof. Since $T_1, T_2 \in S([a,b])$ and $T_1 \leq T \leq T_2$, $T \in S([a,b])$ too.

Let N be the smallest closed subalgebra of $\mathcal{B}(\mathcal{H})$ that contains I, T_1, T_2, and T (this is the closure of the set of polynomials in T_1, T_2, and T), and let E be the corresponding resolution of the identity. Since T_1, T_2, and T are self-adjoint and they commute with each other, N is a normal subalgebra of $\mathcal{B}(\mathcal{H})$.

By Theorem 18,

$$f(\langle T_1 x, x\rangle + \langle T_2 x, x\rangle - \langle Tx, x\rangle) = f((T_1 + T_2 - T)x, x)$$
$$= f\left(\int\limits_{\triangle} T_1 \widehat{+ T_2 -} T dE_{x,x}\right) = f\left(\int\limits_{\triangle} \left(\widehat{T_1} + \widehat{T_2} - \widehat{T}\right) dE_{x,x}\right).$$

Since $\widehat{T_1} \leq \widehat{T} \leq \widehat{T_2}$ and they are continuous functions on the Gelfand topology of $\triangle$, Theorem 12 can be applied and we obtain that

$$f\left(\int\limits_{\triangle} \left(\widehat{T_1} + \widehat{T_2} - \widehat{T}\right) dE_{x,x}\right) \leq \int\limits_{\triangle} f\left(\widehat{T_1} + \widehat{T_2} - \widehat{T}\right) dE_{x,x}$$
$$\leq \int\limits_{\triangle} f\left(\widehat{T_1}\right) dE_{x,x} + \int\limits_{\triangle} f\left(\widehat{T_2}\right) dE_{x,x} - \int\limits_{\triangle} f\left(\widehat{T}\right) dE_{x,x}$$
$$= \langle f(T_1)x, x\rangle + \langle f(T_2)x, x\rangle - \langle f(T)x, x\rangle .$$

The proof is complete. □

Remark 20. *(a) There are generalizations of Mercer's inequality to operator inequalities (see e.g. Matković, Pečarić, Perić, [8]), but in all these results* $T_1 := aI$ *and* $T_2 := bI$.

(b) Inequality (10) is an immediate consequence of the classical operator Jensen's inequality for convex functions (see Mond, Pečarić [10]). Moreover, the commutativity of the operators T_1, T_2 *and* T *is not necessary in this case.*

Corollary 21. *Let* $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ *be a complex Hilbert space,* a, $b \in \mathbb{R}$ *with* $a < b$, *and let* $T \in S([a,b])$. *If* $f : [a,b] \to \mathbb{R}$ *is a convex function and* $x \in \mathcal{H}$ *with* $\|x\| = 1$, *then*

$$f(a + b - \langle Tx, x\rangle) \leq \langle f(aI + bI - T)x, x\rangle$$

$$\leq f(a) + f(b) - \langle f(T)x, x\rangle .$$

Proof. It is an easy consequence of Theorem 19.

The proof is complete. □

References

[1] P. R. Beesack, J. Pečarić, On Jessen's inequality for convex functions, J. Math. Anal. Appl., 110 (1985), 536-552.

[2] W. S. Cheung, A. Matković, J. Pečarić, A variant of Jessen's inequality and generalized means, J. Inequal. Pure Appl. Math. 7 (1) (2006) Article 10.

[3] L. Fuchs, A new proof of an inequality of Hardy, Littlewood and Pólya, Mat. Tidsskr. B. 1947 (1947) 53–54.

[4] L. Horváth, Some notes on Jensen-Mercer's type inequalities; extensions and refinements with applications, Math. Inequal. Appl. (accepted)

[5] L. Horváth, A combinatorial refinement of generalized Jessen's inequality, J. Convex Anal., 29 (2022), No. 1, (accepted)

[6] B. Jessen, "Bemærkninger om konvekse Funktioner og Uligheder imellem Middelværdier. I," Matematisk Tidsskrift B, 2 (1931), 17–28.

[7] A. W. Marshall, I. Olkin, Inequalities: The Theory of Majorizations and Its Applications. New York: Academic Press (1979)

[8] A. Matković, J. Pečarić, I. Perić, A variant of Jensen's inequality of Mercer's type for operators with applications, Linear Algebra Appl., 418 (2006), 551–564.

[9] A. McD. Mercer, A variant of Jensen's inequality, J. Inequal. Pure & Appl. Math. 4 (4) (2003) Article 73.

[10] B. Mond, J. Pečarić, Convex inequalities in Hilbert space, Houston J. Math. 19 (1993), 405-420.

[11] H. R. Moradi, S. Furuichi, Improvement and generalization of some Jensen-Mercer-type inequalities, J. Math. Inequal. 14 (2) (2020), 377–383.

[12] C. Niculescu, L. E. Persson: Convex functions and their applications. A contemporary approach. Springer, Berlin (2006)

[13] M. Niezgoda, A generalization of Mercer's result on convex functions, Nonlinear Anal. 71 (2009), 2771-2779.

[14] W. Rudin, Functional Analysis, McGraw-Hill, Inc. 1991.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PANNONIA
EGYETEM U. 10.
8200 VESZPRÉM, HUNGARY

E-mail: lhorvath@almos.uni-pannon.hu