Pečarić PSR Series on Mathematical Inequalities

Generalizations and Refinements of Sherman's Inequality with Applications

On Sherman's Inequality with Applications

AUTHORS

Slavica Ivelić Bradanović, Marek Niezgoda,
Đilda Pečarić, and Josip Pečarić

PSR Press
Ptolemy Scientific Research Press

EDITORS
Muhammad Imran, Ghulam Farid & Saima Akram

PSR Press
Ptolemy Scientific Research Press
Lahore, Pakistan

Generalizations and refinements of Sherman's inequality with applications

On Sherman's inequality with applications

Authors:
Slavica Ivelić Bradanović
Faculty of Civil Engineering, Architecture and Geodesy, University of Split, Split, Croatia.

Marek Niezgoda
Institute of Mathematics, Pedagogical University of Cracow (Emeritus), Cracow, Poland.

Đilda Pečarić
Department of Media and Communication, University North, Koprivnica, Croatia.

Josip Pečarić
Department of Mathematical, Physical and Chemical Sciences, Croatian Academy of Sciences and Arts, University of Zagreb, Zagreb, Croatia.

List of Authors

Slavica Ivelić Bradanović: Faculty of Civil Engineering, Architecture and Geodesy, University of Split, Split, Croatia.

Marek Niezgoda: Institute of Mathematics, Pedagogical University of Cracow (Emeritus), Cracow, Poland.

Đilda Pečarić: Department of Media and Communication, University North, Koprivnica, Croatia.

Josip Pečarić: Department of Mathematical, Physical and Chemical Sciences, Croatian Academy of Sciences and Arts, University of Zagreb, Zagreb, Croatia.

Editors

Prof. Dr. Muhammad Imran

Government College University Faisalabad, Pakistan

drmimranchaudhry@gcuf.edu.pk

Prof. Dr. Ghulam Farid

Comsats University Islamabad, Pakistan

rgfarid76@gmail.com

Prof. Dr. Saima Akram

Government College Women University Faisalabad.

saimaakram@gcwuf.edu.pk

Ptolemy Institute of Scientific Research and Technology (PISRT)

https://pisrt.org/

ISBN: 978-627-7623-07-4

First printing, June 2026

To our Teachers

Contents

Preface

Inequalities are an essential part of virtually all areas of mathematics. They are one of the most powerful tools for producing equalities, a means of solving linear programming, optimization problems, provide a way of expressing the domain of a function, of solving limits etc. Inequalities have important and useful applications in many other areas of science and engineering. In the 20th century mathematics recognized the power of mathematical inequality which has given rise to a large numbers of new results, problems and has led to new areas of mathematics. The theory of inequalities is one of the central areas of mathematical analysis through the last century and it is a fast growing discipline with increasing applications in different scientific fields. Now there is also a powerful and useful mathematical concept called majorization which is used for finding some nice and applicable inequalities. Moreover, the most important inequalities as Jensen's inequality, the Hermite-Hadamard inequality, the arithmetic-geometric mean inequality can be easily derived by using an argument based on the concept of majorization and strongly related the Schur-convex functions theory.

The concept of majorization was first studied by economists early in the twentieth century as a means for altering the unevenness of distribution of wealth or income. It has many connections with matrix theory, exactly with doubly stochastic matrices, i.e. nonnegative matrices with all rows and columns sums equal to one, which can be strongly related to graph theory and some network problems. The idea of majorization and Schur-convexity is very fundamental to many problems in linear algebra and optimization. Many complicate nonconvex constrained optimization problems involving matrix-valued variables using majorization theory might be transformed into simple problems with scalar variables that can be easily solved. The applications of majorization and various notions of stochastic majorization are extensive in mathematics especially in probability and statistics. Indeed, it has been reinvented and used in different branches of science which are closely related to mathematics (economics, information theory, quantum mechanics and entanglement transformations in quantum computation).

In the space $\mathbb{R}^n$, the term majorization is introduced to compare and detect potential links between vectors. The formal definition of the concept of majorization, that is "$\mathbf{y}$ is majorized by $\mathbf{x}$", in symbol $\mathbf{y} \prec \mathbf{x}$, is given in Definition 2.1.1.. It could be interpreted intuitively by the following sentence by Marshall, Olkin and Arnold (see [106, p. 3]):

> "There is a certain intuitive appeal to the vague notion that the components of a vector $\mathbf{x}$ are "less spread out" or "more nearly equal" than are the components of a vector $\mathbf{y}$."

This intuition is reinforced by noting the following example. Let $\mathbf{e} = (1, ..., 1)$ be the unit vector. Let us consider three vectors in $\mathbb{R}^n$ as follows. For a vector $\mathbf{x} = (x_1, ..., x_n)$ the inner product $\mathbf{x} \cdot \mathbf{e} = \sum_{i=1}^{n} x_i$ is the sum of the components of $\mathbf{x}$. Let $\bar{\mathbf{x}} = (\frac{\mathbf{x}\cdot\mathbf{e}}{n}, ..., \frac{\mathbf{x}\cdot\mathbf{e}}{n})$ and $\mathbf{x}_k = (0, ..., \mathbf{x} \cdot \mathbf{e}, ...,0)$ be the vector such that $\mathbf{x} \cdot \mathbf{e}$ appears in the kth component. Vectors $\mathbf{x}, \bar{\mathbf{x}}$ and $\mathbf{x}_k$ have the same total sum of components but the components of $\bar{\mathbf{x}}$ are more evenly spread out than those of $\mathbf{x}$. In this sense, $\bar{\mathbf{x}}$ is the most evenly spread of this sum of components and $\mathbf{x}_k$ is the most concentrated of this sum. Indeed, we have that $\bar{\mathbf{x}} \prec \mathbf{x} \prec \mathbf{x}_k$ $(k = 1, ..., n)$.

The majorization relation is reflexive and transitive but it is not antisymmetric (see [106, p. 79]) and hence is a preordering but not a partial ordering. The majorization preorder on vectors is known as vector majorization or classical majorization. The notation arises in a variety of contexts. Indeed, many majorization concepts have been reinvented and used in different research areas, as Lorenz or dominance ordering in economics, optimization and graph theory. Naturally extensions of the majorization concept are possible and indeed many of them have been fruitfully introduced.

A mathematical origin of majorization is illustrated by Schur (1923) in his work on the spectra of positive semidefinite Hermitian operators (see [147]). Many key ideas related to the concept of majorization are presented in the monograph *Inequality* (1934) [60], by Hardy, Littlewood, and Pólya, the first monograph on inequalities. However, this monograph attracted only a relatively small number of researchers in terms of problems related to the concept of majorization. The publication of monograph *Inequalities: Theory of Majorization and Its Applications* (1979) [106], by authors Marshall, Olkin and Arnold, contributes a great interest in the potential application of the concept of majorization and Schur convexity in various scientific areas. It provides a systematic overview of past results from the field of mathematical inequalities with a special emphasis on majorization for the first time, and has been complemented by many new results. Here we also refer the reader to the monograph [23] (1997) by Bhatia which contains significant material on majorization theory as well. Other textbooks on matrix and multivariate analysis also include a section on majorization theory, e.g., [63, Sec.4.3], [15, Sec.8.10] and [138]. In last years, a lot of papers in journals in a wide variety of fields, was dedicated to majorization theory.

In majorization theory, the well known Majorization inequality

$$\sum_{i=1}^{m} f(y_i) \leq \sum_{i=1}^{m} f(x_i) \tag{1}$$

plays an important role. It holds for every convex function f and real vectors $\mathbf{x} = (x_1, ..., x_m)$, $\mathbf{y} = (y_1, ..., y_m)$ such that $\mathbf{y} \prec \mathbf{x}$. In literature this inequality is also known as Karamata's or the Hardy-Littlewood-Pólya inequality.

Sherman [149], replacing the classical concept of majorization $\mathbf{y} \prec \mathbf{x}$ with the concept of weighted majorization, in notation $(\mathbf{y}, \mathbf{b}) \prec (\mathbf{x}, \mathbf{a})$, for two pairs $(\mathbf{x}, \mathbf{a})$ and $(\mathbf{y}, \mathbf{b})$ which consist of real vectors $\mathbf{x} = (x_1, ..., x_m)$, $\mathbf{y} = (y_1, ..., y_n)$ and corresponding nonnegative weights $\mathbf{a} = (a_1, ..., a_m)$, $\mathbf{b} = (b_1, ..., b_n)$, obtained a useful generalization of Majorization inequality (1) in the form

$$\sum_{j=1}^{n} b_j f(y_j) \leq \sum_{i=1}^{m} a_i f(x_i). \tag{2}$$

This book is based on some significant papers dealing with Sherman's type inequalities, published in the course of the last fifteen years. More precisely, it deals with series of refinements and generalizations of Sherman's inequality (2), Majorization inequality (1) and related results in various spaces for different classes of functions. The book can serve as a reference and a source of inspiration for researchers working in these and related areas. The methods presented in the book can be further applied to different concepts of generalized convexity, which can contribute to the further development of the methods and provide perspective for further applications.

The book is divided into five chapters. In **Chapter 1**, we give an overview of definitions and important basic results necessary for establishing the results that will follow. We start **Chapter 2** with the concept of majorization and Schur-convexity in order to give some basic results from Theory of majorization that give an important characterization of convex functions. We present Sherman's result with proof of Sherman's inequality (2) as well as its converse. We introduce so called Sherman's functional deduced from (2) and investigate its essential properties. We explore some generalizations of the well known inequalities as the Jensen-Steffensen, Mercer's and Brunk's for (α, β)-convex functions as well as generalization of Fejér's, the Hermite-Hadamard and related inequalities for convex sequences. We also present the usefulness of Sherman method in deriving variants of Sherman's type inequalities for different classes of functions as normalized positive linear maps, superquadratic and strongly convex functions. In **Chapter 3**, applying different interpolation methods in combination with Green's functions, we obtain new generalizations concerning convex functions of higher order. We also get some related results including Čebyšev functional and as outcome we derive new inequalities of the Grüss and Ostrowski type. Using so called exp-convex method, established in [77], we obtain new classes of Cauchy type means and obtained results we interprete in the form of exponentially or in the special case logarithmically convex functions. In **Chapter 4**, using different variants of Sherman's type inequalities derived by Sherman's method, we get new estimates and generalizations for Csiszár f -divergence and in special cases new bounds for Shannon's entropy, the Kullback-Leibler divergence and other related divergences. Also, we observe given results in the context of Zipf's law [162] and its generalized version so called the Zipf-Mandelbrot law [105]. In **Chapter 5** we deal with the Sherman type inequalities for operator convex functions, whose arguments are the bounded self-adjoint operators from C^*-algebra on a Hilbert space and positive linear mappings between C^*-algebras. Using extended idea of convexity to operator functions of several variables, we obtain multidimensional Sherman's operator inequality. We consider some applications to operator inequalities related to connections, solidarities, and multidimensional weighted geometric means. We also extend the results for Csiszár f-divergence functionals to operator Csiszár f -divergences and in consequence, we get the Shannon like inequalities for operators. At the end we investigate vectorial joint majorization with applications to operator Csiszár f-divergence.

In the end of this preface we want to emphasize that this book integrates the whole variety of results from different papers which were previous published in journals by different authors. It was practically impossible to quite unify the notation in the book.

Authors

Lahore, Pakistan

March, 2026

CHAPTER 1.

Basic notation and fundamental results

In this chapter, we give a brief review of some fundamental results on the topics in the sequel and present several basic motivating ideas.

1.1. Convex and n-convex functions

Definition 1.1.1.: Let I be an real interval. Then $\phi : I \to \mathbb{R}$ is said to be **convex** if for all $x, y \in I$ and every $\lambda \in [0,1]$, we have

$$\phi\left((1-\lambda)x + \lambda y\right) \leq (1-\lambda)\phi(x) + \lambda\phi(y). \tag{1.1}$$

If (1.1) is strict for all $x, y \in I$, $x \neq y$ and every $\lambda \in (0,1)$, then ϕ is said to be **strictly convex**.
If in (1.1) the reverse inequality holds, then ϕ is said to be **concave**. If it is strict for all $x, y \in I$, $x \neq y$ and every $\lambda \in (0,1)$, then ϕ is said to be **strictly concave**.

For convex functions the following propositions are valid which exactly define convex functions on equivalent ways.

Remark 1.1.1.: a) Inequality (1.1), for $x_1, x_2, x_3 \in I$, such that $x_1 \leq x_2 \leq x_3$, $x_1 \neq x_3$, can be written in the form

$$\phi\left(x_2\right) \leq \frac{x_3 - x_2}{x_3 - x_1}\phi\left(x_1\right) + \frac{x_2 - x_1}{x_3 - x_1}\phi\left(x_3\right), \tag{1.2}$$

i.e.

$$\left(x_3 - x_2\right)\phi\left(x_1\right) + \left(x_1 - x_3\right)\phi\left(x_2\right) + \left(x_2 - x_1\right)\phi\left(x_3\right) \geq 0, \tag{1.3}$$

by setting $x = x_1,\ y = x_3,\ \lambda = \left(x_2 - x_1\right)/\left(x_3 - x_1\right)$. This inequality is often used as alternative definition of convexity.

b) Another way of writing (1.3) is

$$\frac{\phi\left(x_2\right) - \phi\left(x_1\right)}{x_2 - x_1} \leq \frac{\phi\left(y_2\right) - \phi\left(y_1\right)}{y_2 - y_1}, \tag{1.4}$$

where $x_1 \leq y_1,\ x_2 \leq y_2,\ x_1 \neq x_2$ and $y_1 \neq y_2$.

The following two theorems concern derivatives of convex functions.

Theorem 1.1.1: *(see [138, p. 4]) Let I be an real interval. Let $\phi : I \to \mathbb{R}$ be convex. Then*

(i) ϕ is Lipschitz on any closed interval in I;

(ii) ϕ'_- and ϕ'_+ exist and are increasing on I, and $\phi'_- \leq \phi'_+$ (if ϕ is strictly convex, then these derivatives are strictly increasing);

(iii) ϕ' exists, except possibly on a countable set, and on the complement of which it is continuous.

Remark 1.1.2.: a) If $\phi : I \to \mathbb{R}$ is differentiable function, then ϕ is convex iff a function ϕ' is increasing.

b) If $\phi : I \to \mathbb{R}$ is twice derivable function, then ϕ is convex iff $\phi''(x) \geq 0$ for all $x \in I$. If $\phi''(x) > 0$, then ϕ is strictly convex.

Theorem 1.1.2: *(see [138, p. 5]) Let I be an open interval in $\mathbb{R}$.*

(i) $\phi : I \to \mathbb{R}$ is convex iff there is at least one line of support for ϕ at each $x_0 \in I$, i.e.

for all $x \in I$ we have

$$\phi(x) \geq \phi(x_0) + \lambda(x - x_0),$$

where $\lambda \in \mathbb{R}$ depends on x_0 and is given by $\lambda = \phi'(x_0)$ when $\phi'(x_0)$ exists, and $\lambda \in \left[\phi'_-(x_0), \phi'_+(x_0)\right]$ when $\phi'_-(x_0) \neq \phi'_+(x_0)$.

(ii) $\phi : I \to \mathbb{R}$ is convex if the function $x \longmapsto \phi(x) - \phi(x_0) - \lambda(x - x_0)$, (the difference between the function and its support) is decreasing for $x < x_0$ and increasing for $x > x_0$.

Definition 1.1.2.: Let $\phi : I \to \mathbb{R}$ be a convex function. Then the subdifferential of ϕ at $x \in I$, denoted by $\partial\phi(x)$, is defined by

$$\partial\phi(x) = \{\alpha \in \mathbb{R} : \phi(y) - \phi(x) - \alpha(y - x) \geq 0, y \in I\}.$$

There is a connection between a convex function and its subdifferential. It is well-known that $\partial\phi(x) \neq \emptyset$ for all $x \in IntI$, where $IntI$ is the interior of an interval I. More precisely, at each point $x \in IntI$ we have $-\infty < \phi'_-(x) \leq \phi'_+(x) < \infty$ and

$$\partial\phi(x) \in \left[\phi'_-(x_0), \phi'_+(x_0)\right],$$

while the set on which ϕ is not differentiable is at most countable. Moreover, each function $\varphi : I \to \mathbb{R}$ such that $\varphi(x) \in \partial\phi(x)$, whenever $x \in IntI$, is increasing on $IntI$. For any such function φ and arbitrary $x \in IntI$, $y \in I$, we have

$$\phi(y) - \phi(x) - \varphi(x)(y - x) \geq 0$$

and

$$\begin{aligned} \phi(y) - \phi(x) - \varphi(x)(y - x) &= |\phi(y) - \phi(x) - \varphi(x)(y - x)| \\ &\geq \| \phi(y) - \phi(x)| - |\varphi(x)| \cdot |(y - x)\| . \end{aligned}$$

J. L. W. V. Jensen [78] is considered generally as being the first mathematician who studied convex functions in a systematic way. He defined the concept of convex functions using the inequality (1.5) that are listed in the following definition.

Definition 1.1.3.: A function $\phi : I \to \mathbb{R}$ is called **Jensen-convex** or **J-convex** if for all $x, y \in I$ we have

$$\phi\left(\frac{x+y}{2}\right) \leq \frac{\phi(x)+\phi(y)}{2}. \tag{1.5}$$

Remark 1.1.3.: It can be easily proved that a convex function is J-convex. If $\phi : I \to \mathbb{R}$ is continuous function, then ϕ is convex iff it is J-convex.

Inequality (1.1) can be extended to the convex combinations of finitely many points in I by mathematical induction. This extension is known as **discrete Jensen's inequality**.

Theorem 1.1.3 ***(Discrete Jensen's inequality)*:** *Let I be an real interval and $\phi : I \to \mathbb{R}$ be a convex function. If $\mathbf{x} = (x_1, \ldots, x_n)$ is any n-tuple in I^n and $\mathbf{p} = (p_1, \ldots, p_n) \in [0, \infty)^n$ with $P_n = \sum_{i=1}^n p_i > 0$, then*

$$\phi\left(\frac{1}{P_n}\sum_{i=1}^n p_i x_i\right) \leq \frac{1}{P_n}\sum_{i=1}^n p_i \phi(x_i). \tag{1.6}$$

Remark 1.1.4.: If it is satified that

$$p_1 > 0, \quad p_2, ..., p_n \leq 0, \quad P_n > 0 \quad \text{and} \quad \frac{1}{P_n}\sum_{i=1}^n p_i x_i \in I,$$

then **the reverse of Jensen's inequality**

$$\phi\left(\frac{1}{P_n}\sum_{i=1}^n p_i x_i\right) \geq \frac{1}{P_n}\sum_{i=1}^n p_i \phi(x_i)$$

holds (see for example [138]).

Remark 1.1.5.: Closely connected to Jensen's inequality (1.6) is **the Lah-Ribarič inequality**

$$\frac{1}{P_n}\sum_{i=1}^n p_i \phi(x_i) \leq \frac{\beta - \bar{x}}{\beta - \alpha}\phi(\alpha) + \frac{\bar{x} - \alpha}{\beta - \alpha}\phi(\beta) \tag{1.7}$$

which holds for every convex function $\phi : I \to \mathbb{R}$, $\alpha, \beta \in I$, $\alpha < \beta$, $\mathbf{x} = (x_1, ..., x_n) \in [\alpha, \beta]^n$, $\mathbf{p} = (p_1, ..., p_n) \in [0, \infty)^n$ with $P_n = \sum_{i=1}^n p_i > 0$ and $\bar{x} = \frac{1}{P_n}\sum_{i=1}^n p_i x_i$ (see [100]). It gives an upper bound for the term $\frac{1}{P_n}\sum_{i=1}^{n} p_i\phi\left(x_i\right)$ and is also known as **converse Jensen's inequality**.

The generalization of the notion of a convex function to higher order convexity, i.e. to notion of n-convexity, was defined in terms of divided differences by T. Popoviciu [141] as follows.

Definition 1.1.4.: The second order divided difference of a function $\phi : I \to \mathbb{R}$ at mutually different points $y_0, y_1, y_2 \in I$ is defined recursively by

$$[y_i;\phi] = \phi(y_i),\ \ i = 0, 1, 2,$$

$$[y_i, y_{i+1};\phi] = \frac{\phi(y_{i+1}) - \phi(y_i)}{y_{i+1} - y_i},\ \ i = 0, 1,$$

$$[y_0, y_1, y_2;\phi] = \frac{[y_1, y_2;\phi] - [y_0, y_1;\phi]}{y_2 - y_0}. \tag{1.8}$$

The value $[y_0, y_1, y_2;\phi]$ is independent of the order of the points $y_0, y_1,$ and y_2. By taking limits this definition may be extended to include the cases in which any two or all three points coincide as follows: for all $y_0,\ y_1,\ y_2 \in I$

$$\lim_{y_1\to y_0}[y_0, y_1, y_2;\phi] = [y_0, y_0, y_2;\phi] = \frac{\phi(y_2) - \phi(y_0) - \phi^{'}(y_0)(y_2 - y_0)}{(y_2 - y_0)^2},\quad y_2 \neq y_0,$$

provided that ϕ' exists, and furthermore, taking the limits $y_i \to y_0$, $(i = 1, 2)$, in (1.8), we get

$$[y_0, y_0, y_0;\phi] \ = \lim_{y_i\to y_0}[y_0, y_1, y_2;\phi] = \frac{\phi^{''}(y_0)}{2} \text{ for } \ i = 1, 2$$

provided that $\phi^{''}$ exists.

The previous definition can be generalized as follows.

Definition 1.1.5.: The divided difference of order n, $n \in \mathbb{N}$, of a function $\phi : I \to \mathbb{R}$ at mutually different points $x_0, x_1, ..., x_n \in I$ is defined recursively by

$$[x_i; \phi] = \phi(x_i), \quad i = 0, ..., n, \tag{1.9}$$
$$[x_0, ..., x_n; \phi] = \frac{[x_1, ..., x_n; \phi] - [x_0, ..., x_{n-1}; \phi]}{x_n - x_0}.$$

Divided differences are very important notion by dealing the functions when discussing degree of smoothness. Since the value $[x_0, x_1, ..., x_n; \phi]$ is independent of the order of the points $x_0, ..., x_n$, this definition may be extended to include the case in which some or all the points coincide. Assuming that $\phi^{(j-1)}(x)$ exists, we define

$$[\underbrace{x, ..., x}_{j\text{-times}}; \phi] = \frac{\phi^{(j-1)}(x)}{(j-1)!}. \tag{1.10}$$

We follow the definition of n-convexity given by Karlin [84]:

Definition 1.1.6.: A function $\phi : I \to \mathbb{R}$ is n-**convex**, $n \geq 0$, if for all choices of $(n+1)$ distinct points $x_i \in I$, $i = 0, ..., n$, the nth order divided difference is nonnegative, i.e.

$$[x_0, x_1, ..., x_n; \phi] \geq 0.$$

Remark 1.1.6.: (i) From Definition 1.1.6., it follows that 2-convex functions are just convex functions. Furthermore, 1-convex functions are increasing functions and 0-convex functions are nonnegative functions.

(ii) An n-convex function ϕ need not be n-times differentiable. However, if $\phi^{(n)}$ exists, then ϕ is n-convex iff $\phi^{(n)} \geq 0$. A function ϕ is n-concave if $-\phi$ is n-convex. Also, if ϕ is n-convex for $n \geq 2$, then $\phi^{(k)}$ exists and ϕ is $(n-k)$-convex for $1 \leq k \leq n-2$. For more information see [138].

1.2. Exponentially convex and logarithmically convex functions

Exponentially convex functions are invented by Bernstein in [22] as a subclass of convex functions on an open interval. These functions have many nice properties, for example, they are analytical on their domain. Although we will need only few of these properties we point here that very good reference on general results about exponential convexity is [14].

Here I denotes an open interval in $\mathbb{R}$.

Definition 1.2.7.: A function $\phi : I \to \mathbb{R}$ is **exponentially convex** if it is continuous and

$$\sum_{i,j=1}^{m} p_i p_j \phi\left(x_i + x_j\right) \geq 0$$

holds for every choice of $m \in \mathbb{N}$, $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{R}^m$ and $\mathbf{x} = (x_1, \ldots, x_m) \in I^m$ with $x_i + x_j \in I$, $1 \leq i, j \leq n$.

The notation of n-exponential convexity is introduced in [137].

Definition 1.2.8.: For fixed $n \in \mathbb{N}$, a function $\phi : I \to \mathbb{R}$ is **n-exponentially convex in the Jensen sense** if

$$\sum_{i,j=1}^{n} p_i p_j \phi\left(\frac{x_i + x_j}{2}\right) \geq 0 \tag{1.11}$$

holds for all choices $p_i \in \mathbb{R}$ and $x_i \in I$, $i = 1, ..., n$.

A function $\phi : I \to \mathbb{R}$ is **n-exponentially convex** if it is n-exponentially convex in the Jensen sense and continuous.

Remark 1.2.7.: From Definition 1.2.8. it follows that 1-exponentially convex functions in the Jensen sense are exactly nonnegative functions. Also, n -exponentially convex functions in the Jensen sense are k-exponentially convex in the Jensen sense for every $k \in \mathbb{N}$, $k \leq n$.

Using basic calculus and definition of positive semidefinite matrices we can prove the following proposition.

Proposition 1.2.4: If $\phi : I \to \mathbb{R}$ is an n-exponentially convex function in the Jensen sense, then the matrix

$$\left[\phi\left(\frac{x_i + x_j}{2}\right)\right]_{i,j=1}^{k}$$

is positive semi-definite for all $k \in \mathbb{N}$, $k \leq n$ and $x_i \in I$, $1 \leq i \leq k$.

Particularly,

$$\det\left[\phi\left(\frac{x_i + x_j}{2}\right)\right]_{i,j=1}^{k} \geq 0$$

for all $k \in \mathbb{N}$, $k \leq n$ and $x_i \in I$, $1 \leq i \leq k$.

Definition 1.2.9.: A function $\phi : I \to \mathbb{R}$ is **exponentially convex in the Jensen sense** if it is n-exponentially convex in the Jensen sense for all $n \in \mathbb{N}$.

One of the most important properties of exponentially convex functions is their integral representation (see [14, p. 211]).

Theorem 1.2.5: *A function $\phi : I \to \mathbb{R}$ is exponentially convex iff*

$$\phi(x) = \int_{-\infty}^{\infty} e^{tx} d\sigma(t), \tag{1.12}$$

for some non-decreasing function $\sigma : \mathbb{R} \to \mathbb{R}$.

We give some examples of exponentially convex functions which we use later in the book. For more details see [77].

Example 1.: *The basic example of exponentially convex function is the function $\phi : I \to \mathbb{R}$ defined by $\phi(x) = ce^{kx}$ for every $c \geq 0$ and $k \in \mathbb{R}$.*

The next examples can be deduced using integral representation (1.12) and some results of the Laplace transform [158].

Example 2.:

(ii) The function $\phi : (0, \infty) \to \mathbb{R}$, $\phi(x) = x^{-k}$, is exponentially convex for every $k > 0$ since $x^{-k} = \int_0^{\infty} e^{-xt} \frac{t^{k-1}}{\Gamma(k)} dt$ (see [146, p. 210]).

(iii) The function $\phi:(0,\infty)\to(0,\infty)$, $\phi(x)=e^{-k\sqrt{x}}$, is exponentially convex for every $k>0$ since $e^{-k\sqrt{x}}=\int_0^\infty e^{-xt}e^{-k^2/4t}\frac{k}{2\sqrt{\pi t^3}}dt$ (see [146, p. 214]) .

Remark 1.2.8.: (i) For $n=1$, from (1.11) it follows

$$p_1^2\phi\left(x_1\right)\geq 0,$$

i.e. the exponential convex functions are nonnegative. Also, even derivations of the exponentially convex functions are nonnegative because for every $k\in\mathbb{N}$ we have the following expression

$$\phi^{(2k)}(x)=\int_{-\infty}^{\infty}t^{2k}e^{tx}d\sigma(t).$$

(ii) For $n=2$, from (1.11) it follows

$$\begin{aligned}&\sum_{i,j=1}^{2}p_ip_j\phi\left(\frac{x_i+x_j}{2}\right)\\&=p_1^2\phi\left(x_1\right)+2p_1p_2\phi\left(\frac{x_1+x_2}{2}\right)+p_2^2\phi\left(x_2\right)\geq 0,\end{aligned}\tag{1.13}$$

what is equivalent to

$$\begin{bmatrix}p_1 & p_2\end{bmatrix}\begin{bmatrix}\phi\left(x_1\right) & \phi\left(\frac{x_1+x_2}{2}\right)\\ \phi\left(\frac{x_1+x_2}{2}\right) & \phi\left(x_2\right)\end{bmatrix}\begin{bmatrix}p_1\\ p_2\end{bmatrix}\geq 0,$$

from which we have

$$\left(\phi\left(\frac{x_1+x_2}{2}\right)\right)^2\leq\phi(x_1)\phi(x_2).$$

From the last inequality we conclude that in case when $\phi(x_1)=0$ for some $x_1\in I$, then $\phi(x)=0$ for every $x\in I$. Moreover, if ϕ is not constant (null function), then it is positive on whole domain.

(iii) If we take $p_1=1$ and $p_2=-1$, then from (1.13) we get

$$\phi\left(\frac{x_1+x_2}{2}\right)\leq\frac{\phi(x_1)+\phi(x_2)}{2}.$$

So, exponential convexity implies convexity.

Below we give definition of log convexity.

Definition 1.2.10.: A function $\phi : I \to (0, \infty)$ is said to be **logarithmically convex** or **log-convex** if the function $\log \phi$ is convex, or equivalently, if

$$\phi\left((1-\lambda)s + \lambda t\right) \leq \left(\phi(s)\right)^{1-\lambda} \left(\phi(t)\right)^{\lambda} \tag{1.14}$$

holds for all $s, t \in I, \lambda \in [0, 1]$.

Remark 1.2.9.: If functions ϕ and g are convex and g is increasing, then function $g \circ \phi$ is convex. Since $\phi = \exp \circ \log \circ \phi$ we have that logarithmically convex function is also convex, but the reverse conclusion doesn't hold.

Definition 1.2.11.: A function $\phi : I \to (0, \infty)$ is said to be logarithmically convex in the Jensen sense if

$$\left(\phi\left(\frac{x+y}{2}\right)\right)^2 \leq \phi(x)\phi(y) \tag{1.15}$$

holds for all $x, y \in I$.

Remark 1.2.10.: (i) A function ϕ is log-convex iff for all $x_1, x_2, x_3 \in I$, $x_1 < x_2 < x_3$, it holds

$$\left(\phi(x_2)\right)^{x_3 - x_1} \leq \left(\phi(x_1)\right)^{x_3 - x_2} \left(\phi(x_3)\right)^{x_2 - x_1}.$$

Furthermore, if $x_1, x_2, y_1, y_2 \in I$, $x_1 \leq y_1$, $x_2 \leq y_2$, $x_1 \neq x_2$, $y_1 \neq y_2$, then

$$\left(\frac{\phi(x_2)}{\phi(x_1)}\right)^{\frac{1}{x_2 - x_1}} \leq \left(\frac{\phi(y_2)}{\phi(y_1)}\right)^{\frac{1}{y_2 - y_1}}.$$

These two inequality are easy consequence of (1.2) and (1.4) for the convex function $\log \phi$.
(ii) Inequality (1.15) follows from (1.14) choosing $\lambda = \frac{1}{2}$, i.e. log-convexity implies log-convexity in the Jensen sense. On the other side, if a function is continuous and log-convex in the Jensen sense then it is also log-convex.
(iii) When ϕ is positive exponentially convex function, then it is also log convex because

from the Definition 1.2.8., for $n = 2$, we have

$$\left(\phi\left(\frac{x_1+x_2}{2}\right)\right)^2 \leq \phi(x_1)\phi(x_2)$$

for all $x_1, x_2 \in I$. This inequality defines the log-convexity in the Jensen sense. However an exponentially convex function is continuous, we conclude that the exponential convexity implicates log-convexity.

The converse doesn't hold true what we can see from the example of the function $\phi : (0,1) \to \mathbb{R}$, $\phi(x) = \mathrm{e}^{x^3-x}$. It can be easily proved that ϕ is log convex since

$$(\log\phi(x))'' = \left(x^3 - x\right)'' = 6x > 0 \quad \text{for} \quad x \in (0,1).$$

But, when we calculate $\phi^{(4)}(x)$ we obtain the function

$$\phi^{(4)}(x) = \mathrm{e}^{x^3-x}(81x^8 - 108x^6 + 324x^5 + 54x^4 - 216x^3 + 168x^2 + 36x - 23),$$

which is not nonnegative on whole interval $(0,1)$.

1.3. Mean value theorems

Definition 1.3.12.: A **mean** on I^n, where $I \subseteq \mathbb{R}$ is an interval, is every function $M : I^n \to \mathbb{R}$, with properties:

(M_1) For every choice $\mathbf{x} = (x_1, ..., x_n) \in I^n$,

$$\min\{x_1, x_2, ..., x_n\} \leq M(\mathbf{x}) \leq \max\{x_1, x_2, ..., x_n\}.$$

(M_2) M is symmetric in the sense that for every $\mathbf{x} = (x_1, ..., x_n) \in I^n$,

$$M(\mathbf{x}) = M(\tilde{\mathbf{x}}),$$

whenever $\tilde{\mathbf{x}} = (x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)})$, where $\sigma : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ is any permutation.

(M_3) M is continuous on I^n.

A natural extension of the previous definition is obtained when some of the components of $\mathbf{x} = (x_1, ..., x_n)$ occur more then once, or when, from the practical reason, some of the components of $\mathbf{x} = (x_1, ..., x_n)$ are considered to be more important then others. In that case we introduce an n-tuple of weights $\mathbf{w} = (w_1, ..., w_n)$ and define weighted mean.

The classical weighted generalization is given in the following definition.

Definition 1.3.13.: **The weighted power mean** of order $s \in \mathbb{R}$ of $\mathbf{x}$, where $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{w} = (w_1, ..., w_n)$ are positive real n-tuples such that $\sum_{i=1}^{n} w_i = 1$, is defined by

$$M_s(\mathbf{x}) = M_s(\mathbf{x}, \mathbf{w}) := \begin{cases} \left(\sum\limits_{i=1}^{n} w_i x_i^s\right)^{\frac{1}{s}}, & s \neq 0 \\ \prod\limits_{i=1}^{n} x_i^{w_i}, & s = 0 \\ \min\{x_1, ..., x_n\}, & s \to -\infty \\ \max\{x_1, ..., x_n\} & s \to \infty \end{cases}. \tag{1.16}$$

In that case, the property (M_2) is slightly changed. Mean must be symmetric in the following way:

$(M_{2'})$ for $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{w} = (w_1, ..., w_n)$

$$M_s(\mathbf{x}, \mathbf{w}) = M_s(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})$$

whenever $\tilde{\mathbf{x}} = (x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(n)})$ and $\tilde{\mathbf{w}} = (w_{\sigma(1)}, w_{\sigma(2)}, ..., w_{\sigma(n)})$, where $\sigma : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ is any permutation.

As a special case of the weighted power mean we get:

- **weighted arithmetic mean** $A(\mathbf{x}, \mathbf{w}) = M_1(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^{n} w_i x_i$,
- **weighted geometric mean** $G(\mathbf{x}, \mathbf{w}) = M_0(\mathbf{x}, \mathbf{w}) = \prod\limits_{i=1}^{n} x_i^{w_i}$,
- **weighted harmonic mean** $H(\mathbf{x}, \mathbf{w}) = M_{-1}(\mathbf{x}, \mathbf{w}) = \frac{1}{\sum_{i=1}^{n} \frac{w_i}{x_i}}$.

The classical **HGA inequality,** between the harmonic, geometric and arithmetic means,

$$H(\mathbf{x}, \mathbf{w}) \leq \mathbf{G}(\mathbf{x}, \mathbf{w}) \leq A(\mathbf{x}, \mathbf{w}),$$

can be extended as follows.

If $r < s$, then the **power mean inequality**

$$M_r(\mathbf{x}, \mathbf{w}) \leq M_s(\mathbf{x}, \mathbf{w}),$$

holds with equality iff $x_1 = ... = x_n$.

As examples, we present classes of means that follows from the well-known mean value theorems.

Theorem 1.3.6 ***(Lagrange's mean value theorem)*****:** *If a function $\varphi : [x, y] \to \mathbb{R}$ is continuous on a closed interval $[x, y]$ and differentiable on the open interval (x, y), then there is at least one point $\xi \in (x, y)$ such that*

$$\varphi'(\xi) = \frac{\varphi(y) - \varphi(x)}{y - x}.$$

Under assumption that the function φ' is invertible, from Lagrange's theorem it follows that there is a unique number

$$\xi = (\varphi')^{-1} \left(\frac{\varphi(y) - \varphi(x)}{y - x} \right)$$

which we call **Lagrange's mean** of $[x, y]$.

We can generalize Lagrange's mean using Cauchy's mean value theorem.

Theorem 1.3.7 ***(Cauchy's mean value theorem)*****:** *Let functions $\varphi, \psi : [x, y] \to \mathbb{R}$ be continuous on an interval $[x, y]$ and differentiable on (x, y) and let $\psi'(t) \neq 0$ for all $t \in (x, y)$. Then there is a point $\xi \in (x, y)$ such that*

$$\frac{\varphi'(\xi)}{\psi'(\xi)} = \frac{\varphi(y) - \varphi(x)}{\psi(y) - \psi(x)}.$$

Under assumption that the function $\frac{\varphi'}{\psi'}$ is invertible, from Cauchy's theorem it follows that there is a unique number

$$\xi = \left(\frac{\varphi'}{\psi'}\right)^{-1}\left(\frac{\varphi(y)-\varphi(x)}{\psi(y)-\psi(x)}\right).$$

which we call **Cauchy's mean** of interval $[x, y]$. Continuous expansion gives $\xi = x$ if $y = x$.

If we take $\psi(x) = x$, as a special case of Cauchy's mean we get Lagrange's mean. Moreover, many well known means in mathematics are special cases of Cauchy's mean. Under assumption that $x, y \in (0, \infty)$ and choosing $\varphi(x) = x^v$ and $\psi(x) = x^u$, $u, v \in \mathbb{R}$, $u \neq v$, $u, v \neq 0$, we obtain two-parameter mean

$$E_{u,v}(x,y) = \left(\frac{u(y^v - x^v)}{v(y^u - x^u)}\right)^{\frac{1}{v-u}}$$

firstly introduced by Stolarsky [152, 153] and in literature known as **Stolarsky's mean**.

Stolarsky has also proved that $E_{u,v}(x, y)$ can be extended by continuity as follows:

$$\begin{aligned}
&E_{u,v}(x,y) = \left(\frac{u(y^v - x^v)}{v(y^u - x^u)}\right)^{\frac{1}{v-u}}, \quad u \neq v,\ uv \neq 0,\ x \neq y,\\
&E_{u,0}(x,y) = E_{0,v}(x,y) = \left(\frac{y^u - x^u}{u(\log y - \log x)}\right)^{\frac{1}{u}}, \quad u \neq 0,\ x \neq y,\\
&E_{u,u}(x,y) = e^{-\frac{1}{u}}\left(\frac{y^{y^u}}{x^{x^u}}\right)^{\frac{1}{y^u - x^u}}, \quad x \neq y,\ u \neq 0,\\
&E_{0,0}(x,y) = \sqrt{xy}, \quad x \neq y,\\
&E_{u,v}(x,x) = x.
\end{aligned}$$

This mean is symmetric, i.e. $E_{u,v}(x, y) = E_{v,u}(x, y)$ holds for all choice of numbers $u, v \in \mathbb{R}$, $x, y \in (0, \infty)$. It is also monotonic in both parameters, i.e. for $r, s, u, v \in \mathbb{R}$, such that $r \leq s$, $u \leq v$, we have

$$E_{r,u}(x,y) \leq E_{s,v}(x,y).$$

As special cases of the Stolarsky's mean, we get basic means:

- arithmetic mean $E_{1,2}(x, y) = \dfrac{x+y}{2}$;
- geometric mean $E_{0,0}(x, y) = \sqrt{xy}$;
- harmonic mean $E_{-2,-1}(x, y) = \dfrac{2xy}{x+y}$;

- power mean of order r $E_{r,2r}(x,y) = \left(\dfrac{x^r + y^r}{2}\right)^{\frac{1}{r}}$;
- logarithmic mean $E_{1,0}(x,y) = \dfrac{y-x}{\log y - \log x}$;
- identric mean $\lim_{u\to 1} E_{u,u}(x,y) = \dfrac{1}{\mathrm{e}}\left(\dfrac{y^y}{x^x}\right)^{\frac{1}{y-x}}$, etc.

Following the ideas of generating new means described after Lagrange's mean value theorem and Cauchy's mean value theorem and using the following extreme value theorem, in the third chapter we deduce the Lagrange and Cauchy type mean value theorems, and as consequence we generate new Cauchy's type means.

Theorem 1.3.8 ***(Extreme value theorem)*:** *If $\varphi : [x,y] \to \mathbb{R}$ is continuous function on a closed interval $[x,y]$, then φ is bounded and attains a maximum and minimum value over that interval, i.e. $m \le \varphi(t) \le M$ for all $t \in [x,y]$, where $m = \min_{t\in[x,y]} \varphi(t)$ and $M = \max_{t\in[x,y]} \varphi(t)$.*

1.4. Spaces of integrable, continuous and absolutely continuous functions

Let $[a,b]$ be a finite interval in $\mathbb{R}$. We denote by $L_p[a,b]$, $1 \le p < \infty$, the space of all Lebesgue measurable functions f for which

$$\int_a^b |f(t)|^p dt < \infty,$$

where

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}},$$

and by $L_\infty[a,b]$ the set of all functions measurable and essentially bounded on $[a,b]$ with

$$\|f\|_\infty = ess\sup\{|f(x)| : x \in [a,b]\}.$$

Theorem 1.4.9 ***(Integral Hölder's inequality)*:** *Let $p,q \in \mathbb{R}$ be such that $1 \le p,q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $f,g : [a,b] \to \mathbb{R}$ be integrable functions such that $f \in L_p[a,b]$ and*

$g \in L_q[a,b]$. *Then*

$$\int_a^b |f(t)g(t)|dt \leq \|f\|_p \|g\|_q \ . \tag{1.17}$$

The equality in (1.17) holds iff $A|f(t)|^p = B|g(t)|^q$ *almost everywhere, where* A *and* B *are constants.*

We denote by $C^n([a,b]), n \in \mathbb{N}_0$, the space of functions which are n times continuously differentiable on $[a,b]$, that is

$$C^n([a,b]) = \left\{ f : [a,b] \to \mathbb{R} : f^{(k)} \in C([a,b]), k = 0,1,...,n \right\}$$

with the norm

$$\|f\|_{C^n} = \sum_{k=0}^{n} \left\| f^{(k)} \right\|_C = \sum_{k=0}^{n} \max_{x \in [a,b]} \left| f^{(k)}(x) \right|.$$

In particular, $C^0([a,b]) = C([a,b])$ is the space of continuous functions on $[a,b]$ with the norm

$$\|f\|_C = \max_{x \in [a,b]} |f(x)| \, .$$

Lemma 1.4.1.: *The space* $C^n([a,b])$ *consists of those and only those functions* f *which are represented in the form*

$$f(x) = \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} \varphi(t) dt + \sum_{k=0}^{n-1} c_k (x-a)^k, \tag{1.18}$$

where $\varphi \in C([a,b])$ *and* c_k *are arbitrary constants* $(k = 0,1,...,n-1)$*.*

Moreover,

$$\varphi(t) = f^{(n)}(t), \quad c_k = \frac{f^{(k)}(a)}{k!} \quad (k = 0,1,...,n-1). \tag{1.19}$$

The space of absolutely continuous functions on an interval $[a,b]$ is denote by $AC([a,b])$. It is known that $AC([a,b])$ coincides with the space of primitives of Lebesgue integrable functions $L_1[a,b]$ (see [91]:

$$f \in AC([a,b]) \quad \Leftrightarrow \quad f(x) = f(a) + \int_a^x \varphi(t) dt, \quad \varphi \in L_1[a,b].$$

Therefore, an absolutely continuous function f has an integrable derivatives $f'(x) = \varphi(x)$ almost everywhere on $[a, b]$. We denote by $AC^n([a, b]), n \in \mathbb{N}$, the space

$$AC^n([a, b]) = \{f \in C^{n-1}([a, b]) : f^{(n-1)} \in AC([a, b])\}.$$

In particular, $AC^1([a, b]) = AC([a, b])$.

Lemma 1.4.2.: *The space $AC^n([a, b])$ consists of those and only those functions which can be represented in the form (1.18), where $\varphi \in L_1[a, b]$ and c_k are arbitrary ($k = 0, 1, .., n-1$). Moreover, (1.19) holds.*

The next theorem has numerous applications involving multiple integrals.

Theorem 1.4.10 ***(Fubini's theorem)*:** *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure space and f be $\mu \times \nu$-measurable function on $X \times Y$. If $f \geq 0$, then the following integrals are equal*

$$\int_{X\times Y} f(x, y)d(\mu \times \nu)(x, y), \quad \int_X \left(\int_Y f(x, y)d\nu(y)\right) d\mu(x),$$
$$\int_Y \left(\int_X f(x, y)d\mu(x)\right) d\nu(y).$$

Remark 1.4.11.: The next equality

$$\int_a^b \left(\int_c^d f(x, y)dy\right) dx = \int_c^d \left(\int_a^b f(x, y)dx\right) dy$$

is consequence of the previous theorem.

Theorem 1.4.11 ***(Integral Jensen's inequality)*:** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space with $0 < \mu(\Omega) < \infty$. Let $\phi : \Omega \to \mathbb{R}$ be μ -integrable function and $f : I \to \mathbb{R}$ be a convex function such that $\phi(\Omega) \subset I$ and $f \circ \phi$ be a μ -integrable function. Then*

$$f\left(\frac{1}{\mu(\Omega)} \int_\Omega \phi(x)d\mu(x)\right) \leq \frac{1}{\mu(\Omega)} \int_\Omega f(\phi(x))d\mu(x). \tag{1.20}$$

If f is strictly convex, then (1.20) becomes equality iff ϕ is a constant μ-almost everywhere *on Ω. If f is concave, then (1.20) is reversed.*

Remark 1.4.12.: The discrete Jensen inequality (1.6) is obtained by means of the discrete measure μ on $\Omega = \{1, ..., n\}$, with $\mu(\{i\}) = p_i$ and $\phi(i) = x_i$ for $i \in \Omega$.

At the end of this section, we introduce two recently obtained results involving the Čebyšev functional.

Definition 1.4.14.: For two Lebesgue integrable functions $f, g : [\alpha, \beta] \to \mathbb{R}$, we define **the Čebyšev functional** as

$$T(f,g) := \frac{1}{\beta-\alpha}\int\limits_{\alpha}^{\beta} f(t)g(t)dt - \frac{1}{\beta-\alpha}\int\limits_{\alpha}^{\beta} f(t)dt \cdot \frac{1}{\beta-\alpha}\int\limits_{\alpha}^{\beta} g(t)dt.$$

Theorem 1.4.12: *[32, Theorem 1] Let $f : [\alpha, \beta] \to \mathbb{R}$ be Lebesgue integrable and $g : [\alpha, \beta] \to \mathbb{R}$ be absolutely continuous with $(\cdot - \alpha)(\beta - \cdot)(g')^2 \in L_1[\alpha, \beta]$. Then*

$$|T(f,g)| \leq \frac{1}{\sqrt{2}}[T(f,f)]^{\frac{1}{2}} \frac{1}{\sqrt{\beta-\alpha}} \left(\int_{\alpha}^{\beta} (x-\alpha)(\beta-x)[g'(x)]^2 dx \right)^{\frac{1}{2}}. \tag{1.21}$$

The constant $\frac{1}{\sqrt{2}}$ in (1.21) is the best possible.

Theorem 1.4.13: *[32, Theorem 2] Let $g : [\alpha, \beta] \to \mathbb{R}$ be monotonic nondecreasing and $f : [\alpha, \beta] \to \mathbb{R}$ be absolutely continuous with $f' \in L_\infty[\alpha, \beta]$. Then*

$$|T(f,g)| \leq \frac{1}{2(\beta-\alpha)} \left\| f' \right\|_\infty \int_{\alpha}^{\beta} (x-\alpha)(\beta-x)dg(x). \tag{1.22}$$

The constant $\frac{1}{2}$ in (1.22) is the best possible.

1.5. C*-algebras. Operators on a Hilbert space

Dealing with the Jensen-type functionals whose real arguments are substituted with the bounded self-adjoint operators acting on a Hilbert space, the ones that we are also going to investigate in the sequel, requires additional overview of some basic notions and results. These concern C*-algebras and the theory on operators on a Hilbert space. For a more detailed analysis, the reader is referred to e.g. [14] and [54].

Recall that a linear space $\mathcal{A}$ over the field $\mathbf{F}$, together with the multiplication $(x, y) \mapsto xy$, $x, y \in \mathcal{A}$, constitutes an *algebra* if the mapping $(x, y) \mapsto xy$ possesses the following properties:

$$(xy)z = x(yz),\ x(y+z) = xy + xz,\ (x+y)z = xz + yz,$$

$$(\alpha x)y = \alpha(xy) = x(\alpha y),$$

for all $\alpha \in \mathbf{F}$ and for all $x, y, z \in \mathcal{A}$. If $\mathbf{F} = \mathbb{R}$, algebra is called a real algebra and if $\mathbf{F} = \mathbb{C}$, we call it a complex algebra. If there is an element 1 in $\mathcal{A}$ such that $x \cdot 1 = 1 \cdot x = x$, for all $x \in \mathcal{A}$, then we say that $\mathcal{A}$ is an algebra with the unit 1.

Mapping $x \mapsto \|x\|$ defined on an algebra $\mathcal{A}$ and with the values in $\mathbb{R}$ is a *norm* on $\mathcal{A}$ if the following conditions are satisfied: $x \mapsto \|x\|$ is a norm on the linear space of $\mathcal{A}$, $\|xy\| \leq \|x\|\|y\|$, $x, y \in \mathcal{A}$, and, if $\mathcal{A}$ has a unit 1, then $\|1\| = 1$. Ordered pair $(\mathcal{A}, \|\cdot\|)$ is called a *normed algebra*. Normed algebra is called a *Banach algebra* if the normed space of $\mathcal{A}$ is a Banach space, that is a complete normed vector space.

Involution on algebra $\mathcal{A}$ is a conjugate linear mapping $x \mapsto x^*$ on $\mathcal{A}$, such that $x^{**} = x$ and $(xy)^* = y^* x^*$, for all $x, y \in \mathcal{A}$. Ordered pair $(\mathcal{A}, *)$ is called an *involutive algebra* or a **-algebra*. Now, a *-algebra $\mathcal{A}$ equipped with the complete submultiplicative norm such that $\|x^*\| = \|x\|$, for all $x \in \mathcal{A}$ constitutes a *Banach *-algebra*.

Finally, Banach *-algebra is a *C*-algebra* if for every $x \in \mathcal{A}$, C*-identity $\|x^* x\| = \|x\|^2$ holds. We say that a C*-algebra is unital if it contains the unit 1. The element $a \in \mathcal{A}$ is called: self-adjoint if $a = a^*$, normal if $a^* a = a a^*$ and is called unitary (in the unital algebra) if $a^* a = a a^* = 1$. A standard example of the unital C*-algebra is the scalar field $\mathbb{C}$, where the involution is represented as the complex conjugation.

The environment that is of an interest for us in the sequel is the one of the bounded linear operators on a Hilbert space, which we therefore specify here.

A Hilbert space H is a (complex) vector space H that is complete regarding the metric $d(x, y) = \|x - y\|$ defined by the norm which is induced by an inner product $\|x\| := \langle x, x \rangle^{\frac{1}{2}}$.

In other words, Hilbert space is a complete unitary space.

A linear operator A on a Hilbert space H is bounded if

$$\|A\| := \sup\{\|Ax\| : \|x\| \leq 1, x \in H\} < \infty.$$

We say that $\|A\|$ is an operator norm of A. The sum and the composition of the bounded linear operators is again a bounded linear operator. The mapping $(x, y) \mapsto \langle Ax, y\rangle$ is linear and continuous and according to the Riesz representation theorem (see e.g. [16]), it follows that

$$\langle x, A^*y\rangle = \langle Ax, y\rangle$$

for some uniquely determined vector $A^*y \in H$. Thus another bounded linear operator A^* on H is defined and $A^{**} = A$. Bounded (hence continuous) linear operators on H together with an operator norm and the corresponding involution constitute a C*-algebra denoted with $\mathcal{B}(H)$. *Spectrum* of an operator $A \in \mathcal{B}(H)$ is defined with

$$\sigma(A) := \{\lambda \in \mathbb{C} : A - \lambda 1_H \,\text{not invertible in}\, \mathcal{B}(H)\},$$

where 1_H is a unit operator on H. This set is non-empty and compact for operators in $\mathcal{B}(H)$. A bounded linear operator A on a Hilbert space H is self-adjoint if $A = A^*$. The following characterization holds: $A \in \mathcal{B}(H)$ is self-adjoint if and only if $\langle Ax, x\rangle \in \mathbb{R}$, for all $x \in H$.

Bounded self-adjoint operators constitute the real subspace of the C*-algebra of all bounded linear operators and is denoted with $\mathcal{B}_h(H)$. We induce the partial ordering in $\mathcal{B}_h(H)$.

Definition 1.5.15.: Operator $A \in \mathcal{B}_h(H)$ is positive semidefinite or positive ($A \geq 0$) if $\langle Ax, x\rangle \geq 0$, for all $x \in H$. Positive semidefinite operator $A \in \mathcal{B}_h(H)$ is positive definite or strictly positive ($A > 0$) if there is a real number $m > 0$ such that $\langle Ax, x\rangle \geq m\langle x, x\rangle$, $x \in H$, that is $A \geq m1_H$. For operators $A, B \in \mathcal{B}_h(H)$ is $B \geq A$ or $A \leq B$ if $B - A \geq 0$, that is if $\langle Bx, x\rangle \geq \langle Ax, x\rangle$, for all $x \in H$. Such ordering is called an operator ordering. In particular, for scalars m and M is $m1_H \leq A \leq M1_H$ if $m \leq \langle Ax, x\rangle \leq M$, for every unit vector $x \in H$.

If for a self-adjoint operator A is $\sigma(A) \subseteq [m, M]$, then $m1_H \leq A \leq M1_H$.

The set of all positive operators in $\mathcal{B}_h(H)$ is a convex cone in $\mathcal{B}_h(H)$ which defines the order " $\leq$ " on $\mathcal{B}_h(H)$. This convex cone is denoted with $\mathcal{B}^+(H)$. The set of all strictly positive (or positive invertible) operators in $\mathcal{B}_h(H)$ is denoted with $\mathcal{B}^{++}(H)$.

The *continuous functional calculus* is based on the Gelfand mapping Φ which is a *-isometric isomorphism from the space $C(\sigma(A))$ of all continuous functions that act on the spectrum $\sigma(A)$ of a self-adjoint operator A on H, onto the C*-algebra $C^*(A)$ generated with A and the identity. This mapping has the following properties:

(i) $\Phi(\alpha f + \beta g) = \alpha\Phi(f) + \beta\Phi(g)$,

(ii) $\Phi(fg) = \Phi(f)\Phi(g)$ and $\Phi(\overline{f}) = \Phi(f)^*$,

(iii) $\|\Phi(f)\| = \|f\| := \sup_{t\in\sigma(A)} |f(t)|$,

(iv) $\Phi(f_0) = 1_H$ and $\Phi(f_1) = A$, where $f_0(t) = 1$ and $f_1(t) = t$,

$f,\, g \in C(\sigma(A))$, $\alpha,\, \beta \in \mathbb{C}$.

Thus the continuous functional calculus

$$f(A) = \Phi(f) \tag{1.23}$$

provides acting of the function $f \in C(\sigma(A))$ on the self-adjoint operator A itself.

In that sense, if A is a positive operator and $f_{1/2}(t) = \sqrt{t}$, then $A^{1/2} = f_{1/2}(A)$.

Furthermore, if A is a self-adjoint operator and f is a real valued continuous function defined on $\sigma(A)$ such that $f(t) \geq 0$, for all $t \in \sigma(A)$, then $f(A) \geq 0$, i.e. $f(A)$ is a positive operator.

The following order preserving property is a consequence of the continuous functional calculus.

Lemma 1.5.3.: *Let $A \in \mathcal{B}_h(H)$ and let $f,\, g \in C(\sigma(A))$.*

$$\textit{If } f(t) \geq g(t), \textit{ for all } t \in \sigma(A), \textit{ then } f(A) \geq g(A). \tag{1.24}$$

Additionally, $f(A) = g(A)$ if and only if $f(t) = g(t)$, for all $t \in \sigma(A)$.

1.6. Operator monotone and operator convex functions

Denote with $C([m, M])$ the set of all real valued continuous functions on an interval $[m, M] \subset \mathbb{R}$.

Definition 1.6.16.: We say that the function $f \in C([m, M])$ is **operator monotone** if it preserves the operator order: if $A \leq B$ then $f(A) \leq f(B)$, for all self-adjoint operators A, B on a Hilbert space H, such that $\sigma(A), \sigma(B) \subseteq [m, M]$.

Definition 1.6.17.: We say that the function $f \in C([m, M])$ is **operator convex** if for all self-adjoint operators A, B on a Hilbert space H, such that $\sigma(A), \sigma(B) \subseteq [m, M]$ and for all $\lambda \in [0, 1]$,

$$f((1-\lambda)A + \lambda B) \leq (1-\lambda)f(A) + \lambda f(B). \tag{1.25}$$

We say that the function $f \in C([m, M])$ is **operator concave** if $-f$ is operator convex, that is if inequality (1.25) holds with the reverse sign.

Example 3.: *Function $f: \mathbb{R} \to \mathbb{R}$, $f(t) = \alpha + \beta t$ is operator monotone function, for all $\alpha \in \mathbb{R}$ and $\beta \geq 0$ and is operator convex for all $\alpha, \beta \in \mathbb{R}$, (see [54]).*

Example 4.: *Function $f: (0, \infty) \to \mathbb{R}$, $f(t) = -\frac{1}{t}$ is operator monotone on $(0, \infty)$ (see [54]).*

The following characterization holds.

Theorem 1.6.14: *Let $f: [0, \infty) \to [0, \infty)$ be a continuous function. Then f is operator monotone if and only if f is operator concave.*

The Löwner-Heinz inequality is a very important result which dates from 1934. For more details and the proof, see [54].

Theorem 1.6.15 ***(Löwner-Heinz inequality)*****:** *Let A and B be positive operators on a*

Hilbert space H. If $A \geq B \geq 0$, then $A^r \geq B^r$, for all $r \in [0,1]$.

In other words, function $f\colon [0,\infty) \to \mathbb{R}$, $f(t) = t^r$ is operator monotone on $[0,\infty)$, for all $r \in [0,1]$.

Remark 1.6.13.: If we observe different examples of the operator functions and compare their monotonicity and convexity on $\mathbb{R}$ with their operator monotonicity (convexity), we see that there are some aberrations. Thus for example, the function $f\colon [0,\infty) \to \mathbb{R}$, $f(t) = t^2$ isn't operator monotone, although it is a monotone real valued function when defined on $\mathbb{R}$. Similarly, the function $f\colon [0,\infty) \to \mathbb{R}$, $f(t) = t^3$ isn't operator convex, although it is convex as a real valued function defined on $\mathbb{R}$. For more examples the reader is referred to [54].

To introduce Jensen's type inequalities for operator convex functions, we need the following definition.

Definition 1.6.18.: A map $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ is linear if it is additive and homogeneous, i.e. $\Phi(\lambda X + \mu Y) = \lambda\Phi(X) + \mu\Phi(Y)$ for any $\lambda, \mu \in \mathbb{R}$ and $X, Y \in \mathcal{B}(H)$.
A linear map $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ is said to be positive (resp. strictly positive) if $\Phi(A) \geq 0$ (resp. $\Phi(A) > 0$) whenever A is positive (resp. strictly positive).
It is unital (normalized) if Φ preserves the identity operator, i.e. $\Phi(1_H) = 1_K$.

Remark 1.6.14.: A positive linear map $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ preserves order relation, that is, $A \leq B$ implies $\Phi(A) \leq \Phi(B)$, and preserves adjoint operation, that is, $\Phi(A^*) = \Phi(A)^*$. If $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ is unital, then $m1_H \leq A \leq M1_H$ implies $m1_K \leq \Phi(A) \leq M1_K$. It can be easily seen that a positive linear map Φ is strictly positive if and only if $\Phi(1_H) > 0$.

The following theorem shows the **Davis-Choi-Jensen inequality** for operator convex functions (see [86, Proposition 2.1]).

Theorem 1.6.16: *If $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ is a unital positive linear map and f is an* *operator convex function on an interval J, then*

$$f(\Phi(X)) \leqslant \Phi(f(X)) \tag{1.26}$$

for every self-adjoint operator X with the spectrum in J.

Additionally, if J containing 0 and $f(0) = 0$, then (1.26) holds for every positive linear mapping $\Phi : \mathcal{B}(H) \to \mathcal{B}(K)$ with $0 < \Phi(1_H) \leqslant 1_K$.

Regarding the previous considerations and among many existing variants of operator Jensen's inequalities, we also single out the one proved by B. Mond and J. Pečarić in [114].

Theorem 1.6.17*(Jensen's operator inequality)*: *Let $A_i \in \mathcal{B}(H), i = 1, ..., n$, be self-adjoint operators with the spectra in an interval J and $\Phi_i : \mathcal{B}(H) \to \mathcal{B}(K)$, $i = 1, \dots, n$, are unital positive linear mapping. Let $w_1, \dots, w_n$ be non-negative real numbers with $\sum_{i=1}^{n} w_i =$ 1. If f is operator convex function on J, then*

$$f\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) \leq \sum_{i=1}^{n} w_i \Phi_i\left(f(A_i)\right). \tag{1.27}$$

1.7. Connections and solidarities. Operator means

The theory of operator means for positive linear operators on a Hilbert space was established and for most part developed by T. Ando and F. Kubo, (see [97]). In this process they used the results from the Löwner's theory on operator monotone functions. In the monograph [54] one can find more details on this topic.

Operator means are defined via connections.

Definition 1.7.19.: A binary operation $(A, B) \in \mathcal{B}^+(H) \times \mathcal{B}^+(H) \to A\,\sigma B \in \mathcal{B}^+(H)$ in the cone of positive operators on a Hilbert space H is called a **connection** if the following conditions are satisfied:

(i) monotonicity: $A \leq C$ and $B \leq D$ imply $A\,\sigma\,B \leq C\,\sigma\,D$,

(ii) upper continuity: $A_n \downarrow A$ and $B_n \downarrow B$ imply $A_n\,\sigma\,B_n \downarrow A\,\sigma\,B$,

(iii) transformer inequality: $T^*(A\,\sigma\,B)T \le (T^*AT)\,\sigma\,(T^*BT)$, for every operator T.

If the normalized condition

(iv) $1_H\,\sigma\,1_H = 1_H$,

is satisfied, then we say that the connection is an operator mean.

$A_n \downarrow A$ in condition (ii) denotes the convergence of $\{A_n\}$, $A_n \in \mathcal{B}_h(H)$, $A_1 \ge A_2 \ge \cdots$, to $A \in \mathcal{B}_h(H)$, in the strong operator topology.

If T is an invertible operator, then in (iii) the equality sign holds.

Furthermore, the following homogeneity property holds:

$$\alpha(A\,\sigma\,B) = (\alpha A)\,\sigma\,(\alpha B), \quad \text{for all} \quad \alpha > 0.$$

Connections posses the property of joint concavity. More precisely, the inequality

$$(\lambda A_1 + (1-\lambda)B_1)\ \sigma\ (\lambda A_2 + (1-\lambda)B_2) \ge \lambda(A_1\,\sigma\,A_2) + (1-\lambda)(B_1\,\sigma\,B_2) \tag{1.28}$$

holds for $\lambda \in [0,1]$ and $A_1, A_2,\ B_1, B_2 \in \mathcal{B}^+(H)$.

Let us recall the basic operator means (see [97]).

Definition 1.7.20.: For $A, B \in \mathcal{B}^+(H)$ and $p \in [0,1]$, the p-**arithmetic means** is defined as

$$A\nabla_p B = (1-p)A + pB. \tag{1.29}$$

In particular, for $p = \frac{1}{2}$ we write $A\nabla B$ in place of $A\nabla_{1/2}B$.

Definition 1.7.21.: For $A \in \mathcal{B}^{++}(H)$, $B \in \mathcal{B}^+(H)$ and $p \in [0,1]$, the p-**geometric mean** is defined by

$$A\sharp_p B = A^{1/2}(A^{-1/2}BA^{-1/2})^p A^{1/2}. \tag{1.30}$$

For $p = \frac{1}{2}$ we use the symbol $A\sharp B$ instead of $A\sharp_{1/2}B$.

Particularly, as a special class of means deduced from the operator geometric mean, we may observe **the weighted operator Heinz means** defined by

$$H_p(A,B) = \frac{A\sharp_p B + A\sharp_{1-p} B}{2}. \tag{1.31}$$

Definition 1.7.22.: **Parallel sum** of operators A, $B \in \mathcal{B}^{++}(H)$ is defined by

$$A : B = \left(A^{-1} + B^{-1}\right)^{-1}. \tag{1.32}$$

Definition 1.7.23.: For $A, B \in \mathcal{B}^{++}(H)$ and $p \in [0,1]$, the p-**harmonic mean** is defined as

$$A\,!_p\,B = \left((1-p)A^{-1} + pB^{-1}\right)^{-1}. \tag{1.33}$$

For $p = \frac{1}{2}$ we use the symbol $A!B$ instead of $A!_{1/2}B$.

The linear combination of two connections is defined as follows: if σ, τ are connections and a, b are nonnegative real numbers, then

$$A\,(a\sigma + b\tau)\,B = a(A\,\sigma\,B) + b(A\,\tau\,B). \tag{1.34}$$

Thus, in particular the class of operator means is a convex set.

Furthermore, respecting the order properties, if $\sigma \geq \tau$ then $A\,\sigma\,B \geq A\,\tau\,B$, for all A, $B \in \mathcal{B}^{+}(H)$. Hence, for $A, B \in \mathcal{B}^{++}(H)$ and $p \in [0,1]$ **the operator arithmetic-geometric-harmonic operator inequality (A-G-H operator inequality)** holds:

$$A\,!_p\,B \leq A\,\sharp_p B \leq A\,\nabla_p\,B. \tag{1.35}$$

We also present definitions of some relative operator entropies (see [75, p. 3149]).

Definition 1.7.24.: For $A, B \in \mathcal{B}^{++}(H)$, **the relative operator entropy** is defined by

$$S(A,B) = A^{1/2}\log(A^{-1/2}BA^{-1/2})A^{1/2}. \tag{1.36}$$

Definition 1.7.25.: For $A, B \in \mathcal{B}^{++}(H)$ and $p \in \mathbb{R}$, **the generalized relative opera-**

tor entropy is given by

$$S_p(A,B) = A^{1/2}(A^{-1/2}BA^{-1/2})^p \log(A^{-1/2}BA^{-1/2})A^{1/2}. \tag{1.37}$$

Definition 1.7.26.: For $A, B \in \mathcal{B}^{++}(H)$ and $0 < p \leq 1$, **the Tsallis relative operator entropy** is defined as follows

$$T_p(A,B) = \frac{A\sharp_p B - A}{p} = A^{1/2}\frac{(A^{-1/2}BA^{-1/2})^p - I}{p}A^{1/2}. \tag{1.38}$$

Definition 1.7.27.: For $A, B \in \mathcal{B}^{++}(H)$ and $0 < s \leq 1$, $-1 \leq r \leq 1$, **the parametric Tsallis relative operator entropy** is defined by

$$T_{s,r}(A,B) = A^{1/2}\frac{[(1-s)I + s(A^{-1/2}BA^{-1/2})^r]^{1/r} - I}{s}A^{1/2}. \tag{1.39}$$

The basic result of the theory developed by F. Kubo and T. Ando concerns the isomorphism between connections and the *nonnegative* operator monotone functions.

Theorem 1.7.18*(Kubo-Ando)*: *Let σ be a connection. If for a function $f\colon [0,\infty) \to \mathbb{R}$ and $t \geq 0$*

$$f(t)1_H = 1_H\,\sigma\,(t1_H), \tag{1.40}$$

then f is nonnegative and operator monotone function on $[0,\infty)$ and the following statements are true:

(i) There is an isomorphism between the classes of all connections and all nonnegative operator monotone functions on $[0,\infty)$. Furthermore, for all $A, B \in \mathcal{B}^{+}(H)$ and $t \geq 0$ is

$$A\,\sigma_1\,B \leq A\,\sigma_2\,B \text{ if and only if } f_1(t) \leq f_2(t),$$

where $\sigma_1 \mapsto f_1$, and $\sigma_2 \mapsto f_2$ are the isomorphisms.

(ii) If A is an invertible operator, then $A\,\sigma\,B = A^{\frac{1}{2}}f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}$.

(iii) The connection σ is an operator mean if and only if f is a normalized function, i.e. $f(1) = 1$.

The operator monotone function f is called a *representation function* for the connection σ. Since by Theorem 1.6.14 a continuous function $f\colon [0,\infty) \to [0,\infty)$ is operator monotone if and only if f is operator concave, it follows that the representation function f is an operator concave function.

Hence we have the following representation functions for the arithmetic, harmonic and geometric mean: $f_{\nabla}(t) = \frac{1+t}{2}$, $f_{!}(t) = \frac{2t}{1+t}$ and $f_{\sharp}(t) = \sqrt{t}$. Thus inequalities $f_{\nabla}(t) \geq f_{\sharp}(t) \geq f_{!}(t)$ imply the corresponding operator mean inequalities: $\nabla \geq \sharp \geq !$.

Operator means posses two important properties.

Theorem 1.7.19: *Operator mean σ possesses the property of subadditivity, that is*

$$A\,\sigma\,C + B\,\sigma\,D \leq (A+B)\,\sigma\,(C+D), \tag{1.41}$$

as well as of the joint concavity:

$$\lambda(A\,\sigma\,C) + (1-\lambda)(B\,\sigma\,D) \leq (\lambda A + (1-\lambda)B)\,\sigma\,(\lambda C + (1-\lambda)D), \tag{1.42}$$

where $A, B, C, D \in \mathcal{B}^+(H)$, $\lambda \in [0,1]$.

J. Fujii et al. generalized in [50] the results on connections from Kubo-Ando theory. They investigated the binary operation $s = s_f$ for an *arbitrary* operator monotone function f defined on $[0,\infty)$:

$$A\,s\,B = A^{\frac{1}{2}} f(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}},$$

analogously as in (ii) in Theorem 1.7.18. Domain of s_f is the set of all ordered pairs (A,B) for which $A\,s\,B$ is a bounded operator, respecting f.

Definition 1.7.28.: A binary operation $(A,B) \in \mathcal{D}_s \subseteq \mathcal{B}^+(H) \times \mathcal{B}^+(H) \to (A\,s\,B) \in \mathcal{B}_h(H)$ is called a **solidarity** if it has the following properties:

(i) $B \leq C$ implies $A\,s\,B \leq A\,s\,C$,

(ii) $B_n \downarrow B$ implies $A\,s\,B_n \downarrow A\,s\,B$,

(iii) $A_n \to A$ implies $A_n\,s\,1_H \to A\,s\,1_H$,

(iv) $T^*(A\,s\,B)T \le (T^*AT)\,s\,(T^*BT)$, for all operators T,

where $\mathcal{D}_s$ is the domain of s.

Solidarity s is defined for an arbitrary pair of positive invertible operators, but not for an arbitrary pair of positive operators. Thus $\mathcal{D}_s$ denotes the maximal subset of $\mathcal{B}^+(H)\times\mathcal{B}^+(H)$ on which s is a bounded operator. For more details on the domain $\mathcal{D}_s$ the reader is referred to [50].

In [50] the authors also proved that there was an isomorphism between the solidarities and the operator monotone functions. Furthermore, solidarities posses the joint concavity property. More precisely,

$$(\lambda A_1 + (1-\lambda)B_1)\ s\ (\lambda A_2 + (1-\lambda)B_2) \ge \lambda(A_1\,s\,A_2) + (1-\lambda)(B_1\,s\,B_2), \tag{1.43}$$

where $A_1, A_2,\ B_1, B_2 \in \mathcal{B}^{++}(H)$ and $\lambda \in [0,1]$ and under the assumption that all of the operators included exist as the bounded operators.

Relative operator entropy is a common example among solidarities. More on this topic can be found in [52].

CHAPTER 2.

Sherman type theorems and applications

2.1. About majorization

In this section, we introduce the concepts of majorization and Schur-convexity in order to give some basic results from the Theory of majorization that give an important characterization of convex functions.

For fixed $n \geq 2$, let $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_n)$ denote two real n-tuples. Let

$$x_{[1]} \geq x_{[2]} \geq ... \geq x_{[n]}, \quad y_{[1]} \geq y_{[2]} \geq ... \geq y_{[n]},$$

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}, \quad y_{(1)} \leq y_{(2)} \leq ... \leq y_{(n)}$$

be their ordered components.

Definition 2.1.1. ***(Majorization, [138, p. 319])*****:** We say that $\mathbf{x}$ **majorizes** $\mathbf{y}$ or $\mathbf{y}$ is **majorized** by $\mathbf{x}$ and write

$$\mathbf{y} \prec \mathbf{x}, \tag{2.1}$$

if

$$\begin{aligned} \sum_{i=1}^{m} y_{[i]} &\leq \sum_{i=1}^{m} x_{[i]}, \quad m = 1, 2, ..., n-1, \\ \sum_{i=1}^{n} y_i &= \sum_{i=1}^{n} x_i. \end{aligned} \tag{2.2}$$

Remark 2.1.1.: Note that (2.2) is equivalent to

$$\sum_{i=n-m+1}^{n} y_{(i)} \leq \sum_{i=n-m+1}^{n} x_{(i)}, \quad m = 1, 2, ..., n-1.$$

Example 1.: *If* $\mathbf{x} = (x_1, ..., x_n)$ *is an arbitrary vector with* $x_i \geq 0, i = 1, ..., n$, *and* $\sum_i^n x_i = 1$, *then*

$$\left(\frac{1}{n}, ..., \frac{1}{n}\right) \prec \mathbf{x} \prec (1, 0, ..., 0).$$

This definition defines relation which is reflexive and transitive but it is not antisymmetric (see [106, p. 13]). However, we can consider the set $D = \{\mathbf{x} \in \mathbb{R}^n : x_1 \geq x_2 \geq ... \geq x_n\}$ on which the relation "$\prec$" is a partial order. In assertions or proofs, we can assume that the components of a vector are just arranged in descending order without loss of generality.

It's important to remember that two vectors may not have any majorization relationship, meaning that there are vectors that can not be compared in the sense of the given definition.

Example 2.: *The vectors* $\mathbf{x} = (0.6, 0.6, 0.2)$ *and* $\mathbf{y} = (0.5, 0.4, 0.1)$ *are incomparable in the sense of the previous definition, i.e. no* $\mathbf{y} \prec \mathbf{x}$ *neither* $\mathbf{x} \prec \mathbf{y}$ *are valid.*

The notion of majorization is closely related to Schur-convexity which characterizes functions that preserve preordering $\prec$ on $\mathbb{R}^n$. Exactly, the following notion of Schur-convexity generalizes the definition of convex function via the notion of majorization.

Definition 2.1.2.: A function $F : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is called **Schur-convex** on S if

$$F(\mathbf{y}) \leq F(\mathbf{x}) \tag{2.3}$$

for every $\mathbf{x}$, $\mathbf{y} \in S$ such that

$$\mathbf{y} \prec \mathbf{x}.$$

A function F is called **Schur-concave** on S, if $-F$ is Schur-convex on S.

Issai Schur (see [147], [106, p. 79]) first systematically studied functions that preserve the order of majorization and called them convex functions. Such functions are now called Schur convex functions and convex functions are often referred to convex functions in the sense of Jensen. Each function that is convex and symmetric is also Schur-convex. The opposite implication is not valid. However, all Schur convex functions are symmetric.

The concept of majorization, together with the concept of Schur convexity, provides an important characterization of convex functions and is a powerful and useful mathematical tool that has wide application in many applied sciences.

There are several equivalent characterizations of a majorization relation $\mathbf{x} \succ \mathbf{y}$ in addition to the conditions given in definition of majorization. One is actually the answer of a question posed and answered in [60, 61], the well-known as Majorization theorem. A convenient reference for its proof is given by Marshall and Olkin in [106, p.11] (see also [138, p.320]).

Theorem 2.1.1 ***(Majorization theorem)*****:** *Let I be an real interval and* $\mathbf{x}$,$\mathbf{y}$ *be two n-tuples such that x_i, $y_i \in I$ $(i = 1, ..., n)$. Then*

$$\sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} f(x_i) \tag{2.4}$$

holds for every continuous convex function $f : I \to \mathbb{R}$ iff $\mathbf{y} \prec \mathbf{x}$.

If f is a strictly convex function, then equality in (2.4) is valid iff $x_{[i]} = y_{[i]}$, $i = 1, ..., n$.

The previous result is known as **Majorization inequality**. It can rephrase in the form of the following theorem which gives a relation between one-dimensional convex functions and n-dimensional Schur-convex functions.

Theorem 2.1.2: *Let I be an real interval and $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_n) \in I^n$. Let*

$f : I \to \mathbb{R}$ be a continuous function. Then a function $F : I^n \to \mathbb{R}$, defined by

$$F(\mathbf{x}) = \sum_{i=1}^{n} f(x_i),$$

is Schur-convex on I^n iff f is convex on I.

The concept of majorization admits an order-free characterization based on the notion of doubly stochastic matrix. The definition of stochastic matrices is given below.

Definition 2.1.3.: A $n \times m$ real matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{nm}(\mathbb{R})$ is called **column stochastic** if all of its entries are greater or equal to zero, i.e. $a_{ij} \geq 0$ for $i = 1, ..., n$, $j = 1, ..., m$ and the sum of the entries in each column is equal to 1, i.e. $\sum\limits_{i=1}^{n} a_{ij} = 1$ for $j = 1, ..., m$.

On the other hand, a matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{nm}(\mathbb{R})$ is called **row stochastic** if all of its entries are greater or equal to zero, i.e. $a_{ij} \geq 0$ for $i = 1, ..., n$, $j = 1, ..., m$ and the sum of the entries in each row is equal to 1, i.e. $\sum\limits_{j=1}^{m} a_{ij} = 1$ for $i = 1, ..., n$.

A square matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_m(\mathbb{R})$ is called **double stochastic** if both $\mathbf{A}$ and its transpose $\mathbf{A}^T = (a_{ji})$ are row stochastic. In other words, $\mathbf{A} = (a_{ij}) \in \mathcal{M}_m(\mathbb{R})$ is called *double stochastic* if all of its entries are greater or equal to zero, i.e. $a_{ij} \geq 0$ for $i, j = 1, ..., m$, and the sum of the entries in each column and each row is equal to 1, i.e. $\sum\limits_{i=1}^{m} a_{ij} = 1$ for $j = 1, ..., m$ and $\sum\limits_{j=1}^{m} a_{ij} = 1$ for $i = 1, ..., m$.

An interesting characterization of the concept of majorization is the well known Rado's theorem [106, p. 33]: for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is valid

$$\mathbf{y} \prec \mathbf{x} \quad \text{iff} \quad \mathbf{y} = \mathbf{x}\mathbf{A}$$

for some double stochastic matrix $\mathbf{A}$.

In fact, the previous characterization implies that the set of vectors $\mathbf{y}$ that satisfy $\mathbf{y} \prec \mathbf{x}$ is the convex hull spanned by the $n!$ points formed from the permutations of the elements of $\mathbf{y}$.

By Birkhoff's theorem [106, p. 30], each $n \times n$ doubly stochastic matrix $\mathbf{A}$ has the form

$$\mathbf{A} = \sum_{i=1}^{k} \alpha_i \mathbf{P}_i$$

for some positive integer k, $n \times n$ permutation matix $\mathbf{P}_i$, $i = 1, 2, ..., k$, and $\alpha_i \geq 0$, with $\sum_{i=1}^{k} \alpha_i = 1$.

Taking into account previous observations, Majorization theorem can be summarized in the form of the following theorem.

Theorem 2.1.3*(Hardy-Littlewood-Pólya [60, 61])*: *Let I be an real interval and $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_n) \in I^n$. Then the following statements are equivalent:*

(i) $\mathbf{y} \prec \mathbf{x}$ *;*

(ii) There is a doubly stochastic matrix $\mathbf{A}$ such that $\mathbf{y} = \mathbf{x}\mathbf{A}$;

(iii) The inequality

$$\sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} f(x_i) \tag{2.5}$$

holds for each continuous convex function $f : I \to \mathbb{R}$.

Remark 2.1.2.: For many purposes, the condition $\mathbf{y} = \mathbf{x}\mathbf{A}$ is more convenient than the partial sums conditions defining majorization.

Majorization theorems on convex functions and standard majorization, due to Hardy, Littlewood, Pólya and Karamata, have numerous applications in many disciplines of mathematics and applied sciences (see the monograph [106]). In this chapter, we develop similar results involving the concept of weighted majorization which we introduce in the sequel. Material of this chapter is quoted mainly from [17, 64, 65, 72, 124, 128, 130, 132, 133].

2.2. Sherman's theorem and its converse

Sherman [149] considered the weighted concept of majorization for two pairs $(\mathbf{x}, \mathbf{a})$ and $(\mathbf{y}, \mathbf{b})$ which consist of vectors $\mathbf{x} = (x_1, ..., x_n) \in I^n$, $\mathbf{y} = (y_1, ..., y_m) \in I^m$, where I is a real interval, and corresponding nonnegative weights $\mathbf{a} = (a_1, ..., a_n)$ and $\mathbf{b} = (b_1, ..., b_m)$.

The concept of weighted majorization is defined by assumption of existence of column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nm}(\mathbb{R})$ such that

$$\mathbf{y} = \mathbf{x}\mathbf{S} \quad \text{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T, \tag{2.6}$$

i.e.

$$y_j = \sum_{i=1}^{n} x_i s_{ij}, \quad j = 1, ..., m \quad \text{and} \quad a_i = \sum_{j=1}^{m} b_j s_{ij}, \quad i = 1, ..., n.$$

For such pairs $(\mathbf{y}, \mathbf{b})$ and $(\mathbf{x}, \mathbf{a})$, under assumption of existence of column stochastic matrix $\mathbf{S}$ such that (2.6) holds, we use the notation

$$(\mathbf{y}, \mathbf{b}) \prec (\mathbf{x}, \mathbf{a}) \tag{2.7}$$

and say that the pair $(\mathbf{y}, \mathbf{b})$ is **weighted majorized** by $(\mathbf{x}, \mathbf{a})$. It presents a generalization of the classical concept of majorization in notation $\mathbf{y} \prec \mathbf{x}$.

Instead the assumptions (2.6), we may assume

$$\mathbf{y} = \mathbf{x}\mathbf{S}^T \text{ and } \mathbf{a} = \mathbf{b}\mathbf{S} \tag{2.8}$$

for some row stochastic matrix $\mathbf{S}$.

Under the previous assumptions Sherman [149] proved a useful generalization of Majorization theorem (Theorem 2.1.1, Theorem 2.1.3). We quote Sherman's result in the following form.

Theorem 2.2.4 ***(Sherman's inequality)*****:** *Let $f : I \to \mathbb{R}$ be a convex function. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in I^n$, $\mathbf{y} = (y_1, y_2, \ldots, y_m) \in I^m$, $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in [0, \infty)^n$ and $\mathbf{b} = (b_1, b_2, \ldots, b_m) \in [0, \infty)^m$. If*

$$\mathbf{y} = \mathbf{x}\mathbf{S} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{2.9}$$

holds for some $n \times m$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nm}(\mathbb{R})$, then

$$\sum_{j=1}^{m} b_j f(y_j) \le \sum_{i=1}^{n} a_i f(x_i). \tag{2.10}$$

Proof: *Using the assumptions (2.9) and the fact that f is convex function, we have*

$$\begin{aligned}
\sum_{j=1}^{m} b_j f(y_j) &= \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} x_i s_{ij}\right) \\
&\le \sum_{j=1}^{m} b_j \sum_{i=1}^{n} s_{ij} f(x_i) \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} b_j s_{ij}\right) f(x_i) = \sum_{i=1}^{n} a_i f(x_i),
\end{aligned}$$

what we need to finish the proof. □

From now, on the inequality (2.10) we refer as **Sherman's inequality.** It is interesting that a reverse implication is valid too (for details see [25, 149] and also [31, 126]).

Remark 2.2.3.: It is worth emphasizing that if $\mathbf{S}$ is doubly stochastic and $\mathbf{b} = (1, 1, \ldots, 1) \in \mathbb{R}^m$, then (2.9) implies (2.4), because $\mathbf{S}^T$ is column stochastic and, in consequence,

$$\mathbf{a} = \mathbf{b}\mathbf{S}^T = (1, 1, \ldots, 1)\mathbf{S}^T = (1, 1, \ldots, 1).$$

In this situation, Sherman's inequality reduces to the fact that the function

$$(x_1, \ldots, x_n) \mapsto \sum_{j=1}^{n} f(x_j) \quad \text{for } (x_1, \ldots, x_n) \in I^n$$

is Schur-convex on I^n.

Also, choosing $n = 1$ and $\mathbf{b} = [1]$, Sherman's inequality (2.10) reduces to Jensen's inequality (1.6).

Next we present the converse of Sherman's inequality (2.10).

Theorem 2.2.5 ***(Converse Sherman's inequality, [17]):*** *Let $f : I \to \mathbb{R}$ be a convex* *function. Let* $\alpha, \beta \in I$, $\alpha < \beta$, $\mathbf{x} = (x_1, ..., x_n) \in [\alpha, \beta]^n$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_n) \in [0, \infty)^n$ *and* $\mathbf{b} = (b_1, ..., b_m) \in [0, \infty)^m$ *be such that* (2.9) *holds for some column stochastic matrix* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nm}(\mathbb{R})$*. Then*

$$\sum_{j=1}^{m} b_j f(y_j) \le \sum_{i=1}^{n} a_i f(x_i) \le \sum_{j=1}^{m} b_j \frac{f(\alpha)(\beta - y_j) + f(\beta)(y_j - \alpha)}{\beta - \alpha}. \tag{2.11}$$

Proof: Under the assumptions, Sherman's inequality (2.10) holds. Further, from (1.7), setting $p_i = s_{ij}$, for $i = 1, ..., n$, we have

$$\begin{aligned}
\sum_{j=1}^{m} b_j f(y_j) &\le \sum_{i=1}^{n} a_i f(x_i) \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} b_j s_{ij} \right) f(x_i) \\
&= \sum_{j=1}^{m} b_j \left(\sum_{i=1}^{n} s_{ij} f(x_i) \right) \\
&\le \sum_{j=1}^{m} b_j \left(\frac{\beta - \sum\limits_{i=1}^{n} x_i s_{ij}}{\beta - \alpha} f(\alpha) + \frac{\sum\limits_{i=1}^{n} x_i s_{ij} - \alpha}{\beta - \alpha} f(\beta) \right),
\end{aligned}$$

which is what was to be proved. □

Remark 2.2.4.: Choosing $n = m$ and $\mathbf{b} = \mathbf{e} = (1, ..., 1)$, the condition $(\mathbf{y}, \mathbf{b}) \prec (\mathbf{x}, \mathbf{a})$ implies $\mathbf{y} = \mathbf{xS}$ with some doubly stochastic matrix $\mathbf{S}$ and $\mathbf{a} = \mathbf{bS}^T = \mathbf{e} = (1, ..., 1)$, and as direct consequence of the previous result we get Majorization inequality and **its converse**

$$\begin{aligned}
\sum_{j=1}^{n} f(y_j) \;\le\; \sum_{i=1}^{n} f(x_i) &\le \sum_{j=1}^{n} \frac{f(\alpha)(\beta - y_j) + f(\beta)(y_j - \alpha)}{\beta - \alpha} \\
&= \frac{\bar{\beta} - \sum_{j=1}^{n} y_j}{\beta - \alpha} f(\alpha) + \frac{\sum_{j=1}^{n} y_j - \bar{\alpha}}{\beta - \alpha} f(\beta) \\
&= \frac{\bar{\beta} - \sum_{i=1}^{n} x_i}{\beta - \alpha} f(\alpha) + \frac{\sum_{i=1}^{n} x_i - \bar{\alpha}}{\beta - \alpha} f(\beta)
\end{aligned} \tag{2.12}$$

where $\bar{\beta} = m\beta$ and $\bar{\alpha} = m\alpha$ and because of (2.9) we have $\sum_{j=1}^{n} y_j = \sum_{j=1}^{n} \sum_{i=1}^{n} x_i s_{ij} = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} s_{ij} = \sum_{i=1}^{n} x_i$.

2.3. Sherman's functional with estimations

In this section, we define Sherman's functional deduced from Sherman's inequality (2.10). We present some of its essential properties. We also establish lower and upper bounds for it. In proving those results, we use the following interpolating inequalities for Jensen's functional proved by J. Pečarić (see [111, p. 717]), which we state with proofs due to their importance. Results from this section are quoted from [72].

Theorem 2.3.6: *Let $f : I \to \mathbb{R}$ be a convex function, $\boldsymbol{x} = (x_1, \ldots, x_n) \in I^n$, $\boldsymbol{p} = (p_1, \ldots, p_n)$ and $\boldsymbol{q} = (q_1, \ldots, q_n)$ be nonnegative n-tuples such $\boldsymbol{p} \geq \boldsymbol{q}$, i.e. $p_i \geq q_i, i = 1, ..., n$, and $P_n = \sum_{i=1}^n p_i > 0$, $Q_n = \sum_{i=1}^n q_i > 0$. Then*

$$\sum_{i=1}^{n} q_i f\left(x_i\right) - Q_n f\left(\frac{1}{Q_n} \sum_{i=1}^{n} q_i x_i\right) \leq \sum_{i=1}^{n} p_i f\left(x_i\right) - P_n f\left(\frac{1}{P_n} \sum_{i=1}^{n} p_i x_i\right). \tag{2.13}$$

Proof: *In the case $\boldsymbol{p} = \boldsymbol{q}$, the assertion holds trivially. So, without lost of generality we can assume that $\boldsymbol{p} \neq \boldsymbol{q}$. Hence $\sum\limits_{i=1}^{n} (p_i - q_i) > 0$ since $\boldsymbol{p} \geq \boldsymbol{q}$.*

Applying Jensen's inequality we have

$$\begin{aligned} &\sum_{i=1}^{n} (p_i - q_i) f\left(\frac{1}{\sum\limits_{i=1}^{n} (p_i - q_i)} \sum_{i=1}^{n} (p_i - q_i) x_i\right) \\ &\leq \sum_{i=1}^{n} (p_i - q_i) \frac{1}{\sum\limits_{i=1}^{n} (p_i - q_i)} \sum_{i=1}^{n} (p_i - q_i) f(x_i) = \sum_{i=1}^{n} (p_i - q_i) f(x_i). \end{aligned} \tag{2.14}$$

On the other side, applying the reverse of Jensen's inequality we get

$$\begin{aligned}
&\sum_{i=1}^{n}(p_i-q_i)\, f\left(\frac{1}{\sum\limits_{i=1}^{n}(p_i-q_i)}\sum_{i=1}^{n}(p_i-q_i)\,x_i\right) \qquad (2.15)\\
&= (P_n-Q_n)\, f\left(\frac{P_n\left(\frac{1}{P_n}\sum\limits_{i=1}^{n}p_ix_i\right)-Q_n\left(\frac{1}{Q_n}\sum\limits_{i=1}^{n}q_ix_i\right)}{P_n-Q_n}\right)\\
&\geq (P_n-Q_n)\left(\frac{P_n}{P_n-Q_n}f\left(\frac{1}{P_n}\sum_{i=1}^{n}p_ix_i\right)-\frac{Q_n}{P_n-Q_n}f\left(\frac{1}{Q_n}\sum_{i=1}^{n}q_ix_i\right)\right)\\
&= P_nf\left(\frac{1}{P_n}\sum_{i=1}^{n}p_ix_i\right)-Q_nf\left(\frac{1}{Q_n}\sum_{i=1}^{n}q_ix_i\right).
\end{aligned}$$

Now, combining (2.14) and (2.15), we obtain (2.13). □

As easy consequences of the previous theorem the following lemmas hold.

Lemma 2.3.1.: *Let $f : I \to \mathbb{R}$ be a convex function. Let $\boldsymbol{x} = (x_1, \ldots, x_n) \in I^n$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$, $\boldsymbol{q} = (q_1, \ldots, q_n)$ be positive n-tuples with $P_n = \sum_{i=1}^n p_i$, $Q_n = \sum_{i=1}^n q_i$ and $\bar{x}_P = \frac{1}{P_n}\sum_{i=1}^n p_ix_i$, $\bar{x}_Q = \frac{1}{Q_n}\sum_{i=1}^n q_ix_i$. Then*

$$\begin{aligned}
0 &\leq \min_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n}q_if(x_i)-Q_nf(\bar{x}_Q)\right)\leq\sum_{i=1}^{n}p_if(x_i)-P_nf(\bar{x}_P)\\
&\leq \max_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n}q_if(x_i)-Q_nf(\bar{x}_Q)\right). \qquad (2.16)
\end{aligned}$$

Proof: *Consider n-tuples $\boldsymbol{p} = (p_1, ..., p_n)$ and $\boldsymbol{q} = (q_1, ..., q_n)$. Note that*

$$\boldsymbol{p} = (p_1, ..., p_n) = \left(q_1\frac{p_1}{q_1}, \ldots, q_n\frac{p_n}{q_n}\right).$$

Let us denote $d_* = \min_{1\leqslant i\leqslant n}\left\{\frac{p_i}{q_i}\right\}$. We define

$$\boldsymbol{q}_{\min} = (q_1d_*, \ldots, q_nd_*).$$

We calculate $\bar{x}_Q = \frac{1}{\sum_{i=1}^n q_id_*}\sum_{i=1}^n q_id_*x_i = \frac{1}{Q_n}\sum\limits_{i=1}^{n} q_ix_i$.

It is obviously $\boldsymbol{q}_{\min} \leq \boldsymbol{p}$. *Applying Theorem 2.3.6 to* n*-tuples* $\boldsymbol{q}_{\min}$ *and* $\mathbf{p}$, *we get*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} q_i d_* f(x_i) - d_* Q_n f(\bar{x}_Q) - c d_* \sum_{i=1}^{n} q_i (x_i - \bar{x}_Q)^2 \\ &\leq \sum_{i=1}^{n} p_i f(x_i) - P f(\bar{x}_P) - c \sum_{i=1}^{n} p_i (x_i - \bar{x}_P)^2, \end{aligned}$$

what is equivalent to the second inequality in (2.16).

If we consider $\boldsymbol{p}$ *and* n*- tuple* $\boldsymbol{q}_{\max} = (q_1 d^*, \ldots, q_n d^*)$, *where* $d^* = \max_{1\leqslant i\leqslant n}\left\{\frac{p_i}{q_i}\right\}$, *then* $\boldsymbol{p} \leq \boldsymbol{q}_{\max}$. *Applying Theorem 2.3.6 to* n*-tuples* $\mathbf{p}$ *and* $\boldsymbol{q}_{\max}$ *we get the third inequality in (2.16).* □

Lemma 2.3.2.: *Let* $f : I \to \mathbb{R}$ *be a convex function,* $\boldsymbol{x} = (x_1, \ldots, x_n) \in I^n$ *and* $\boldsymbol{p} = (p_1, \ldots, p_n)$ *be nonnegative* n*-tuple such that* $P_n = \sum_{i=1}^{n} p_i > 0$. *Then*

$$\begin{aligned} \min_{1\leqslant i\leqslant n}\{p_i\}\left[\sum_{i=1}^{n} f(x_i) - n f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right] &\leq \sum_{i=1}^{n} p_i f(x_i) - P_n f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \qquad (2.17) \\ &\leq \max_{1\leqslant i\leqslant n}\{p_i\}\left[\sum_{i=1}^{n} f(x_i) - n f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right]. \end{aligned}$$

Proof: *Let us consider* n*-tuple* $\boldsymbol{p} = (p_1, \ldots, p_n)$ *and a constant* n*- tuple* $\mathbf{p}_{\min} = (\min_{1\leqslant i\leqslant n}\{p_i\}, \ldots, \min_{1\leqslant i\leqslant}$ *Then* $\boldsymbol{p} \geq \mathbf{p}_{\min}$. *Applying Theorem 2.3.6 we have*

$$\begin{aligned} &\sum_{i=1}^{n} p_i f(x_i) - P_n f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \\ \geq\ & \sum_{i=1}^{n} \min_{1\leqslant i\leqslant n}\{p_i\} f(x_i) - \sum_{i=1}^{n} \min_{1\leqslant i\leqslant n}\{p_i\} f\left(\frac{1}{\sum_{i=1}^{n} \min_{1\leqslant i\leqslant n}\{p_i\}} \sum_{i=1}^{n} \min_{1\leqslant i\leqslant n}\{p_i\} x_i\right) \\ =\ & \min_{1\leqslant i\leqslant n}\{p_i\} \sum_{i=1}^{n} f(x_i) - n \cdot \min_{1\leqslant i\leqslant n}\{p_i\} f\left(\frac{\min_{1\leqslant i\leqslant n}\{p_i\}}{n \cdot \min_{1\leqslant i\leqslant n}\{p_i\}} \sum_{i=1}^{n} x_i\right) \\ =\ & \min_{1\leqslant i\leqslant n}\{p_i\}\left[\sum_{i=1}^{n} f(x_i) - n f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right], \end{aligned}$$

i.e. we get the first inequality in (2.17). The second inequality can be obtained analogously by exchanging the roles of min and max . □

Remark 2.3.5.: Note that we can get (2.17) directly from (2.16) by choosing n-tuple $\boldsymbol{q}$ with $q_i = \frac{1}{n}$, $i = 1, ..., n$.

Few years later, Dragomir et el. [42], obtained the analogous result as in Theorem 2.3.6 but as a consequence of a quite different approach, via superadditivity. They considered Jensen's functional $J(f, \mathbf{x}, \mathbf{p})$, deduced from Jensen's inequality (1.6), defined as follows.

Definition 2.3.4. *([42])***:** For a function $f : I \to \mathbb{R}$, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in I^n$ and $\mathbf{p} \in \mathcal{P}_n^0$, where

$$\mathcal{P}_n^0 = \left\{ \mathbf{p} = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n : p_i \geq 0, \;\; P_n = \sum_{i=1}^{n} p_i > 0 \right\},$$

the Jensen functional is given by

$$J(f, \mathbf{x}, \mathbf{p}) = \sum_{i=1}^{n} p_i f(x_i) - P_n f\left(\frac{1}{P_n} \sum_{i=1}^{n} p_i x_i \right). \tag{2.18}$$

From (2.18), for $\mathbf{p} = \mathbf{e} = (1, ..., 1)$, we get

$$J(f, \mathbf{x}, \mathbf{e}) = \sum_{i=1}^{n} f\left(x_i\right) - n f\left(\frac{1}{n} \sum_{i=1}^{n} x_i \right). \tag{2.19}$$

Dragomir [39] continued the investigation of the properties for normalized Jensen's functional, i.e. functional (2.18) satisfying $P_n = 1$,

$$\bar{J}(f, \mathbf{x}, \mathbf{p}) = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i \right). \tag{2.20}$$

Using different technique in proving, he obtained analogous results as in Lemma 2.3.1. and Lemma 2.3.2. in the form of lower and upper bound for the normalized Jensen's functionals (2.20) as follows

$$(0 \leq) \min_{1 \leqslant i \leqslant n} \left\{ \frac{p_i}{q_i} \right\} \bar{J}(f, \mathbf{x}, \mathbf{q}) \leq \bar{J}(f, \mathbf{x}, \mathbf{p}) \leq \max_{1 \leqslant i \leqslant n} \left\{ \frac{p_i}{q_i} \right\} \bar{J}(f, \mathbf{x}, \mathbf{q}), \tag{2.21}$$

and

$$(0 \leq) \min_{1 \leqslant i \leqslant n} \{p_i\} J(f, \mathbf{x}, \mathbf{e}) \leq \bar{J}(f, \mathbf{x}, \mathbf{p}) \leq \max_{1 \leqslant i \leqslant n} \{p_i\} J(f, \mathbf{x}, \mathbf{e}). \tag{2.22}$$

We also recommend to the reader the monograph [96], in which results related to the concept of the Jensen functional are collected.

2.3.1. Sherman's functional and its properties

In this subsection, we denote with $\mathcal{C}_{nm}([0,\infty))$ the set of column stochastic matrices of size $n \times m$. Specially, with $\mathcal{C}_{nm}((0,\infty))$ we denote the set of column stochastic matrices with positive entries.

Motivated by Sherman's inequality (2.10), i.e.

$$\begin{aligned} 0 &\leqslant \sum_{i=1}^{n} a_i f(x_i) - \sum_{j=1}^{m} b_j f(y_j) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} b_j s_{ij} \right) f(x_i) - \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} x_i s_{ij} \right), \end{aligned}$$

we define Sherman's functional as follows.

Let $\mathcal{F}(I)$ denotes the set of all convex functions defined on I. Under the assumptions of Sherman's theorem, we define **Sherman's functional** $\mathcal{S} : \mathcal{F}(I) \times I^n \times [0,\infty)^m \times \mathcal{C}_{nm}([0,\infty)) \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S}) &= \sum_{i=1}^{n}\sum_{j=1}^{m} b_j s_{ij} f(x_i) - \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} x_i s_{ij} \right) \\ &= \sum_{j=1}^{m} b_j \left(\sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} x_i s_{ij} \right) \right). \end{aligned} \tag{2.23}$$

We study properties of Sherman's functional.

Theorem 2.3.7 *([72])*: *Sherman's functional $\mathcal{S}$ satisfies the following properties:*

a) $\mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S})$ *is nonnegative, i.e.*

$$\mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S}) \geq 0; \tag{2.24}$$

b) $\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \cdot)$ *is concave on the convex set* $\mathcal{C}_{nm}([0,\infty))$, *i.e. for all* $\lambda \in [0,1]$

$$\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \lambda\mathbf{S} + (1-\lambda)\mathbf{T}) \geq \lambda\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \mathbf{S}) + (1-\lambda)\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \mathbf{T}); \tag{2.25}$$

c) $\mathcal{S}(f, \mathbf{x}, \cdot, \mathbf{S})$ *is linear mapping, i.e.*

$$\begin{aligned} \mathcal{S}(f, \mathbf{x}, \mathbf{b}+\mathbf{d}, \mathbf{S}) &= \mathcal{S}(f, \mathbf{x}, \mathbf{b}, \mathbf{S}) + \mathcal{S}(f, \mathbf{x}, \mathbf{d}, \mathbf{S}), \\ \mathcal{S}(f, \mathbf{x}, \lambda\mathbf{b}, \mathbf{S}) &= \lambda\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \mathbf{S}); \end{aligned} \tag{2.26}$$

d) $\mathcal{S}(f, \mathbf{x}, \cdot, \mathbf{S})$ *is increasing on* $[0,\infty)^m$, *i.e. if* $\mathbf{b}, \mathbf{c} \in [0,\infty)^m$ *such that* $\mathbf{b} \geq \mathbf{c}$, *i.e.* $b_j \geq c_j, j = 1, ..., m$, *then*

$$\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \mathbf{S}) \geq \mathcal{S}(f, \mathbf{x}, \mathbf{c}, \mathbf{S}) \geq 0. \tag{2.27}$$

Proof: *a) Positivity of functional* $\mathcal{S}$ *follows from Sherman's inequality (2.10).*

b) We have

$$\begin{aligned} &\mathcal{S}(f, \mathbf{x}, \mathbf{b}, \lambda\mathbf{S} + (1-\lambda)\mathbf{T}) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} b_j \left(\lambda s_{ij} + (1-\lambda)t_{ij}\right) f(x_i) - \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} \left(\lambda s_{ij} + (1-\lambda)t_{ij}\right) x_i\right) \\ &= \lambda \sum_{i=1}^{n}\sum_{j=1}^{m} b_j s_{ij} f(x_i) + (1-\lambda)\sum_{i=1}^{n}\sum_{j=1}^{m} b_j t_{ij} f(x_i) - \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} \left(\lambda s_{ij} + (1-\lambda)t_{ij}\right) x_i\right). \end{aligned} \tag{2.28}$$

Further, by convexity of f *we have*

$$\begin{aligned} &\sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} \left(\lambda s_{ij} + (1-\lambda)t_{ij}\right) x_i\right) \\ &\leq \sum_{j=1}^{m} b_j \left[\lambda f\left(\sum_{i=1}^{n} s_{ij}x_i\right) + (1-\lambda) f\left(\sum_{i=1}^{n} t_{ij}x_i\right)\right] \\ &= \lambda \sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} s_{ij}x_i\right) + (1-\lambda)\sum_{j=1}^{m} b_j f\left(\sum_{i=1}^{n} t_{ij}x_i\right). \end{aligned} \tag{2.29}$$

Combining (2.28) and (2.29), we prove the concavity property b).

c) It is obvious.

d) Case $\mathbf{b}=\mathbf{c}$ *is trivial. Let* $\mathbf{b}>\mathbf{c}$*. Since*

$$\mathbf{b}=(\mathbf{b}-\mathbf{c})+\mathbf{c},$$

using the property linearity (2.26) and positivity (2.24), we have

$$\begin{aligned}\mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S}) &= \mathcal{S}(f,\mathbf{x},(\mathbf{b}-\mathbf{c})+\mathbf{c},\mathbf{S})\\ &= \mathcal{S}(f,\mathbf{x},\mathbf{b}-\mathbf{c},\mathbf{S})+\mathcal{S}(f,\mathbf{x},\mathbf{c},\mathbf{S})\\ &\geq \mathcal{S}(f,\mathbf{x},\mathbf{c},\mathbf{S})\geq 0,\end{aligned}$$

as wanted. This ends the proof. □

2.3.2. Lower and upper bounds for Sherman's functional

In the following theorems we establish lower and upper bounds for Sherman's functional (2.23).

Theorem 2.3.8*([72])*: *For* $\mathbf{b}=(b_1,...,b_m)\in[0,\infty)^m$ *and* $\mathbf{S}=(s_{ij})\in\mathcal{C}_{nm}([0,\infty))$*, we define*

$$\tilde{m}=\sum_{j=1}^{m}b_j\min_{1\leqslant i\leqslant n}\{s_{ij}\}\quad \textit{and}\quad \tilde{M}=\sum_{j=1}^{m}b_j\max_{1\leqslant i\leqslant n}\{s_{ij}\}.$$

Then

$$0\leq\tilde{m}\cdot J(f,\mathbf{x},\mathbf{e})\leqslant\mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S})\leqslant\tilde{M}\cdot J(f,\mathbf{x},\mathbf{e}).\tag{2.30}$$

Proof: *The first inequality in (2.30) is consequence of nonnegativity of all entries of* $\mathbf{b}$*,* $\mathbf{S}$ *and Jensen's functional* $J(f,\mathbf{x},\mathbf{e})$ *given by (2.19).*

Now we prove

$$\min_{1\leqslant i\leqslant n}\{s_{ij}\}\cdot J(f,\mathbf{x},\mathbf{e})\leqslant\sum_{i=1}^{n}s_{ij}f(x_i)-f\left(\sum_{i=1}^{n}x_is_{ij}\right)\leqslant\max_{1\leqslant i\leqslant n}\{s_{ij}\}\cdot J(f,\mathbf{x},\mathbf{e}),\tag{2.31}$$

for all $j=1,...,m$.

Let us denote $\tilde{m}_j=\min_{1\leqslant i\leqslant n}\{s_{ij}\}$ *and* $\tilde{M}_j=\max_{1\leqslant i\leqslant n}\{s_{ij}\}$*,* $j\in\{1,...,m\}$*. Then*

(2.31) can be written as

$$\tilde{m}_j \cdot J(f,\mathbf{x},\mathbf{e}) \leqslant \sum_{i=1}^{n} s_{ij} f\left(x_i\right) - f\left(\sum_{i=1}^{n} x_i s_{ij}\right) \leqslant \tilde{M}_j \cdot J(f,\mathbf{x},\mathbf{e}). \tag{2.32}$$

Since $n \cdot \tilde{m}_j + \sum\limits_{i=1}^{n}\left(s_{ij} - \tilde{m}_j\right) = 1$, *applying Jensen's inequality we have*

$$\begin{aligned} f\left(\sum_{i=1}^{n} x_i s_{ij}\right) &= f\left(n \cdot \tilde{m}_j \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + \sum_{i=1}^{n}\left(s_{ij} - \tilde{m}_j\right) x_i\right) \\ &\leq n \cdot \tilde{m}_j f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + \sum_{i=1}^{n}\left(s_{ij} - \tilde{m}_j\right) f(x_i) \\ &= \sum_{i=1}^{n} s_{ij} f(x_i) - \tilde{m}_j \left(\sum_{i=1}^{n} f(x_i) - n f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right) \\ &= \sum_{i=1}^{n} s_{ij} f(x_i) - \tilde{m}_j \cdot J(f,\mathbf{x},\mathbf{e}) \end{aligned}$$

which gives the left hand side of (2.32).

Further, since $\tilde{M}_j > 0$ *and* $\frac{1}{\tilde{M}_j}\sum\limits_{i=1}^{n}\left(\frac{\tilde{M}_j}{n} - \frac{s_{ij}}{n}\right) + \frac{1}{n\tilde{M}_j} = 1$, *applying Jensen's inequality we have*

$$\begin{aligned} f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) &= f\left(\frac{1}{\tilde{M}_j}\sum_{i=1}^{n}\left(\frac{\tilde{M}_j}{n} - \frac{s_{ij}}{n}\right) x_i + \frac{1}{n\tilde{M}_j}\left(\sum_{i=1}^{n} s_{ij} x_i\right)\right) \\ &\leq \frac{1}{\tilde{M}_j}\sum_{i=1}^{n}\left(\frac{\tilde{M}_j}{n} - \frac{s_{ij}}{n}\right) f(x_i) + \frac{1}{n\tilde{M}_j} f\left(\sum_{i=1}^{n} s_{ij} x_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} f(x_i) - \frac{1}{n\tilde{M}_j}\left(\sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} s_{ij} x_i\right)\right), \end{aligned}$$

which gives the right hand side of (2.32).

Therefore, the inequality (2.31) holds. Now, multiplying with b_j *and summing over* $j = 1, ..., m$, *we get (2.30) what we need to prove.* □

Theorem 2.3.9*([72])*: *For* $\mathbf{b} = (b_1, ..., b_m) \in [0,\infty)^m$ *and* $\mathbf{S} = (s_{ij}) \in \mathcal{C}_{nm}([0,\infty))$ *and*

$\mathbf{T} = (t_{ij}) \in \mathcal{C}_{nm}((0,\infty))$, *we define*

$$\bar{\mathbf{b}}_{\min} = \left(b_1 \min_{1\le i\le n} \left\{ \frac{s_{i1}}{t_{i1}} \right\}, ..., b_m \min_{1\le i\le n} \left\{ \frac{s_{im}}{t_{im}} \right\} \right),$$
$$\bar{\mathbf{b}}_{\max} = \left(b_1 \max_{1\le i\le n} \left\{ \frac{s_{i1}}{t_{i1}} \right\}, ..., b_m \max_{1\le i\le n} \left\{ \frac{s_{im}}{t_{im}} \right\} \right).$$

Then

$$(0 \le)\ \ \mathcal{S}(f,\mathbf{x},\bar{\mathbf{b}}_{\min},\mathbf{T}) \le \mathcal{S}(f,\mathbf{x},\mathbf{b},\mathbf{S}) \le \mathcal{S}(f,\mathbf{x},\bar{\mathbf{b}}_{\max},\mathbf{T}). \tag{2.33}$$

Proof: *According to the notation, we need to prove*

$$\begin{aligned}
&\sum_{j=1}^{m} b_j \min_{1\le i\le n} \left\{ \frac{s_{ij}}{t_{ij}} \right\} \left(\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} x_i t_{ij} \right) \right) \\
&\le \sum_{j=1}^{m} b_j \left(\sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} x_i s_{ij} \right) \right) \\
&\le \sum_{j=1}^{m} b_j \max_{1\le i\le n} \left\{ \frac{s_{ij}}{t_{ij}} \right\} \left(\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} x_i t_{ij} \right) \right)
\end{aligned}$$

Let us consider n-tuples $\boldsymbol{p}$ and $\mathbf{q}_{\min}$ such that

$$\begin{aligned}
\boldsymbol{p} &= \left(t_{1j}\frac{s_{1j}}{t_{1j}}, \ldots, t_{nj}\frac{s_{nj}}{t_{nj}} \right) = (s_{1j}, ..., s_{nj}), \\
\boldsymbol{q}_{\min} &= \left(t_{1j} \min_{1\le i\le n} \left\{ \frac{s_{ij}}{t_{ij}} \right\}, \ldots, t_{nj} \min_{1\le i\le n} \left\{ \frac{s_{ij}}{t_{ij}} \right\} \right).
\end{aligned}$$

It is obviously $\boldsymbol{q}_{\min} \le \mathbf{p}$. *Applying Theorem 2.3.6 to n-tuples* $\boldsymbol{q}_{\min}$ *and* $\mathbf{p}$, *we get*

$$\begin{aligned}
&\sum_{i=1}^{n} s_{ij} f(x_i) - \sum_{i=1}^{n} s_{ij} f\left(\frac{1}{\sum_{i=1}^{n} s_{ij}} \sum_{i=1}^{n} s_{ij} x_i\right) \\
&= \sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} s_{ij} x_i\right) \\
&\ge \sum_{i=1}^{n} t_{ij} \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\} f(x_i) \\
&\quad - \sum_{i=1}^{n} t_{ij} \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\} f\left(\frac{1}{\sum_{i=1}^{n} t_{ij} \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}} \sum_{i=1}^{n} t_{ij} \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\} x_i\right) \\
&= \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}\left[\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} t_{ij} x_i\right)\right],
\end{aligned}$$

i.e. we get

$$\sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} s_{ij} x_i\right) \ge \min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}\left[\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} t_{ij} x_i\right)\right].$$

Exchanging the roles of min and max, we get the inequality

$$\sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} s_{ij} x_i\right) \le \max_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}\left[\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} t_{ij} x_i\right)\right].$$

So, we have

$$\begin{aligned}
&\min_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}\left[\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} t_{ij} x_i\right)\right] \\
&\le \sum_{i=1}^{n} s_{ij} f(x_i) - f\left(\sum_{i=1}^{n} s_{ij} x_i\right) \\
&\le \max_{1 \le i \le n}\left\{\frac{s_{ij}}{t_{ij}}\right\}\left[\sum_{i=1}^{n} t_{ij} f(x_i) - f\left(\sum_{i=1}^{n} t_{ij} x_i\right)\right].
\end{aligned}$$

Now multiplying with b_j *and summing over* $j = 1, ..., m$, *we get the required result.* □

Remark 2.3.6.: Choosing $m = 1$ and setting $b_1 = 1$ and $s_{i1} = \frac{p_i}{P_n}$, for $i = 1, ..., n$, Sherman's functional (2.23) reduces to Jensen's functional that corresponds to (2.18). Therefore, applying Theorem 2.3.8, we obtain the bounds of Jensen's functional in the form (2.18), i.e.

$$0 \leq \min_{1\leq i\leq n}\{p_i\}J(f,\mathbf{x},\mathbf{e}) \leq J(f,\mathbf{x},\mathbf{p}) \leq \max_{1\leq i\leq n}\{p_i\}J(f,\mathbf{x},\mathbf{e}).$$

Applying Theorem 2.3.7, the concavity property b), for $\lambda = \frac{1}{2}$, with $m = 1$, $b_1 = 1$, $s_{i1} = p_i$, $t_{i1} = q_i$, for $i = 1, ..., n$, $P_n = Q_n = 1$, it implies for Jensen's functional (2.20) the properties superadditivity

$$\bar{J}(f,\mathbf{x},\mathbf{p}+\mathbf{q}) \geq \bar{J}(f,\mathbf{x},\mathbf{p}) + \bar{J}(f,\mathbf{x},\mathbf{q}), \quad \mathbf{p},\mathbf{q} \in \mathcal{P}_n^0, \tag{2.34}$$

and monotonicity

$$\bar{J}(f,\mathbf{x},\mathbf{p}) \geq \bar{J}(f,\mathbf{x},\mathbf{q}) \geq 0, \quad , \quad \mathbf{p},\mathbf{q} \in \mathcal{P}_n^0,\ \mathbf{p} \geq \mathbf{q}. \tag{2.35}$$

Further, applying Theorem 2.3.9, with $m = 1$, $b_1 = 1$ and $s_{i1} = p_i$, $t_{i1} = q_i$, for $i = 1, ..., n$, the inequality (2.33) reduces to the inequality (2.21).

2.4. Schur-concavity of some sums generated by the Jensen and Jensen-Mercer functionals

In this section, we prove Sherman type theorem for $\leq_{\mathcal{C}}$-convex functions, where $\leq_{\mathcal{C}}$ is a cone preordering on a linear space induced by a convex cone $\mathcal{C}$. As a specification, we obtain a majorization theorem for such functions. We apply the obtained results to the Jensen and Jensen-Mercer type functionals and investigate their properties. Thus we get some extensions of the property of the superadditivity and monotonicity of these functionals. In order to develop these properties, of the functionals, we introduce a vectorial majorization ordering for comparing two tuples of vectors of a real linear space. Next, we prove a Sherman type inequality for a vector-valued $\leq_C$-convex function f. In consequence, we get a Hardy-Littlewood-Pólya-Karamata like inequality. Finally, we prove that some sums related to the Jensen and Jensen-Mercer functionals are Schur-concave with respect to their weight vectors. Results of this section are cited mainly from [133].

2.4.1. Motivation

We intend to show an extended approach to Sherman type theorems via vectorial majorization. This is motivated by the natural problem of studying some sums of Jensen functionals, which requires to use n-tuples of vectorial parameters in place of scalar ones.

Remark 2.4.7.: Equation (2.18) can be rewritten as

$$J(f,\mathbf{x},\mathbf{p}) = \langle \mathbf{p}, f(\mathbf{x})\rangle - \langle \mathbf{p},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p},\mathbf{x}\rangle}{\langle \mathbf{p},\mathbf{e}\rangle}\right), \tag{2.36}$$

where $f(\mathbf{x}) = (f(x_1), f(x_2), \ldots, f(x_k))$ with $\mathbf{e} = (1,1,\ldots,1) \in \mathbb{R}^k$, and $\langle\cdot,\cdot\rangle$ is the standard inner product on $\mathbb{R}^k$. The celebrated Jensen's inequality says that if $f : I \to \mathbb{R}$ is a convex function, then

$$J(f,\mathbf{x},\mathbf{p}) \geq 0.$$

As in the previous section, the set of the described nonnegative real k-tuples $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, with $P_k = \sum_{i=1}^{k} p_i > 0$, will be denoted by $\mathcal{P}_k^0$.

Theorem 2.4.10*([42])*: *If $f : I \to \mathbb{R}$ is a convex function, then the function $\mathbf{p} \mapsto J(f,\mathbf{x},\mathbf{p})$, $\mathbf{p} \in \mathcal{P}_k^0$, is superadditive for any $\mathbf{x} \in I^k$, i.e.*

$$J(f,\mathbf{x},\mathbf{p}+\mathbf{q}) \geq J(f,\mathbf{x},\mathbf{p}) + J(f,\mathbf{x},\mathbf{q}), \quad \textit{for } \mathbf{p},\mathbf{q} \in \mathcal{P}_k^0. \tag{2.37}$$

In consequence, the function $\mathbf{p} \mapsto J(f,\mathbf{x},\mathbf{p})$, $\mathbf{p} \in \mathcal{P}_k^0$, is monotone for any $\mathbf{x} \in I^k$, i.e.

$$J(f,\mathbf{x},\mathbf{p}+\mathbf{q}) \geq J(f,\mathbf{x},\mathbf{p}), \quad \textit{for } \mathbf{p},\mathbf{q} \in \mathcal{P}_k^0.$$

Remark 2.4.8.: It can be shown that property (2.37) is equivalent to the subadditivity of the function

$$\mathbf{p} \mapsto \langle \mathbf{p},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p},\mathbf{x}\rangle}{\langle \mathbf{p},\mathbf{e}\rangle}\right), \quad \text{for } \mathbf{p} \in \mathcal{P}_k^0,$$

which means

$$\langle \mathbf{p}+\mathbf{q},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}+\mathbf{q},\mathbf{x}\rangle}{\langle \mathbf{p}+\mathbf{q},\mathbf{e}\rangle}\right) \leq \langle \mathbf{p},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p},\mathbf{x}\rangle}{\langle \mathbf{p},\mathbf{e}\rangle}\right) + \langle \mathbf{q},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{q},\mathbf{x}\rangle}{\langle \mathbf{q},\mathbf{e}\rangle}\right), \quad \text{for } \mathbf{p},\mathbf{q} \in \mathcal{P}_k^0.$$

We now discuss the Jensen-Mercer functional.

Definition 2.4.5.*([95])*: Let $f : I \to \mathbb{R}$ be a function on an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{x} = (x_1, x_2, \dots, x_k) \in [a, b]^k$ and $\mathbf{p} = (p_1, p_2, \dots, p_k) \in \mathcal{P}_k^0$. **The Jensen-Mercer functional** is defined by

$$M(f, \mathbf{x}, \mathbf{p}) = P_k(f(a) + f(b)) - \sum_{i=1}^{k} p_i f(x_i) - P_k f\left(a + b - \frac{1}{P_k}\sum_{i=1}^{k} p_i x_i\right).$$

It is not hard to verify that

$$M(f, \mathbf{x}, \mathbf{p}) = (f(a) + f(b))\langle \mathbf{p}, \mathbf{e}\rangle - \langle \mathbf{p}, f(\mathbf{x})\rangle - \langle \mathbf{p}, \mathbf{e}\rangle f\left(\frac{(a+b)\langle \mathbf{p}, \mathbf{e}\rangle - \langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle}\right),$$

where $f(\mathbf{x}) = (f(x_1), f(x_2), \dots, f(x_k))$ and $\mathbf{e} = (1, 1, \dots, 1) \in \mathbb{R}^k$.

Mercer's inequality [107], also known as Jensen-Mercer inequality (2.97), states that if $f : [a, b] \to \mathbb{R}$ is a convex function, then

$$M(f, \mathbf{x}, \mathbf{p}) \geq 0.$$

Theorem 2.4.11*([95])*: *If $f : I \to \mathbb{R}$ is a convex function, where $I = [a, b]$, then the function $\mathbf{p} \mapsto M(f, \mathbf{x}, \mathbf{p})$, $\mathbf{p} \in \mathcal{P}_k^0$, is superadditive for any $\mathbf{x} \in I^k$, i.e.*

$$M(f, \mathbf{x}, \mathbf{p} + \mathbf{q}) \geq M(f, \mathbf{x}, \mathbf{p}) + M(f, \mathbf{x}, \mathbf{q}), \quad \textit{for } \mathbf{p}, \mathbf{q} \in \mathcal{P}_k^0. \tag{2.38}$$

In consequence, the function $\mathbf{p} \mapsto M(f, \mathbf{x}, \mathbf{p})$, $\mathbf{p} \in \mathcal{P}_k^0$, is monotone for any $\mathbf{x} \in I^k$, i.e.

$$M(f, \mathbf{x}, \mathbf{p} + \mathbf{q}) \geq M(f, \mathbf{x}, \mathbf{p}), \quad \textit{for } \mathbf{p}, \mathbf{q} \in \mathcal{P}_k^0.$$

Remark 2.4.9.: The inequality (2.38) is equivalent to the subadditivity of the function

$$\mathbf{p} \mapsto \langle \mathbf{p}, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}, \mathbf{y}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle}\right), \quad \text{for } \mathbf{p} \in \mathcal{P}_k^0,$$

where $\mathbf{y} = (a+b)\mathbf{e} - \mathbf{x}$, i.e.

$$\langle \mathbf{p}+\mathbf{q},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}+\mathbf{q},\mathbf{y}\rangle}{\langle \mathbf{p}+\mathbf{q},\mathbf{e}\rangle}\right) \leq \langle \mathbf{p},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{p},\mathbf{y}\rangle}{\langle \mathbf{p},\mathbf{e}\rangle}\right) + \langle \mathbf{q},\mathbf{e}\rangle f\left(\frac{\langle \mathbf{q},\mathbf{y}\rangle}{\langle \mathbf{q},\mathbf{e}\rangle}\right), \quad \text{for } \mathbf{p},\mathbf{q} \in \mathcal{P}_k^0.$$

In the sequel, for a given $\mathbf{c} \in \mathcal{P}_k^0$, consider two sum decompositions

$$\mathbf{c} = \mathbf{p}_1 + \mathbf{p}_2 + \ldots + \mathbf{p}_m \quad \text{and} \quad \mathbf{c} = \mathbf{q}_1 + \mathbf{q}_2 + \ldots + \mathbf{q}_n,$$

where $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m \in \mathcal{P}_k^0$ and $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n \in \mathcal{P}_k^0$.

Our goal is to give a condition on the vectorial tuples $(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m)$ and $(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n)$ for the following inequalities to hold:

$$\begin{aligned} & J(f,\mathbf{x},\mathbf{p}_1) + J(f,\mathbf{x},\mathbf{p}_2) + \ldots + J(f,\mathbf{x},\mathbf{p}_m) \\ \leq\ & J(f,\mathbf{x},\mathbf{q}_1) + J(f,\mathbf{x},\mathbf{q}_2) + \ldots + J(f,\mathbf{x},\mathbf{q}_n), \end{aligned} \tag{2.39}$$

and

$$\begin{aligned} & M(f,\mathbf{x},\mathbf{p}_1) + M(f,\mathbf{x},\mathbf{p}_2) + \ldots + M(f,\mathbf{x},\mathbf{p}_m) \\ \leq\ & M(f,\mathbf{x},\mathbf{q}_1) + M(f,\mathbf{x},\mathbf{q}_2) + \ldots + M(f,\mathbf{x},\mathbf{q}_n). \end{aligned} \tag{2.40}$$

So, for the Jensen and Jensen-Mercer functionals, we are interested in inequalities of the Hardy-Littlewood-Pólya-Karamata and Sherman types.

To do so, in Subsection 2.4.2. we introduce a *vectorial majorization* relation for comparing tuples of vectors. This develops the classical approach to majorization theory.

In Subsections 2.4.3. and 2.4.4. we provide majorization generalizations of Theorems 2.4.10 and 2.4.11 for the Jensen and Jensen-Mercer functionals, respectively.

2.4.2. Vectorial majorization and Sherman type inequalities

Throughout this subsection $\mathcal{V}$ and $\mathcal{W}$ are real linear spaces.

Definition 2.4.6.: An n-tuple $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in \mathcal{V}^n$ is said to be **pre-majorized**

by an m-tuple $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) \in \mathcal{V}^m$, written as $\mathbf{Y} \prec_p \mathbf{X}$, if there exists an $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$ such that

$$(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)\mathbf{S} \tag{2.41}$$

in the standard sense

$$\mathbf{y}_j = s_{1j}\mathbf{x}_1 + s_{2j}\mathbf{x}_2 + \ldots + s_{mj}\mathbf{x}_m \quad \text{for } j = 1, 2, \ldots, n. \tag{2.42}$$

(see [35] for a similar notion of matrix majorization).

Definition 2.4.7.: An n-tuple $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in \mathcal{V}^n$ is said to be **majorized** by an n-tuple $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in \mathcal{V}^n$, written as $\mathbf{Y} \prec \mathbf{X}$, if there exists an $n \times n$ doubly stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_n(\mathbb{R})$ such that

$$(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\mathbf{S}. \tag{2.43}$$

Remark 2.4.10.: It is obvious that if $\mathcal{V} = \mathbb{R}$ then $\prec$ reduces to the standard majorization ordering for n-tuples in $\mathbb{R}^n$ (see [106, pp. 8 and 33]).

Definition 2.4.8.: A nonempty subset $\mathcal{C}$ of $\mathcal{W}$ is said to be a **convex cone** if:

(i) $\mathbf{a}, \mathbf{b} \in \mathcal{C}$ implies $\mathbf{a} + \mathbf{b} \in \mathcal{C}$,

(ii) $\mathbf{a} \in \mathcal{C}$ and $0 \leq t \in \mathbb{R}$ imply $t\mathbf{a} \in \mathcal{C}$.

Definition 2.4.9.: If $\mathcal{C} \subset \mathcal{W}$ is a convex cone, then the relation $\leq_{\mathcal{C}}$ on $\mathcal{W}$ defined by

$$\mathbf{b} \leq_{\mathcal{C}} \mathbf{a} \quad \text{iff} \quad \mathbf{a} - \mathbf{b} \in \mathcal{C}, \quad \text{for} \quad \mathbf{a}, \mathbf{b} \in \mathcal{W},$$

is called **cone preordering** (induced by $\mathcal{C}$).

Definition 2.4.10.: For a given convex cone $\mathcal{C}$ in a linear space $\mathcal{W}$, a function $f : \mathcal{I} \to \mathcal{W}$ is called **$\leq_{\mathcal{C}}$-convex** on a convex set $\mathcal{I} \subset \mathcal{V}$, if

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq_{\mathcal{C}} \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y})$$

for $\mathbf{x}, \mathbf{y} \in \mathcal{I}$ and $\alpha \in [0,1]$.

A function $f : \mathcal{I} \to \mathcal{W}$ is called $\leq_{\mathcal{C}}$**-concave** if $-f$ is $\leq_{\mathcal{C}}$-convex.

In Theorem 2.4.12 we present a Sherman type inequality (2.45) for tuples of vectors (cf. [149], see also [25, 31, 126]). This gives an extension of Theorem 2.2.4 to the vector case.

In the sequel we denote $\mathbb{R}_+ = [0, \infty)$.

Theorem 2.4.12 *([133])*: *Let $f : \mathcal{I} \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function defined on a convex set $\mathcal{I} \subset \mathcal{V}$. Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) \in \mathcal{I}^m$, $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in \mathcal{I}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$.*

If

$$\mathbf{Y} = \mathbf{X}\mathbf{S} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{2.44}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$, then

$$\sum_{j=1}^{n} b_j f(\mathbf{y}_j) \leq_{\mathcal{C}} \sum_{i=1}^{m} a_i f(\mathbf{x}_i). \tag{2.45}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.45) is reversed.

Proof: (Cf. [126].) By (2.44) we can see that

$$(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)\mathbf{S}.$$

Due to (2.42) we have $\mathbf{y}_j = \sum\limits_{i=1}^{m} s_{ij}\mathbf{x}_i$ with $\sum\limits_{i=1}^{m} s_{ij} = 1$, $j = 1, 2, \ldots, n$ and $s_{ij} \geq 0$. By the $\leq_{\mathcal{C}}$-convexity of f, we derive

$$f(\mathbf{y}_j) = f\left(\sum_{i=1}^{m} s_{ij}\mathbf{x}_i\right) \leq_{\mathcal{C}} \sum_{i=1}^{m} s_{ij} f(\mathbf{x}_i), \quad \text{for } j = 1, 2, \ldots, n.$$

Consequently,

$$\sum_{j=1}^{n} b_j f(\mathbf{y}_j) \leq_{\mathcal{C}} \sum_{j=1}^{n} b_j \sum_{i=1}^{m} s_{ij} f(\mathbf{x}_i).$$

This gives

$$\sum_{j=1}^{n} b_j f(\mathbf{y}_j) \leq_{\mathcal{C}} \sum_{i=1}^{m} \sum_{j=1}^{n} b_j s_{ij} f(\mathbf{x}_i).$$

Again, by (2.44), $\mathbf{a} = \mathbf{b}\mathbf{S}^T$, which forces $a_i = \sum_{j=1}^{n} b_j s_{ij}$, $i = 1, 2, \ldots, m$.

In summary, we deduce that

$$\sum_{j=1}^{n} b_j f(\mathbf{y}_j) \leq_{\mathcal{C}} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} b_j s_{ij} \right) f(\mathbf{x}_i) = \sum_{i=1}^{m} a_i f(\mathbf{x}_i).$$

Thus the inequality (2.45) is proven. □

It is interesting that the statement (2.44) says that the pairs $(\mathbf{X}, \mathbf{a})$ and $(\mathbf{Y}, \mathbf{b})$ are *weighted majorized* (see [31]).

A corollary to Theorem 2.4.12 is the following majorization theorem for n-tuples of vectors.

Theorem 2.4.13 *([133])*: *Let $f : \mathcal{I} \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function defined on a convex set $\mathcal{I} \subset \mathcal{V}$. Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in \mathcal{I}^n$ and $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in \mathcal{I}^n$.*

If $\mathbf{Y} \prec \mathbf{X}$, then

$$\sum_{i=1}^{n} f(\mathbf{y}_i) \leq_{\mathcal{C}} \sum_{i=1}^{n} f(\mathbf{x}_i). \tag{2.46}$$

Proof: We apply Theorem 2.4.12 with $m = n$ and $\mathbf{b} = \mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^n$. The vectorial majorization $\mathbf{Y} \prec \mathbf{X}$ means that $\mathbf{Y} = \mathbf{X}\mathbf{S}$ for some doubly stochastic matrix $\mathbf{S}$ (cf. (2.43)). By putting $\mathbf{a} = \mathbf{b}\mathbf{S}^T$ as in (2.44), we get $\mathbf{a} = \mathbf{e}\mathbf{S}^T = \mathbf{e}$. Now, it follows from (2.45) that the required inequality (2.46) is satisfied. □

Remark 2.4.11.: Theorem 2.4.13 reduces to the classical *Majorization Theorem* (Theorem 2.1.1), whenever $\mathcal{V} = \mathcal{W} = \mathbb{R}$ and $\mathcal{C} = [0, \infty)$.

Definition 2.4.11.: A function $F : \mathcal{I}^n \to \mathcal{W}$, with a convex set $\mathcal{I} \subset \mathcal{V}$, is called **$\leq_{\mathcal{C}}$-Schur-convex** on $\mathcal{I}^n$ if

$$F(\mathbf{Y}) \leq_{\mathcal{C}} F(\mathbf{X})$$

for every $\mathbf{X}, \mathbf{Y} \in \mathcal{I}^n$ such that

$$\mathbf{Y} \prec \mathbf{X}.$$

A function F is called $\leq_{\mathcal{C}}$**-Schur-concave** on $\mathcal{I}^n$, if $-F$ is $\leq_{\mathcal{C}}$-Schur-convex on $\mathcal{I}^n$.

According to Theorem 2.4.13 one sees that if f is $\leq_{\mathcal{C}}$-convex on $\mathcal{I}$, then the function

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i=1}^{n} f(\mathbf{x}_i) \quad \text{for } (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in \mathcal{I}^n$$

is $\leq_{\mathcal{C}}$-Schur-convex on $\mathcal{I}^n$.

2.4.3. Applications to the Jensen functional

In this subsection $\mathcal{V} = \mathbb{R}^k$, $\mathcal{I} = \mathcal{P}_k^0$ and $\mathcal{W}$ is a linear space with a convex cone $\mathcal{C} \subset \mathcal{W}$ inducing preordering $\leq_{\mathcal{C}}$ on $\mathcal{W}$.

Lemma 2.4.3. *([133]): Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_l) \in \mathcal{I}^l$, where $\mathcal{I} = \mathcal{P}_k^0$.*

Then for any fixed $\mathbf{x} \in I^k$,

$$\left\langle \sum_{i=1}^{l} \mathbf{p}_i, \mathbf{e} \right\rangle f \left(\frac{\left\langle \sum_{i=1}^{l} \mathbf{p}_i, \mathbf{x} \right\rangle}{\left\langle \sum_{i=1}^{l} \mathbf{p}_i, \mathbf{e} \right\rangle} \right) \leq_{\mathcal{C}} \sum_{i=1}^{l} \langle \mathbf{p}_i, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right). \tag{2.47}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.47) is reversed.

Proof: It is not hard to check that

$$\frac{\left\langle \sum_{i=1}^{l} \mathbf{p}_i, \mathbf{x} \right\rangle}{\left\langle \sum_{i=1}^{l} \mathbf{p}_i, \mathbf{e} \right\rangle} = \sum_{i=1}^{l} \alpha_i \frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}, \tag{2.48}$$

where

$$\alpha_i = \frac{\langle \mathbf{p}_i, \mathbf{e} \rangle}{\sum_{j=1}^{l} \langle \mathbf{p}_j, \mathbf{e} \rangle} \quad \text{for } i = 1, 2, \ldots, l. \tag{2.49}$$

Obviously, $\sum\limits_{i=1}^{l} \alpha_i = 1$ with $\alpha_i \geq 0$ for $i = 1, 2, \ldots, l$.

By using (2.48)-(2.49) and the $\leq_{\mathcal{C}}$-convexity on I of the function $f : I \to \mathcal{W}$, we find that

$$f\left(\frac{\left\langle \sum\limits_{i=1}^{l} \mathbf{p}_i, \mathbf{x} \right\rangle}{\left\langle \sum\limits_{i=1}^{l} \mathbf{p}_i, \mathbf{e} \right\rangle}\right) = f\left(\sum_{i=1}^{l} \alpha_i \frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right) \tag{2.50}$$

$$\leq_{\mathcal{C}} \sum_{i=1}^{l} \alpha_i f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right)$$

$$= \sum_{i=1}^{l} \frac{\langle \mathbf{p}_i, \mathbf{e} \rangle}{\sum\limits_{j=1}^{l} \langle \mathbf{p}_j, \mathbf{e} \rangle} f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right) = \frac{1}{\sum\limits_{j=1}^{l} \langle \mathbf{p}_j, \mathbf{e} \rangle} \sum_{i=1}^{l} \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right).$$

It is now seen that (2.50) implies (2.47). □

The forthcoming result is a Sherman like inequality (2.52) for the function $\mathbf{p} \mapsto \langle \mathbf{p}, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}, \mathbf{x} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle}\right)$, $\mathbf{p} \in \mathcal{P}_k^0$.

Theorem 2.4.14 *([133])*: *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m) \in \mathcal{I}^m$, $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If

$$\mathbf{Q} = \mathbf{P}\mathbf{S} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{2.51}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$, then for any fixed $\mathbf{x} \in I^k$,

$$\sum_{j=1}^{n} b_j \langle \mathbf{q}_j, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{q}_j, \mathbf{x} \rangle}{\langle \mathbf{q}_j, \mathbf{e} \rangle}\right) \leq_{\mathcal{C}} \sum_{i=1}^{m} a_i \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right). \tag{2.52}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.52) is reversed.

Proof: We prove that $\frac{\langle \mathbf{p}, \mathbf{x} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle} \in I$ for any fixed $\mathbf{x} \in I^k$. In fact, $\mathbf{x} \in I^k$ implies $x_i \in I$, i.e., $a \leq x_i \leq b$, for $i = 1, 2, \ldots, k$. Equivalently, $a\mathbf{e} \leq \mathbf{x} \leq b\mathbf{e}$, where $\leq$ denotes the componentwise ordering on $\mathbb{R}^k$. Thus we get

$$\frac{\langle \mathbf{p}, a\mathbf{e}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle} \leq \frac{\langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle} \leq \frac{\langle \mathbf{p}, b\mathbf{e}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle},$$

and next

$$a \leq \frac{\langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle} \leq b.$$

Hence $\frac{\langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle} \in I$, as required.

Therefore the function

$$\mathbf{p} \to \langle \mathbf{p}, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle}\right), \quad \mathbf{p} \in \mathcal{P}_k^0, \tag{2.53}$$

is well-defined. So, both the sides of the inequality (2.52) are well-defined.

We shall prove the $\leq_{\mathcal{C}}$-convexity of the function (2.53) on the convex set $\mathcal{I} = \mathcal{P}_k^0$. To show this, we apply Lemma 2.4.3. for $l = 2$. Namely, for any $\mathbf{p}, \mathbf{q} \in \mathcal{P}_k^0$ and $\alpha \in (0,1)$, we have $\alpha\mathbf{p}, (1-\alpha)\mathbf{q} \in \mathcal{P}_k^0$. By using (2.47) for the vectors $\alpha\mathbf{p}$ and $(1-\alpha)\mathbf{q}$ in place of $\mathbf{p}_1$ and $\mathbf{p}_2$, we get

$$\begin{aligned}
&\langle \alpha\mathbf{p} + (1-\alpha)\mathbf{q}, \mathbf{e}\rangle f\left(\frac{\langle \alpha\mathbf{p} + (1-\alpha)\mathbf{q}, \mathbf{x}\rangle}{\langle \alpha\mathbf{p} + (1-\alpha)\mathbf{q}, \mathbf{e}\rangle}\right)\\
&\leq_{\mathcal{C}} \langle \alpha\mathbf{p}, \mathbf{e}\rangle f\left(\frac{\langle \alpha\mathbf{p}, \mathbf{x}\rangle}{\langle \alpha\mathbf{p}, \mathbf{e}\rangle}\right) + \langle (1-\alpha)\mathbf{q}, \mathbf{e}\rangle f\left(\frac{\langle (1-\alpha)\mathbf{q}, \mathbf{x}\rangle}{\langle (1-\alpha)\mathbf{q}, \mathbf{e}\rangle}\right)\\
&= \alpha\langle \mathbf{p}, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}, \mathbf{x}\rangle}{\langle \mathbf{p}, \mathbf{e}\rangle}\right) + (1-\alpha)\langle \mathbf{q}, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{q}, \mathbf{x}\rangle}{\langle \mathbf{q}, \mathbf{e}\rangle}\right).
\end{aligned} \tag{2.54}$$

Furthermore, for $\alpha = 0$ or $\alpha = 1$ inequality (2.54) holds trivially.

Now, in light of Theorem 2.4.12 applied to the $\leq_{\mathcal{C}}$-convex function (2.53), the inequality (2.52) is valid. □

As a specialization of Theorem 2.4.14, we obtain a Hardy-Littlewood-Pólya-Karamata type result.

Theorem 2.4.15 ***([133])*:** *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) \in \mathcal{I}^n$ and $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If $\mathbf{Q} \prec \mathbf{P}$*, then for any fixed* $\mathbf{x} \in I^k$*,*

$$\sum_{i=1}^{n} \langle \mathbf{q}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{q}_i, \mathbf{x} \rangle}{\langle \mathbf{q}_i, \mathbf{e} \rangle}\right) \leq_{\mathcal{C}} \sum_{i=1}^{n} \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right). \tag{2.55}$$

If f *is* $\leq_{\mathcal{C}}$*-concave, then the inequality (2.55) is reversed.*

Proof: Let $m = n$ and $\mathbf{b} = \mathbf{e} = (1, \dots, 1) \in \mathbb{R}^n$. Since $(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n) \prec (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n)$, we deduce that $(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n) = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n)\mathbf{S}$ for some $n \times n$ doubly stochastic matrix $\mathbf{S}$.

For $\mathbf{a} = \mathbf{b}\mathbf{S}^T$, we have $\mathbf{a} = \mathbf{e}\mathbf{S}^T = \mathbf{e}$.

Now, by employing Theorem 2.4.14, eq. (2.52), we derive the wanted inequality (2.55). □

The statement (2.55) states that for any fixed $\mathbf{x} \in I^k$, the function

$$F(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n) = \sum_{i=1}^{n} \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle}\right), \quad \mathbf{p}_i \in \mathcal{I},$$

is $\leq_{\mathcal{C}}$-Schur-convex on $\mathcal{I}^n$.

The right-hand side of inequality (2.55) can be thought of as the f-divergence of $\langle \mathbf{P}, \mathbf{x} \rangle = (\langle \mathbf{p}_1, \mathbf{x} \rangle, \dots, \langle \mathbf{p}_n, \mathbf{x} \rangle)$ and $\langle \mathbf{P}, \mathbf{e} \rangle = (\langle \mathbf{p}_1, \mathbf{e} \rangle, \dots, \langle \mathbf{p}_n, \mathbf{e} \rangle)$ [38]. A similar observation holds true for the left-hand side of inequality (2.55).

We now extend the notion of Jensen functional for $\mathcal{W}$-valued functions.

Definition 2.4.12.: Let $f : I \to \mathcal{W}$ be a $\mathcal{W}$-valued function. Let $\mathbf{p} = (p_1, p_2, \dots, p_k) \in \mathbb{R}_+^k$ and $\mathbf{x} = (x_1, x_2, \dots, x_k) \in I^k$. **The Jensen functional** is defined as follows:

$$J(f, \mathbf{x}, \mathbf{p}) = [\mathbf{p}, f(\mathbf{x})] - \langle \mathbf{p}, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}, \mathbf{x} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle}\right), \tag{2.56}$$

where $f(\mathbf{x}) = (f(x_1), f(x_2), \dots, f(x_k))$, $\mathbf{e} = (1, 1, \dots, 1) \in \mathbb{R}^k$, and $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^k$, and $[\mathbf{p}, f(\mathbf{x})] = \sum\limits_{i=1}^{k} p_i f(x_i) \in \mathcal{W}$.

We now illustrate the previous results in terms of the Jensen functional.

Corollary 2.4.1. *([133]):* *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m) \in \mathcal{I}^m$, $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If

$$\mathbf{Q} = \mathbf{PS} \quad \textit{and} \quad \mathbf{a} = \mathbf{bS}^T \tag{2.57}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$, then for any fixed $\mathbf{x} \in I^k$,

$$\sum_{j=1}^{n} b_j J(f, \mathbf{x}, \mathbf{q}_j) \geq_{\mathcal{C}} \sum_{i=1}^{m} a_i J(f, \mathbf{x}, \mathbf{p}_i). \tag{2.58}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.58) is reversed.

Proof: It follows from (2.56) that

$$\begin{aligned} \sum_{j=1}^{n} b_j J(f, \mathbf{x}, \mathbf{q}_j) &= \sum_{j=1}^{n} b_j \left([\mathbf{q}_j, f(\mathbf{x})] - \langle \mathbf{q}_j, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{q}_j, \mathbf{x} \rangle}{\langle \mathbf{q}_j, \mathbf{e} \rangle} \right) \right) \\ &= \left[\sum_{j=1}^{n} b_j \mathbf{q}_j, f(\mathbf{x}) \right] - \sum_{j=1}^{n} b_j \langle \mathbf{q}_j, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{q}_j, \mathbf{x} \rangle}{\langle \mathbf{q}_j, \mathbf{e} \rangle} \right). \end{aligned} \tag{2.59}$$

Also, we have

$$\begin{aligned} \sum_{i=1}^{m} a_i J(f, \mathbf{x}, \mathbf{p}_i) &= \sum_{i=1}^{m} a_i \left([\mathbf{p}_i, f(\mathbf{x})] - \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right) \right) \\ &= \left[\sum_{i=1}^{m} a_i \mathbf{p}_i, f(\mathbf{x}) \right] - \sum_{i=1}^{m} a_i \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right). \end{aligned} \tag{2.60}$$

By (2.57) it can be shown that

$$\sum_{j=1}^{n} b_j \mathbf{q}_j = \sum_{i=1}^{m} a_i \mathbf{p}_i.$$

Finally, by (2.59) and (2.60) and thanks to Theorem 2.4.14, we obtain inequality (2.58), completing the proof. □

Corollary 2.4.2. *([133]):* *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) \in \mathcal{I}^n$ and $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If $\mathbf{Q} \prec \mathbf{P}$*, then for any fixed* $\mathbf{x} \in I^k$*,*

$$\sum_{i=1}^{n} J(f, \mathbf{x}, \mathbf{q}_i) \geq_{\mathcal{C}} \sum_{i=1}^{n} J(f, \mathbf{x}, \mathbf{p}_i). \tag{2.61}$$

If f *is* $\leq_{\mathcal{C}}$*-concave, then the inequality (2.61) is reversed.*

Proof: We take $m = n$ and $\mathbf{b} = \mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^n$. From $\mathbf{Q} \prec \mathbf{P}$ we have $\mathbf{Q} = \mathbf{PS}$ for some doubly stochastic matrix $\mathbf{S}$ (cf. (2.43)). By letting $\mathbf{a} = \mathbf{bS}^T$ as in (2.57), we get $\mathbf{a} = \mathbf{eS}^T = \mathbf{e}$. We now apply Corollary 2.4.1.. The desired inequality (2.61) follows from (2.58). This completes the proof. □

It is interesting in (2.61) that for any fixed $\mathbf{x} \in I^k$, the function

$$J(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) = \sum_{i=1}^{n} J(f, \mathbf{x}, \mathbf{p}_i), \quad \mathbf{p}_i \in I,$$

is $\leq_{\mathcal{C}}$-Schur-concave on $\mathcal{I}^n$.

2.4.4. Applications to the Jensen-Mercer functional

In this subsection we assume that $\mathcal{V} = \mathbb{R}^k$, $\mathcal{I} = \mathcal{P}_k^0$ and $\mathcal{W}$ is a linear space with a convex cone $\mathcal{C} \subset \mathcal{W}$ inducing preordering $\leq_{\mathcal{C}}$ on $\mathcal{W}$.

Here $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^k$.

For $\mathbf{e} = (1, 1, \ldots, 1) \in \mathbb{R}^k$ and $f(\mathbf{x}) = (f(x_1), f(x_2), \ldots, f(x_k))$, we use notation

$$[\mathbf{p}, f(\mathbf{x})] = \sum_{i=1}^{k} p_i f(x_i) \in \mathcal{W},$$

$$(f(a) + f(b))\mathbf{e} = (f(a) + f(b), f(a) + f(b), \ldots, f(a) + f(b)) \in \mathcal{W}^k,$$

$$[\mathbf{p}, (f(a) + f(b))\mathbf{e}] = \sum_{i=1}^{k} p_i (f(a) + f(b)) \in \mathcal{W}.$$

Definition 2.4.13.: Let $f : I \to \mathcal{W}$ be a $\mathcal{W}$-valued function with an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{p} = (p_1, p_2, \ldots, p_k) \in \mathcal{I}$ and $\mathbf{x} = (x_1, x_2, \ldots, x_k) \in I^k$. **The Jensen-Mercer**

functional is defined by

$$M(f, \mathbf{x}, \mathbf{p}) = [\mathbf{p}, (f(a) + f(b))\mathbf{e}] - [\mathbf{p}, f(\mathbf{x})] - \langle \mathbf{p}, \mathbf{e} \rangle f \left(\frac{(a+b)\langle \mathbf{p}, \mathbf{e} \rangle - \langle \mathbf{p}, \mathbf{x} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle} \right).$$

Remark 2.4.12.: Clearly,

$$M(f, \mathbf{x}, \mathbf{p}) = [\mathbf{p}, \widetilde{f}(\mathbf{y})] - \langle \mathbf{p}, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{p}, \mathbf{y} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle} \right),$$

where $\mathbf{y} = (a+b)\mathbf{e} - \mathbf{x}$ and $\widetilde{f}(\mathbf{y}) = (f(a) + f(b))\mathbf{e} - f(\mathbf{x})$.

In the next theorem we show a Sherman like inequality (2.64) connected with the function

$$\mathbf{p} \mapsto \langle \mathbf{p}, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{p}, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{p}, \mathbf{e} \rangle} \right), \quad \mathbf{p} \in \mathcal{P}_k^0. \tag{2.62}$$

Theorem 2.4.16*([133])*: *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a, b]$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m) \in \mathcal{I}^m$, $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If

$$\mathbf{Q} = \mathbf{PS} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{2.63}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{mn}(\mathbb{R})$, then for any fixed $\mathbf{x} \in I^k$,

$$\sum_{j=1}^{n} b_j \langle \mathbf{q}_j, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{q}_j, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{q}_j, \mathbf{e} \rangle} \right) \leq_{\mathcal{C}} \sum_{i=1}^{m} a_i \langle \mathbf{p}_i, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{p}_i, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right). \tag{2.64}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.64) is reversed.

Proof: It can be easily shown that the vector $\mathbf{y} = (a+b)\mathbf{e} - \mathbf{x}$ lies in I^k. Thus the function (2.62) is well-defined, and both the sides of the inequality (2.64) are well-defined, too.

By making use of Theorem 2.4.14 for the vector $\mathbf{y} \in I^k$, we find that

$$\sum_{j=1}^{n} b_j \langle \mathbf{q}_j, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{q}_j, \mathbf{y} \rangle}{\langle \mathbf{q}_j, \mathbf{e} \rangle} \right) \leq_{\mathcal{C}} \sum_{i=1}^{m} a_i \langle \mathbf{p}_i, \mathbf{e} \rangle f \left(\frac{\langle \mathbf{p}_i, \mathbf{y} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right).$$

This is the desired inequality (2.64). $\square$

Theorem 2.4.17 *([133])*: *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a,b]$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) \in \mathcal{I}^n$ and $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If $\mathbf{Q} \prec \mathbf{P}$, then or any fixed $\mathbf{x} \in I^k$,

$$\sum_{i=1}^{n} \langle \mathbf{q}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{q}_i, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{q}_i, \mathbf{e} \rangle} \right) \leq_{\mathcal{C}} \sum_{i=1}^{n} \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right). \tag{2.65}$$

Proof: Because $\mathbf{Q} = \mathbf{PS}$ for some doubly stochastic matrix $\mathbf{S}$, it follows that $\mathbf{e} = \mathbf{e}\mathbf{S}^T$. By taking $m = n$ and $\mathbf{a} = \mathbf{b} = \mathbf{e}$, we also get $\mathbf{a} = \mathbf{b}\mathbf{S}^T$. Finally we infer that the required inequality (2.65) follows from (2.64) in Theorem 2.4.16. $\square$

We interpret (2.65) by saying that the function

$$F(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) = \sum_{i=1}^{n} \langle \mathbf{p}_i, \mathbf{e} \rangle f\left(\frac{\langle \mathbf{p}_i, (a+b)\mathbf{e} - \mathbf{x} \rangle}{\langle \mathbf{p}_i, \mathbf{e} \rangle} \right), \quad \mathbf{p}_i \in \mathcal{I}$$

is $\leq_{\mathcal{C}}$-Schur-convex on $\mathcal{I}^n$.

In the next two corollaries we present inequalities for the Jensen-Mercer functional.

Corollary 2.4.3. *([133])*: *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a,b]$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m) \in \mathcal{I}^m$, $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If

$$\mathbf{Q} = \mathbf{PS} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{2.66}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij})$, then for any fixed $\mathbf{x} \in I^k$,

$$\sum_{j=1}^{n} b_j M(f, \mathbf{x}, \mathbf{q}_j) \geq_{\mathcal{C}} \sum_{i=1}^{m} a_i M(f, \mathbf{x}, \mathbf{p}_i). \tag{2.67}$$

If f is $\leq_{\mathcal{C}}$-concave, then the inequality (2.67) is reversed.

Proof: By setting $\mathbf{y} = (a+b)\mathbf{e} - \mathbf{x}$ and $\widetilde{f}(\mathbf{y}) = (f(a) + f(b))\mathbf{e} - f(\mathbf{x})$, we observe from

(2.64) that

$$\sum_{j=1}^{n} b_j \langle \mathbf{q}_j, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{q}_j, \mathbf{y}\rangle}{\langle \mathbf{q}_j, \mathbf{e}\rangle}\right) \leq_{\mathcal{C}} \sum_{i=1}^{m} a_i \langle \mathbf{p}_i, \mathbf{e}\rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{y}\rangle}{\langle \mathbf{p}_i, \mathbf{e}\rangle}\right). \tag{2.68}$$

We also have

$$\begin{aligned} \sum_{j=1}^{n} b_j M(f, \mathbf{x}, \mathbf{q}_j) &= \sum_{j=1}^{n} b_j \left([\mathbf{q}_j, \widetilde{f}(\mathbf{y})] - \langle \mathbf{q}_j, \mathbf{e}\rangle\, f\left(\frac{\langle \mathbf{q}_j, \mathbf{y}\rangle}{\langle \mathbf{q}_j, \mathbf{e}\rangle}\right)\right) \\ &= \left[\sum_{j=1}^{n} b_j \mathbf{q}_j, \widetilde{f}(\mathbf{y})\right] - \left\langle \sum_{j=1}^{n} b_j \mathbf{q}_j, \mathbf{e}\right\rangle f\left(\frac{\langle \mathbf{q}_j, \mathbf{y}\rangle}{\langle \mathbf{q}_j, \mathbf{e}\rangle}\right). \end{aligned} \tag{2.69}$$

Likewise,

$$\begin{aligned} \sum_{i=1}^{m} a_i M(f, \mathbf{x}, \mathbf{p}_i) &= \sum_{i=1}^{m} a_i \left(\left[\mathbf{p}_i, \widetilde{f}(\mathbf{y})\right] - \langle \mathbf{p}_i, \mathbf{e}\rangle\, f\left(\frac{\langle \mathbf{p}_i, \mathbf{y}\rangle}{\langle \mathbf{p}_i, \mathbf{e}\rangle}\right)\right) \\ &= \left[\sum_{i=1}^{m} a_i \mathbf{p}_i, \widetilde{f}(\mathbf{y})\right] - \left\langle \sum_{i=1}^{m} a_i \mathbf{p}_i, \mathbf{e}\right\rangle f\left(\frac{\langle \mathbf{p}_i, \mathbf{y}\rangle}{\langle \mathbf{p}_i, \mathbf{e}\rangle}\right). \end{aligned} \tag{2.70}$$

Due to (2.66), it can be proven that

$$\sum_{j=1}^{n} b_j \mathbf{q}_j = \sum_{i=1}^{m} a_i \mathbf{p}_i.$$

Now, the desired inequality (2.67) is a consequence of (2.68), (2.69) and (2.70). This completes the proof. □

Corollary 2.4.4. *([133]):* *Let $f : I \to \mathcal{W}$ be a $\leq_{\mathcal{C}}$-convex function on an interval $I = [a, b]$. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) \in \mathcal{I}^n$ and $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n) \in \mathcal{I}^n$, where $\mathcal{I} = \mathcal{P}_k^0$.*

If $\mathbf{Q} \prec \mathbf{P}$, then for any fixed $\mathbf{x} \in I^k$,

$$\sum_{i=1}^{n} M(f, \mathbf{x}, \mathbf{q}_i) \geq_{\mathcal{C}} \sum_{i=1}^{n} M(f, \mathbf{x}, \mathbf{p}_i). \tag{2.71}$$

Proof: It is enough to appeal to Corollary 2.4.3. with $\mathbf{a} = \mathbf{b} = \mathbf{e}$, because $\mathbf{Q} = \mathbf{PS}$ and $\mathbf{a} = \mathbf{bS}^T$ for some doubly stochastic matrix $\mathbf{S}$. This leads to (2.71), as wanted. □

The statement (2.71) states that the function

$$M(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n) = \sum_{i=1}^{n} M(f, \mathbf{x}, \mathbf{p}_i), \quad \mathbf{p}_i \in \mathcal{I},$$

is $\leq_{\mathcal{C}}$-Schur-concave on $\mathcal{I}^n$.

2.5. Jessen's functional and majorization

It's well known that Jensen's inequality can be generalized using positive linear functionals acting on a linear class of real valued functions. Such version of Jensen's inequality is called Jessen's inequality. In this section, we focus on Sherman's type inequality related to the Jessen functional deduced from Jessen's inequality. We apply a majorization method. We derive the Hardy-Littlewood-Pólya-Karamata type inequality. It says that some n-sums induced by the Jessen's functional are Schur-concave with respect to its n-tuple of weight vectors. This section is based on [124].

Definition 2.5.14.: Let E be a nonempty set. Assume that $\mathcal{L}(E, \mathbb{R})$ is a linear class of real functions $f : E \to \mathbb{R}$ having the following properties:

$(L1)$ $f, g \in \mathcal{L}(E, \mathbb{R})$ implies $\alpha f + \beta g \in \mathcal{L}(E, \mathbb{R})$ for $\alpha, \beta \in \mathbb{R}$;

$(L2)$ $1 \in \mathcal{L}(E, \mathbb{R})$, i.e., the constant function $f(t) = 1$ for $t \in E$ is in $\mathcal{L}(E, \mathbb{R})$.

A real functional $A : \mathcal{L}(E, \mathbb{R}) \to \mathbb{R}$ is called:

$(A1)$ **linear**, if $A(\alpha f + \beta g) = \alpha A(f) + \beta A(g)$ for all $f, g \in \mathcal{L}(E, \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$;

$(A2)$ **positive**, if $f \in \mathcal{L}(E, \mathbb{R})$, $f(t) \geq 0$ for all $t \in E$ implies $A(f) \geq 0$;

$(A3)$ **normalized**, if $A(1) = 1$.

We introduce the set

$$\mathcal{P}^0_{E,A} = \{p \in \mathcal{L}(E, \mathbb{R}) : p \geq 0,\ A(p) > 0\}.$$

B. Jessen [79] proved the following generalization of Jensen's inequality involving linear positive normalized functionals (see also [138, p. 47]).

Theorem 2.5.18 ***(Jessen's inequality)*:** *Let $\Phi : I \to \mathbb{R}$ be a continuous convex function on interval $I \subset \mathbb{R}$ and A be a normalized positive linear functional on $\mathcal{L}(E, \mathbb{R})$. Then, for all $f \in \mathcal{L}(E, \mathbb{R})$ such that $\Phi(f) \in \mathcal{L}(E, \mathbb{R})$, we have*

$$\Phi(A(f)) \leq A(\Phi(f)). \tag{2.72}$$

Remark 2.5.13.: The inequality (2.72) is called **Jessen's inequality.** Note that $\Phi(A(f))$ in (2.72) is well defined. Namely, if $I = [\alpha, \beta]$, for some $\alpha, \beta \in \mathbb{R}, \alpha < \beta$, then the assumptions $f \in \mathcal{L}(E, \mathbb{R})$ *such that* $\Phi(f) \in \mathcal{L}(E, \mathbb{R})$ implies that $\alpha \leqslant f(t) \leqslant \beta$, for every $t \in E$, and $\alpha = A(\alpha \cdot 1) \leqslant A(f) \leqslant A(\beta \cdot 1) = \beta$, i.e. $A(f) \in I$. So $\Phi(A(f))$ is well defined.

In the above theorem the assumption that Φ is continuous can not be weakened. There are counterexamples which show that Jessen's inequality (2.72) does not hold for some functionals A if Φ is not continuous. For example, if we choose

$$\mathcal{L}(E, \mathbb{R}) = \{f\colon [0,1] \to \mathbb{R} : \lim_{t \to 1^-} f(t) < \infty\}, \quad A(f) = \lim_{t \to 1^-} f(t), \quad \Phi(x) = \begin{cases} 0, & x \in [0,1) \\ 1, & x = 1 \end{cases},$$

then for the identity function $id_{[0,1]}$ on $[0,1]$ we have

$$\Phi(A(id_{[0,1]})) = 1 > A(\Phi(id_{[0,1]})) = 0.$$

In most common examples the continuity condition can be omitted.

Definition 2.5.15. ***([92])*:** For a function $\Phi : I \to \mathbb{R}$ defined on an interval $I \subset \mathbb{R}$, for a positive linear functional A on $\mathcal{L}(E, \mathbb{R})$ and $p \in \mathcal{P}^0_{E,A}$, **the Jessen functional** is defined by

$$J(\Phi, f, p, A) = A(p\Phi(f)) - A(p)\Phi\left(\frac{A(pf)}{A(p)}\right). \tag{2.73}$$

From now on it is assumed that the Jessen functional is well-defined, i.e., $f, pf, p\,\Phi(f) \in \mathcal{L}(E, \mathbb{R})$ and $f \in I^E$.

Clearly, Jessen's inequality (2.72) guarantees that for every continuous convex function

$\Phi : I \to \mathbb{R}$, it holds

$$J(\Phi, f, p, A) \geq 0.$$

Theorem 2.5.19*([92, Theorem 1])*: *Let $\Phi : I = [a, b] \to \mathbb{R}$ be a continuous convex function, A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$, and the map $p \mapsto J(\Phi, f, p, A)$, $p \in \mathcal{P}^0_{E,A}$, be well-defined.*

Then the map $p \mapsto J(\Phi, f, p, A)$ is superadditive, i.e.

$$J(\Phi, f, p+q, A) \geq J(\Phi, f, p, A) + J(\Phi, f, q, A) \quad \textit{for } p, q \in \mathcal{P}^0_{E,A}. \tag{2.74}$$

In consequence, the map $p \mapsto J(\Phi, f, p, A)$ is monotone, i.e.

$$J(\Phi, f, p+q, A) \geq J(\Phi, f, p, A) \quad \textit{for } p, q \in \mathcal{P}^0_{E,A}.$$

Remark 2.5.14.: It is not hard to check that inequality (2.74) amounts to the subadditivity of the function

$$p \to A(p)\Phi\left(\frac{A(pf)}{A(p)}\right) \quad \text{for } p \in \mathcal{P}^0_{E,A},$$

i.e.

$$A(p+q)\Phi\left(\frac{A((p+q)f)}{A(p+q)}\right) \leq A(p)\Phi\left(\frac{A(pf)}{A(p)}\right) + A(q)\Phi\left(\frac{A(qf)}{A(q)}\right) \quad \text{for } p, q \in \mathcal{P}^0_{E,A}.$$

In the next subsection, some conditions will be provided for real scalars $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_n$ and for vectors $p_1, p_2, \ldots, p_m \in \mathcal{P}^0_{E,A}$, $q_1, q_2, \ldots, q_n \in \mathcal{P}^0_{E,A}$, so that the following Sherman type inequality (2.75) will be fulfilled:

$$\begin{aligned} & a_1 J(\Phi, f, p_1, A) + a_2 J(\Phi, f, p_2, A) + \ldots + a_m J(\Phi, f, p_m, A) \\ \leq\ & b_1 J(\Phi, f, q_1, A) + b_2 J(\Phi, f, q_2, A) + \ldots + b_n J(\Phi, f, q_n, A). \end{aligned} \tag{2.75}$$

In consequence, an inequality of the Hardy-Littlewood-Pólya-Karamata type can be obtained for the Jessen's functional, which is a complement of Theorem 2.5.19.

2.5.1. Sherman inequality for Jessen's functional

Throughout we use the abbreviations $\mathcal{L} = \mathcal{L}(E, \mathbb{R})$ and $\mathcal{P} = \mathcal{P}^0_{E,A}$.

Lemma 2.5.4. *([124]): Let $\Phi : I \to \mathbb{R}$ be a convex function on an interval $I = [a, b] \subset \mathbb{R}$. Let $(p_1, p_2, \ldots, p_l) \in \mathcal{P}^l$. Let A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, and $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$.*

If the below expressions are well-defined, then

$$A\left(\sum_{i=1}^{l} p_i\right) \Phi\left(\frac{A\left(\sum_{i=1}^{l} p_i f\right)}{A\left(\sum_{i=1}^{l} p_i\right)}\right) \leq \sum_{i=1}^{l} A(p_i) \Phi\left(\frac{A(p_i f)}{A(p_i)}\right). \tag{2.76}$$

If Φ is concave, then the inequality (2.76) is reversed.

Proof: (See the proof of [92, Theorem 1].) For the coefficients

$$\alpha_i = \frac{A(p_i)}{\sum_{j=1}^{l} A(p_j)} \quad \text{for } i = 1, 2, \ldots, l, \tag{2.77}$$

the following identity holds:

$$\frac{A\left(\sum_{i=1}^{l} p_i f\right)}{A\left(\sum_{i=1}^{l} p_i\right)} = \sum_{i=1}^{l} \alpha_i \frac{A(p_i f)}{A(p_i)}. \tag{2.78}$$

Because $\sum_{i=1}^{l} \alpha_i = 1$ with $\alpha_i \geq 0$ for $i = 1, 2, \ldots, l$, the convexity of $\Phi : I \to \mathbb{R}$ on I and (2.77)-(2.78) imply that

$$\begin{aligned}
\Phi\left(\frac{A\left(\sum_{i=1}^{l} p_i f\right)}{A\left(\sum_{i=1}^{l} p_i\right)}\right) &= \Phi\left(\sum_{i=1}^{l} \alpha_i \frac{A(p_i f)}{A(p_i)}\right) \leq \sum_{i=1}^{l} \alpha_i \Phi\left(\frac{A(p_i f)}{A(p_i)}\right) \\
&= \sum_{i=1}^{l} \frac{A(p_i)}{\sum_{j=1}^{l} A(p_j)} \Phi\left(\frac{A(p_i f)}{A(p_i)}\right) = \frac{1}{\sum_{j=1}^{l} A(p_j)} \sum_{i=1}^{l} A(p_i) \Phi\left(\frac{A(p_i f)}{A(p_i)}\right).
\end{aligned} \tag{2.79}$$

Thus (2.79) leads to (2.76), as wanted. □

Definition 2.5.16.: An n-tuple $\mathbf{Q} = (q_1, q_2, \ldots, q_n) \in \mathcal{L}^n$ is said to **be majorized** by n-tuple $\mathbf{P} = (p_1, p_2, \ldots, p_n) \in \mathcal{L}^n$, written as $\mathbf{Q} \prec \mathbf{P}$, if there exists an $n \times n$ doubly stochastic matrix $\mathbf{S} = (s_{ij})$ such that

$$(q_1, q_2, \ldots, q_n) = (p_1, p_2, \ldots, p_n)\mathbf{S} \tag{2.80}$$

in the sense that

$$q_j = s_{1j}p_1 + s_{2j}p_2 + \ldots + s_{nj}p_n \quad \text{for } j = 1, 2, \ldots, n. \tag{2.81}$$

In the forthcoming theorem we provide a Sherman like inequality (2.83) for the functional $p \mapsto A(p)\Phi\left(\frac{A(pf)}{A(p)}\right)$, $p \in \mathcal{P}$. The basic requirement for (2.83) to hold is condition (2.82) which defines the notion of *weighted majorization* for pairs $(\mathbf{P}, \mathbf{a})$ and $(\mathbf{Q}, \mathbf{b})$ (see [25, 31, 133] for details).

Theorem 2.5.20 ***([124]):*** *Let $\Phi : I \to \mathbb{R}$ be a convex function on an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{P} = (p_1, p_2, \ldots, p_m) \in \mathcal{P}^m$, $\mathbf{Q} = (q_1, q_2, \ldots, q_n) \in \mathcal{P}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{P} = \mathcal{P}_{E,A}^0$ and $m, n \in \mathbb{N}$. Let A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, and $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$.*

If

$$\mathbf{Q} = \mathbf{PS} \quad \textit{and} \quad \mathbf{a} = \mathbf{bS}^T \tag{2.82}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij})$, and the below expressions are well-defined, then

$$\sum_{j=1}^{n} b_j A(q_j)\Phi\left(\frac{A(q_j f)}{A(q_j)}\right) \leq \sum_{i=1}^{m} a_i A(p_i)\Phi\left(\frac{A(p_i f)}{A(p_i)}\right). \tag{2.83}$$

If Φ is concave, then the inequality (2.83) is reversed.

Proof: Observe that $\frac{A(pf)}{A(p)} \in I = [a, b]$ for $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$. Indeed, we have $f(t) \in I$, i.e., $a \leq f(t) \leq b$ for $t \in E$, which means $a1 \leq f \leq b1$ with $\leq$ denoting the componentwise ordering on $\mathcal{L}(E,\mathbb{R})$. Next, we find that

$$\frac{A(pa1)}{A(p)} \leq \frac{A(pf)}{A(p)} \leq \frac{A(pb1)}{A(p)}.$$

Therefore

$$a \leq \frac{A(pf)}{A(p)} \leq b,$$

and further $\frac{A(pf)}{A(p)} \in I$, as claimed.

For this reason the function

$$p \to \Psi(p) = A(p)\Phi\left(\frac{A(pf)}{A(p)}\right), \quad p \in \mathcal{P}^0_{E,A}, \tag{2.84}$$

is well-defined, and both the sides of the inequality (2.83) are well-defined.

It follows that the function (2.84) is convex on $\mathcal{P}$. To see this, we utilize Lemma 2.5.4. for $l = 2$. In fact, for any $p, q \in \mathcal{P}$ and $\alpha \in (0,1)$, by (2.76) applied to the vectors $\alpha p \in \mathcal{P}$ and $(1-\alpha)q \in \mathcal{P}$ in place of p_1 and p_2, we get

$$\begin{aligned}
\Psi(\alpha p + (1-\alpha)q) &= A(\alpha p + (1-\alpha)q)\Phi\left(\frac{A(\alpha p + (1-\alpha)qf)}{A(\alpha p + (1-\alpha)q)}\right) \\
&\leq A(\alpha p)\Phi\left(\frac{A(\alpha pf)}{A(\alpha p)}\right) + A((1-\alpha)q)\Phi\left(\frac{A((1-\alpha)qf)}{A((1-\alpha)q)}\right) \\
&= \alpha A(p)\Phi\left(\frac{A(pf)}{A(p)}\right) + (1-\alpha)A(q)\Phi\left(\frac{A(qf)}{A(q)}\right) \\
&= \alpha\Psi(p) + (1-\alpha)\Psi(q).
\end{aligned} \tag{2.85}$$

Additionally, for $\alpha = 0$ or $\alpha = 1$, inequality (2.85) holds trivially. In summary, the function Ψ is convex on $\mathcal{P}$.

By (2.81)-(2.82), we have $(q_1, q_2, \ldots, q_n) = (p_1, p_2, \ldots, p_m)\mathbf{S}$, and therefore $q_j = \sum\limits_{i=1}^{m} s_{ij}p_i$ with $\sum\limits_{i=1}^{m} s_{ij} = 1$, $j = 1, 2, \ldots, n$, and $s_{ij} \geq 0$. So, the convexity of Ψ yields

$$\Psi(q_j) = \Psi\left(\sum_{i=1}^{m} s_{ij}p_i\right) \leq \sum_{i=1}^{m} s_{ij}\Psi(p_i) \quad \text{for } j = 1, 2, \ldots, n.$$

Since $b_j \geq 0$, we obtain

$$\sum_{j=1}^{n} b_j \Psi(q_j) \leq \sum_{j=1}^{n} b_j \sum_{i=1}^{m} s_{ij} \Psi(p_i) = \sum_{i=1}^{m} \sum_{j=1}^{n} b_j s_{ij} \Psi(p_i).$$

Moreover, $a_i = \sum\limits_{j=1}^{n} b_j s_{ij}$ $(i = 1, 2, \ldots, m)$ by (2.82). So, we find that

$$\sum_{j=1}^{n} b_j \Psi(q_j) \leq \sum_{i=1}^{m} a_i \Psi(p_i).$$

This together with (2.84) leads to inequality (2.83). $\square$

As a corollary to Theorem 2.5.20 we give the following Hardy-Littlewood-Pólya-Karamata like result.

Theorem 2.5.21 *([124]): Let $\Phi : I \to \mathbb{R}$ be a convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (p_1, p_2, \ldots, p_n) \in \mathcal{P}^n$ and $\mathbf{Q} = (q_1, q_2, \ldots, q_n) \in \mathcal{P}^n$, where $\mathcal{P} = \mathcal{P}^0_{E,A}$ and $n \in \mathbb{N}$. Let A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, and $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$.*

If $\mathbf{Q} \prec \mathbf{P}$ and the below expressions are well-defined, then

$$\sum_{i=1}^{n} A(q_i) \Phi\left(\frac{A(q_i f)}{A(q_i)}\right) \leq \sum_{i=1}^{n} A(p_i) \Phi\left(\frac{A(p_i f)}{A(p_i)}\right). \tag{2.86}$$

If f is concave, then the inequality (2.86) is reversed.

Proof: With $m = n$ and $\mathbf{b} = \mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^n$, the vector majorization $(q_1, q_2, \ldots, q_n) \prec (p_1, p_2, \ldots, p_n)$ ensures that $(q_1, q_2, \ldots, q_n) = (p_1, p_2, \ldots, p_n)\mathbf{S}$ for some doubly stochastic matrix $\mathbf{S}$ (see (2.80)).

By putting $\mathbf{a} = \mathbf{b}\mathbf{S}^T$, it follows that $\mathbf{a} = \mathbf{e}\mathbf{S}^T = \mathbf{e}$. So, applying Theorem 2.5.20, Eq. (2.83), provides the wanted inequality (2.86). $\square$

Definition 2.5.17.: A function $F : \mathcal{P}^n \to \mathbb{R}$, with a convex set $\mathcal{P} \subset \mathcal{L}$, is called **Schur-convex** on $\mathcal{P}^n$ if

$$F(\mathbf{Q}) \leq F(\mathbf{P})$$

for all $\mathbf{P}, \mathbf{Q} \in \mathcal{P}^n$ such that

$$\mathbf{Q} \prec \mathbf{P}.$$

A function F is called **Schur-concave** on $\mathcal{P}^n$, if $-F$ is Schur-convex on $\mathcal{P}^n$.

Notice that the statement (2.86) says that the function

$$F(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} A(p_i)\Phi\left(\frac{A(p_i f)}{A(p_i)}\right), \quad p_i \in \mathcal{P},$$

is Schur-convex on $\mathcal{P}^n$.

We now interpret the previous results by using the Jessen's functional.

Corollary 2.5.5. *([124]):* *Let $\Phi : I \to \mathbb{R}$ be a convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (p_1, p_2, \ldots, p_m) \in \mathcal{P}^m$, $\mathbf{Q} = (q_1, q_2, \ldots, q_n) \in \mathcal{P}^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, where $\mathcal{P} = \mathcal{P}_{E,A}^0$ and $m, n \in \mathbb{N}$. Let A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, and $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$.*

If

$$\mathbf{Q} = \mathbf{PS} \quad \textit{and} \quad \mathbf{a} = \mathbf{bS}^T \tag{2.87}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij})$, and the below expressions are well-defined, then

$$\sum_{j=1}^{n} b_j J(\Phi, f, q_j, A) \geq \sum_{i=1}^{m} a_i J(\Phi, f, p_i, A). \tag{2.88}$$

If f is concave, then the inequality (2.88) is reversed.

Proof: From (2.73) we obtain

$$\begin{aligned} \sum_{j=1}^{n} b_j J(\Phi, f, q_j, A) &= \sum_{j=1}^{n} b_j \left(A(q_j \Phi(f)) - A(q_j)\Phi\left(\frac{A(q_j f)}{A(q_j)}\right)\right) \\ &= A\left(\sum_{j=1}^{n} b_j q_j \Phi(f)\right) - \sum_{j=1}^{n} b_j A(q_j)\Phi\left(\frac{A(q_j f)}{A(q_j)}\right) \end{aligned} \tag{2.89}$$

and

$$\sum_{i=1}^{m} a_i J(\Phi, f, p_i, A) = \sum_{i=1}^{m} a_i \left(A(p_i \Phi(f)) - A(p_i) \Phi \left(\frac{A(p_i f)}{A(p_i)} \right) \right) \tag{2.90}$$
$$= A \left(\sum_{i=1}^{m} a_i p_i \Phi(f) \right) - \sum_{i=1}^{m} a_i A(p_i) \Phi \left(\frac{A(p_i f)}{A(p_i)} \right).$$

By a simple computation, from (2.87) we get

$$\sum_{j=1}^{n} b_j q_j = \sum_{i=1}^{m} a_i p_i.$$

Now, by combining (2.89) and (2.90), and employing Theorem 2.5.20, we conclude that inequality (2.88) is satisfied, as was to be proved. □

With the help of a doubly stochastic matrix $\mathbf{S}$ and $\mathbf{a} = \mathbf{b} = \mathbf{e}$ for the all ones vector $\mathbf{e}$, we get the following result from Corollary 2.5.5..

Corollary 2.5.6. *([124]): Let $\Phi : I \to \mathbb{R}$ be a convex function on a closed interval $I \subset \mathbb{R}$. Let $\mathbf{P} = (p_1, p_2, \ldots, p_n) \in \mathcal{P}^n$ and $\mathbf{Q} = (q_1, q_2, \ldots, q_n) \in \mathcal{P}^n$, where $\mathcal{P} = \mathcal{P}^0_{E,A}$ and $n \in \mathbb{N}$. Let A be a positive linear functional on $\mathcal{L}(E, \mathbb{R})$, and $f \in I^E \cap \mathcal{L}(E, \mathbb{R})$.*

If $\mathbf{Q} \prec \mathbf{P}$ and the below expressions are well-defined, then

$$\sum_{i=1}^{n} J(\Phi, f, q_i, A) \geq \sum_{i=1}^{n} J(\Phi, f, p_i, A). \tag{2.91}$$

If Φ is concave, then the inequality (2.91) is reversed.

Proof: Since $\mathbf{Q} \prec \mathbf{P}$, we obtain $\mathbf{Q} = \mathbf{PS}$ for some doubly stochastic matrix $\mathbf{S}$ (cf. (2.80)). By setting $m = n$, $\mathbf{b} = \mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^n$, and $\mathbf{a} = \mathbf{b}\mathbf{S}^T$ (see 2.87), we find that $\mathbf{a} = \mathbf{e}\mathbf{S}^T = \mathbf{e}$. By applying inequality (2.88) in Corollary 2.5.5., we get the inequality (2.91), as required. □

It is worth to mention that the statement (2.91) asserts that the function

$$J(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} J(\Phi, f, p_i, A), \quad p_i \in I,$$

is Schur-concave on $\mathcal{P}^n$.

Remark 2.5.15.: Corollaries 2.5.5. and 2.5.6. extend similar results proved for the Jensen's functional and Jensen-Mercer's functional (see [42, 95, 127, 133] for details).

2.6. Sherman like inequalities for (α, β)-convex functions

In this section, a Sherman like inequality is presented for the class of (α, β)-convex functions. In particular, Sherman's type results corresponding to the Jensen-Steffensen, the Mercer-Steffensen and Brunk's inequalities are established. The results of this section are cited from [130].

2.6.1. Sherman type results for (α, β)-convex functions

In this section we again denote by $\langle \cdot, \cdot \rangle$ the standard inner product on $\mathbb{R}^m$.

Definition 2.6.18.*([130])*: For given vectors $\alpha, \beta \in \mathbb{R}^m$, we say that a function $f : I \to \mathbb{R}$, defined on an interval $I \subset \mathbb{R}$, is (α, β)-**convex** on a set $S \subset I^m$ if the following inequality holds

$$f(\langle \alpha, \mathbf{x} \rangle) \leq \langle \beta, f(\mathbf{x}) \rangle \quad \text{for } \mathbf{x} \in S, \tag{2.92}$$

where $f(\mathbf{x}) = (f(x_1), f(x_2), \ldots, f(x_m))$ for $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in S$. A function f is (α, β)-**concave** on $S \subset I^m$ if reversed inequality in (2.92) holds.

Functions satisfying inequality (2.92) (with $\alpha = \beta$ or $\alpha \neq \beta$) can be found among convex functions, subadditive functions, starshaped functions, Breckner's s-convex functions [26, p. 254], Godunova-Levin functions [155], P-functions [136, 155], h-convex functions [155], Toader's c-convex functions [89], (k, h)-convex functions [104, 108, 127], etc. For more details see the following example.

Example 3.: *(i) Jensen's inequality (1.6) says that if $f : I \to \mathbb{R}$ is a convex function on an interval $I \subset \mathbb{R}$, $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in S = I^m$ and $p_i \geq 0, i = 1, ..., m$, with*

$P_m = \sum_{i=1}^{m} p_i > 0$, then

$$f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \leq \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i). \tag{2.93}$$

Clearly, each convex function is (α, β)-convex for $\alpha = (\alpha_1, \ldots, \alpha_m)$ and $\beta = (\beta_1, \ldots, \beta_m)$ with $\alpha_i = \beta_i = \frac{p_i}{P_m}$, $i = 1, \ldots, m$.

(ii) Jensen-Steffensen inequality (see [5], [138, p. 57] states that if $f : I \to \mathbb{R}$ is a convex function on an interval $I \subset \mathbb{R}$, such that $[a, b] \subset I$ with $a < b$, and if $a \leq x_1 \leq x_2 \leq \ldots \leq x_m \leq b$ and

$$0 \leq W_i \leq W_m\,, \quad W_m > 0 \quad \text{for } i = 1, \ldots, m, \tag{2.94}$$

where $W_i = \sum_{j=1}^{i} w_j$, $i = 1, \ldots, m$, then

$$f\left(\frac{1}{W_m}\sum_{j=1}^{m} w_j x_j\right) \leq \frac{1}{W_m}\sum_{j=1}^{m} w_j f(x_j). \tag{2.95}$$

Inequalities (2.94) are referred to as Steffensen's condition. Here $S = \{\mathbf{x} = (x_1, \ldots, x_m) \in I^m : a \leq x_1 \leq \ldots \leq x_m \leq b\}$. It is easily seen that each convex function is (α, β)-convex for $\alpha = (\alpha_1, \ldots, \alpha_m)$ and $\beta = (\beta_1, \ldots, \beta_m)$ with $\alpha_i = \beta_i = \frac{w_i}{W_m}$, $i = 1, \ldots, m$, satisfying (2.94).

(iii) Brunk's inequality [27] states that if $f : [0, x_1] \to \mathbb{R}$ is a convex function on $[0, x_1]$ with $f(0) \leq 0$, and if $1 \geq h_1 \geq h_2 \geq \ldots \geq h_m \geq 0$ and $x_1 \geq x_2 \geq \ldots \geq x_m \geq 0$, then

$$f\left(\sum_{i=1}^{m}(-1)^{i-1} h_i x_i\right) \leq \sum_{i=1}^{m}(-1)^{i-1} h_i f(x_i). \tag{2.96}$$

(iv) Jensen-Mercer inequality [107, Theorem 1.2] states that if $f : I \to \mathbb{R}$ is a convex function on an interval $I \subset \mathbb{R}$, and if $x_i \in I$, $i = 1, \ldots, m$, such that $0 < x_1 \leq x_2 \leq \ldots \leq x_m$, then

$$f(x_1 - \sum_{i=1}^{m} w_i x_i + x_m) \leq f(x_1) - \sum_{i=1}^{m} w_i f(x_i) + f(x_m), \tag{2.97}$$

where $\sum_{i=1}^{m} w_i = 1$ with $w_i > 0, i = 1, ..., m$.

(v) Mercer-Steffensen inequality [5, Theorem 2] states that if $f : I \to \mathbb{R}$ is a convex function on an interval I in $\mathbb{R}$, $[a, b] \subset I$ with $a < b$ and $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{w} = (w_1, \ldots, w_m)$ are real m-tuples such that $a \leq x_1 \leq x_2 \leq \ldots \leq x_m \leq b$ and

$$w_i \neq 0 \quad \text{and} \quad 0 \leq W_i \leq W_m \,, \quad W_m > 0 \quad \text{for } i = 1, \ldots, m, \tag{2.98}$$

where $W_i = \sum_{j=1}^{i} w_j$, $i = 1, \ldots, m$, then

$$f(a - \frac{1}{W_m} \sum_{i=1}^{m} w_i x_i + b) \leq f(a) - \frac{1}{W_m} \sum_{i=1}^{m} w_i f(x_i) + f(b). \tag{2.99}$$

(vi) Generalized Mercer's type inequality [121, Theorem 2.1] states that if $f : I \to \mathbb{R}$ is a continuous convex function on interval $I \subset \mathbb{R}$, $\mathbf{a} = (a_1, \ldots, a_n) \in I^n$ and $\mathbf{X} = (x_{ij})$ is a real $m \times n$ matrix such that $x_{ij} \in I$ for all i, j. If $\mathbf{a}$ majorizes each row of $\mathbf{X}$, i.e., $(x_{i1}, \ldots, x_{in}) \prec (a_1, \ldots, a_n)$ for each $i = 1, \ldots, m$, then

$$f\left(\sum_{j=1}^{n} a_j - \sum_{j=1}^{n-1} \sum_{i=1}^{m} w_i x_{ij}\right) \leq \sum_{j=1}^{n} f(a_j) - \sum_{j=1}^{n-1} \sum_{i=1}^{m} w_i f(x_{ij}), \tag{2.100}$$

where $\sum_{i=1}^{m} w_i = 1$ with $w_i \geq 0$.

(vii) Let $c \in [0, 1]$. Toader's c-convex functions are functions $f : I = [0, b] \to \mathbb{R}$ defined by

$$f(t_1 x_1 + \ldots + t_{m-1} x_{m-1} + c t_m x_m) \leq t_1 f(x_1) + \ldots + t_{m-1} f(x_{m-1}) + c t_m f(x_m) \tag{2.101}$$

for $(x_1, \ldots, x_m) \in S = [0, b]^m$ and $t_1, \ldots, t_m \in [0, 1]$ with $\sum_{i=1}^{m} t_i = 1$ (see [89]).

As noted in [89], (2.101) provides the standard convexity for $c = 1$, and the starshapedness for $n = 2$ and $c = 0$. It is obvious that Toader's c-functions are (α, β)-convex with $\alpha = \beta = (t_1, t_2, \ldots, c t_m)$.

(viii) Let $s \in (0, 1]$. Breckner's s-convex functions are functions $f : I \to \mathbb{R}_+$, $I \subset \mathbb{R}$, defined by

$$f(t_1 x_1 + \ldots + t_m x_m) \leq t_1^s f(x_1) + \ldots + t_m^s f(x_m) \tag{2.102}$$

for all $(x_1, \ldots, x_m) \in S = I^m$ *and* $t_1, \ldots, t_m \in (0,1)$ *with* $t_1 + \ldots + t_m = 1$ *(see [26, p. 254], cf. [155, p. 304]).*

Each s-convex function is (α, β)*-convex for* $\alpha = (\alpha_1, \ldots, \alpha_m)$ *and* $\beta = (\beta_1, \ldots, \beta_m)$ *with* $\alpha_i = t_i$ *and* $\beta_i = t_i^s$, $i = 1, \ldots, m$.

(ix) P-functions$f : I \to \mathbb{R}_+$, *where* $I \subset \mathbb{R}$, *are characterized by the inequality*

$$f(t_1x_1 + \ldots + t_mx_m) \leq f(x_1) + \ldots + f(x_m) \tag{2.103}$$

for all $(x_1, \ldots, x_m) \in S = I^m$ *and* $t_1, \ldots, t_m \in (0,1)$ *with* $\sum\limits_{i=1}^{m} t_i = 1$ *(cf. [136],[155, p. 304]).*

Each P-convex function is (α, β)*-convex for* $\alpha = (\alpha_1, \ldots, \alpha_m)$ *and* $\beta = (\beta_1, \ldots, \beta_m)$ *with* $\alpha_i = t_i$ *and* $\beta_i = 1$, $i = 1, \ldots, m$.

(x) Godunova-Levin functions$f : I \to \mathbb{R}_+$, *where* $I \subset \mathbb{R}$, *are described by the condition*

$$f(t_1x_1 + \ldots + t_mx_m) \leq \frac{1}{t_1}f(x_1) + \ldots + \frac{1}{t_m}f(x_m) \tag{2.104}$$

for all $(x_1, \ldots, x_m) \in S = I^m$ *and* $t_1, \ldots, t_m \in (0,1)$ *with* $\sum\limits_{i=1}^{m} t_i = 1$ *(cf. [155, p. 303]).*

Therefore each Godunova-Levin function is (α, β)*-convex for* $\alpha = (\alpha_1, \ldots, \alpha_m)$ *and* $\beta = (\beta_1, \ldots, \beta_m)$ *with* $\alpha_i = t_i$ *and* $\beta_i = \frac{1}{t_i}$, $i = 1, \ldots, m$.

(xi) Let $h : (0,1) \to \mathbb{R}_+$ *be a given function. An h-convex function* $f : I \to \mathbb{R}_+$, $I \subset \mathbb{R}$, *is defined by the inequality*

$$f(t_1x_1 + \ldots + t_mx_m) \leq h(t_1)f(x_1) + \ldots + h(t_m)f(x_m) \tag{2.105}$$

for all $(x_1, \ldots, x_m) \in S = I^m$ *and* $t_1, \ldots, t_m \in (0,1)$ *with* $\sum\limits_{i=1}^{m} t_i = 1$ *(cf. [155, p. 304]).*

It is clear that each h-convex function is (α, β)*-convex for* $\alpha = (\alpha_1, \ldots, \alpha_m)$ *and* $\beta = (\beta_1, \ldots, \beta_m)$ *with* $\alpha_i = t_i$ *and* $\beta_i = h(t_i)$, $i = 1, \ldots, m$.

The following theorem provides a result of Sherman type for (α, β)-convex functions.

Theorem 2.6.22 *([130])*: *Assume that* $\mathbf{A} = (a_{ij})$ *and* $\mathbf{B} = (b_{ij})$ *are* $n \times m$*-matrices with ith rows* α_i *and* β_i, $i = 1, 2, \ldots, n$, *respectively. Let* $f : I \to \mathbb{R}$ *be an* (α_i, β_i)*-convex* *function on non-empty set $S \subset I^m$, where $I \subset \mathbb{R}$, i.e.*

$$f(\langle \alpha_i, \mathbf{x}\rangle) \leq \langle \beta_i, f(\mathbf{x})\rangle \quad \textit{for } \mathbf{x} \in S, \quad i = 1, 2, \ldots, n. \tag{2.106}$$

Let $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in S$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in I^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n_+$.

If

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{B} \tag{2.107}$$

then

$$\sum_{i=1}^{n} b_i f(y_i) \leq \sum_{j=1}^{m} a_j f(x_j). \tag{2.108}$$

If f is (α_i, β_i)-concave, $i = 1, 2, \ldots, n$, then the inequality (2.108) is reversed.

Proof: We have

$$\sum_{i=1}^{n} b_i f(y_i) = \langle \mathbf{b}, f(\mathbf{y})\rangle, \tag{2.109}$$

where $f(\mathbf{y}) = (f(y_1), f(y_2), \ldots, f(y_n))$.

Since $\mathbf{y} = \mathbf{x}\mathbf{A}^T$, we can write

$$y_i = \sum_{j=1}^{m} a_{ij} x_j, \quad i = 1, 2, \ldots, n.$$

It follows from (2.106) that

$$f(y_i) = f\left(\sum_{j=1}^{m} a_{ij} x_j\right) \leq \sum_{j=1}^{m} b_{ij} f(x_j), \quad i = 1, 2, \ldots, n.$$

Therefore,

$$f(\mathbf{y}) \leq_C \left(\sum_{j=1}^{m} b_{1j} f(x_j), \sum_{j=1}^{m} b_{2j} f(x_j), \ldots, \sum_{j=1}^{m} b_{nj} f(x_j)\right) = f(\mathbf{x})\mathbf{B}^T,$$

where $f(\mathbf{x}) = (f(x_1), f(x_2), \ldots, f(x_m))$ and $\leq_C$ denotes the componentwise preorder on $\mathbb{R}^n$.

From (2.109) and (2.107) and by the positivity of $\mathbf{b} \in \mathbb{R}_+^n$, we derive

$$\begin{aligned}\sum_{i=1}^{n} b_i f(y_i) &= \langle \mathbf{b}, f(\mathbf{y})\rangle \leq \langle \mathbf{b}, f(\mathbf{x})\mathbf{B}^T\rangle \\ &= \langle \mathbf{bB}, f(\mathbf{x})\rangle = \langle \mathbf{a}, f(\mathbf{x})\rangle = \sum_{j=1}^{m} a_j f(x_j),\end{aligned}$$

which completes the proof. □

Definition 2.6.19. ***(Cf. [108, Def. 2.1], [127], [130])*:** Let $k : (0, 1) \to \mathbb{R}$ be a given function and $m \geq 2$ be a given positive integer. Then a set $I \subset \mathbb{R}$ is said to be $(k; m)$**-convex** if

$$k(t_1)x_1 + \ldots + k(t_m)x_m \in I$$

for all $x_1, \ldots, x_m \in I$ and $t_1, \ldots, t_m \in (0, 1)$ with $t_1 + \ldots + t_m = 1$.

Definition 2.6.20. ***(Cf. [108, Def. 2.4], [26, p. 254], [127], [130])*:** Let $k, h : (0, 1) \to \mathbb{R}$ be two given functions and $m \geq 2$ be a given positive integer. A function $f : I \to \mathbb{R}$ defined on a $(k; m)$-convex set $I \subset \mathbb{R}$ is said to be $(k, h; m)$**-convex** if

$$f(k(t_1)x_1 + \ldots + k(t_m)x_m) \leq h(t_1)f(x_1) + \ldots + h(t_m)f(x_m)$$

for all $x_1, \ldots, x_m \in I$ and $t_1, \ldots, t_m \in (0, 1)$ with $t_1 + \ldots + t_m = 1$.

For a given function $g : (0, 1) \to \mathbb{R}$, and an $n \times m$ matrix $\mathbf{T} = (t_{ij})$ such that $t_{ij} \in (0, 1)$ for all i, j, we define $g(\mathbf{T})$ to be the matrix $(g(t_{ij}))$.

Corollary 2.6.7. ***([130])*:** *Let $k, h : (0, 1) \to \mathbb{R}$ be two given functions and $m \geq 2$ be a given positive integer. Assume that f is a real $(k, h; m)$-convex function defined on a $(k; m)$-convex set $I \subset \mathbb{R}$.*

If $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in I^m$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in I^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, and

$$\mathbf{y} = \mathbf{x}(k(\mathbf{T}))^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}(h(\mathbf{T}))$$

for some $n \times m$ row stochastic matrix $\mathbf{T}$, *then*

$$\sum_{i=1}^{n} b_i f(y_i) \leq \sum_{j=1}^{m} a_j f(x_j). \tag{2.110}$$

Proof: It is enough to invoke to Theorem 2.6.22 with $\mathbf{A} = k(\mathbf{T})$ and $\mathbf{B} = h(\mathbf{T})$. $\square$

Example 4. *([130])*: *By taking $k(t) = t$ and $h(t) = t^s$, $t, s \in (0,1)$, we use Corollary 2.6.20. to Breckner's s-convex functions $f : (0,\infty) \to \mathbb{R}_+$ defined on $I = (0,\infty)$, $s \in (0,1)$ (see (2.102)).*

If $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in I^m$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in I^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$, and

$$\mathbf{y} = \mathbf{x}\mathbf{T}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{T}^s \tag{2.111}$$

for some $n \times m$ row stochastic matrix $\mathbf{T} = (t_{ij})$, then (2.110) is met with (2.111).

2.6.2. The case $\mathbf{A} = \mathbf{B}$

Our purpose in this subsection is to give some applications of Theorem 2.6.22 when $\mathbf{A} = \mathbf{B}$.

Remark 2.6.16.: Theorem 2.1.1 is a special case of Majorization theorem (Theorem 2.6.22) with $\mathbf{A} = \mathbf{B}$ related to the classical Jensen's inequality (2.93).

The forthcoming result is connected with the Jensen-Steffensen inequality (2.95).

Corollary 2.6.8. *([130])*: *Let $f : I \to \mathbb{R}$ be a real convex function defined on an interval $I = [a,b] \subset \mathbb{R}$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in I^m$ with $a \leq x_1 \leq x_2 \leq \ldots \leq x_m \leq b$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in I^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$.*

If $\mathbf{y} = \mathbf{x}\mathbf{A}^T$ and $\mathbf{a} = \mathbf{b}\mathbf{A}$ for some $n \times m$ matrix $\mathbf{A} = (a_{ij})$ with rows satisfying Steffensen's condition

$$0 \leq \sum_{j=1}^{k} a_{ij} \leq 1 = \sum_{j=1}^{m} a_{ij} \quad \textit{for } i = 1, 2, \ldots, n, \ k = 1, 2, \ldots, m, \tag{2.112}$$

then (2.108) holds.

Proof: Thanks to the Jensen-Steffensen inequality (2.95) for convex f, we have the (α_i, α_i)-convexity of f, i.e.,

$$f(\langle\alpha_i, \mathbf{x}\rangle) \le \langle\alpha_i, f(\mathbf{x})\rangle, \quad i = 1, 2, \dots, n,$$

where α_i is the ith row of $\mathbf{A}$. Now, it suffices to appeal to Theorem 2.6.22 with $\mathbf{A} = \mathbf{B}$. □

We now give an extension of the inequality of Brunk (2.96).

Example 5. *([130])*: *Let f be a real convex function defined on interval $I = [0, x_1]$ with $f(0) \le 0$. Assume that $\mathbf{x} = (x_1, x_2, \dots, x_m) \in I^m$ with $x_1 \ge x_2 \ge \dots \ge x_m > 0$, and $\mathbf{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n_+$.*

If $\mathbf{A} = (a_{ij}) = ((-1)^{j-1}h_{ij})$ is an $n \times m$ matrix with rows such that

$$1 \ge h_{i1} \ge h_{i2} \ge \dots \ge h_{in} \ge 0 \quad \text{for } i = 1, 2, \dots, n,$$

then

$$\sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{m}(-1)^{j-1}h_{ij}x_j\right) \le \sum_{j=1}^{m} a_j f(x_j),$$

where $\mathbf{a} = \mathbf{bA}$, i.e., $a_j = (-1)^{j-1}\sum_{i=1}^{n} b_i h_{ij}$.

In the next example we use Toader's c-convexity (2.101) with $m = 2$.

Example 6. *([130])*: *Let f be a real Toader's c-convex function defined on an interval $I = [0, b) \subset \mathbb{R}$. Suppose that $\mathbf{x} = (x_1, x_2) \in I^2$ and $\mathbf{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n_+$.*

If $\mathbf{A}$ is an $n \times 2$ matrix with rows $\alpha_i = (a_{i1}, ca_{i2})$ such that

$$a_{i1} + a_{i2} = 1, \quad a_{i1}, a_{i2} \ge 0 \quad \text{for } i = 1, 2, \dots, n,$$

then

$$\sum_{i=1}^{n} b_i f\left(a_{i1}x_1 + ca_{i2}x_2\right) \le a_1 f(x_1) + a_2 f(x_2),$$

where $a_1 = \sum_{i=1}^{n} b_i a_{i1}$ *and* $a_2 = c\sum_{i=1}^{n} b_i a_{i2}$.

We now present an extension of the Mercer-Steffensen inequality (2.99).

Corollary 2.6.9. *([130]):* *Let $f : I \to \mathbb{R}$ be a real convex function defined on an interval $I = [a, b] \subset \mathbb{R}$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m) \in I^m$ with $a \le x_1 \le x_2 \le \ldots \le x_m \le b$, and $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n_+$.*

If $\mathbf{A} = (a_{ij})$ is an $n \times m$ matrix with rows satisfying Steffensen's condition (2.112), then

$$\sum_{i=1}^{n} b_i f(a - \sum_{j=1}^{m} a_{ij}x_j + b) \le \sum_{i=1}^{n} b_i f(a) - \sum_{j=1}^{m} a_j f(x_j) + \sum_{i=1}^{n} b_i f(b). \tag{2.113}$$

where $a_j = \sum_{i=1}^{n} a_{ij} b_i$.

Proof: The symbol $\mathbf{1}_n$ stands for the (column) n-tuple of ones. By setting $\mathbf{y} = (y_1, y_2, \ldots, y_n) = \mathbf{x}\mathbf{A}^T \in I^n$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) = \mathbf{b}\mathbf{A} \in \mathbb{R}^m$, $\mathbf{y} = a\mathbf{1}_n^T - \mathbf{y} + b\mathbf{1}_n^T$, $\tilde{\mathbf{b}} = \mathbf{b}$, and

$$\tilde{\mathbf{A}} = (\mathbf{1}_n, -\mathbf{A}, \mathbf{1}_n) \quad \text{and} \quad \tilde{\mathbf{x}} = (a, \mathbf{x}, b) = (a, x_1, x_2, \ldots, x_m, b) \in I^{m+2},$$

we find that $\tilde{\mathbf{y}} = \tilde{\mathbf{x}}\tilde{\mathbf{A}}^T$ and $\tilde{\mathbf{a}} = \tilde{\mathbf{b}}\tilde{\mathbf{A}}$.

By the Mercer-Steffensen inequality (2.99) we establish

$$f(a - \langle \alpha_i, \mathbf{x}\rangle + b) \le f(a) - \langle \alpha_i, f(\mathbf{x})\rangle + f(b), \quad i = 1, 2, \ldots, n,$$

where α_i denotes the ith row of the matrix $\mathbf{A}$. Thus we get

$$f(\langle \tilde{\alpha}_i, \tilde{\mathbf{x}}\rangle) \le \langle \tilde{\alpha}_i, f(\tilde{\mathbf{x}})\rangle, \quad i = 1, 2, \ldots, n,$$

where $\tilde{\alpha}_i$ denotes the ith row of the matrix $\tilde{\mathbf{A}}$. Hence f is $(\tilde{\alpha}_i, \tilde{\alpha}_i)$-convex.

Now, by making use of Theorem 2.6.22 we deduce that

$$\sum_{i=1}^{n} \tilde{b}_i f(\tilde{y}_i) \le \sum_{j=1}^{m+2} \tilde{a}_j f(\tilde{x}_j). \tag{2.114}$$

Consequently, for $a_j = \sum\limits_{i=1}^{n} a_{ij} b_i$,

$$\sum_{i=1}^{n} b_i f(\tilde{y}_i) \leq (\sum_{i=1}^{n} b_i) f(a) - \sum_{j=1}^{m} a_j f(x_j) + (\sum_{i=1}^{n} b_i) f(b), \tag{2.115}$$

because

$$\tilde{y}_i = a - y_i + b = a - \sum_{j=1}^{m} a_{ij} x_j + b,$$

$$\tilde{a}_1 = \tilde{a}_{m+2} = \sum_{i=1}^{n} b_i, \quad \tilde{a}_{j+1} = -\sum_{i=1}^{n} a_{ij} b_i = -a_j, \quad j = 1, 2, \ldots, m.$$

Therefore inequality (2.113) follows from (2.114)-(2.115). This completes the proof. □

2.7. Sherman like inequalities for convex sequences and non-decreasing convex functions

Here our aim is to prove Sherman like inequalities for convex sequences and nondecreasing convex functions. To do so, we develop some results by S. Wu and L. Debnath [159]. We derive discrete versions of Petrović' and Giaccardi's inequalities. We obtain some generalization of Fejér inequality for convex sequences. We also deal with inequalities of the Hermite-Hadamard type. We extend some recent results of Latreuch and Belaïdi [102]. The results of the present section are quoted from [131].

2.7.1. Introduction and preliminaries

In this expository subsection we demonstrate some relevant notation and facts for convex functions and sequences.

We begin this presentation with Fejér inequality (2.116) for convex functions (see [2, 88, 110]).

Theorem 2.7.23 *([2])***:** *Let $f : I \to \mathbb{R}$ be a convex function, $a, b \in I, a < b$, and let $p : [a, b] \to \mathbb{R}$ be a nonnegative integrable weight on I. Assume that p is symmetric about*

$\frac{a+b}{2}$. *Then*

$$f\left(\frac{a+b}{2}\right)\int_a^b p(t)dt \le \int_a^b f(t)p(t)dt \le \frac{f(a)+f(b)}{2}\int_a^b p(t)dt. \tag{2.116}$$

Remark 2.7.17.: By letting $p(t) = 1$ for $t \in [a,b]$ into (2.116), we obtain *Hermite-Hadamard inequality* (2.117). That is, if $f : I \to \mathbb{R}$ is a continuous convex function on an interval $I \subset \mathbb{R}$, $a, b \in I$ with $a < b$, then

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(x)\,dx \le \frac{f(a)+f(b)}{2}. \tag{2.117}$$

For more informations see [26, p. 137], [119, 138].

In what follows, we are interesting in inequalities for *convex sequences* (see [120, 139, 159]).

Definition 2.7.21.*([159, p. 526])*: A sequence $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{R}^n$ is said to be **convex sequence** if

$$z_i \le \frac{z_{i-1}+z_{i+1}}{2} \quad \text{for } i = 2, \ldots, n-1.$$

For convex sequence $z = (z_1, ..., z_n)$ we consider the function $\varphi_{\mathbf{z}} : [1, n] \to \mathbb{R}$ defined by

$$\varphi_{\mathbf{z}}(t) = \begin{cases} z_1 + (z_2 - z_1)(t-1), & t \in [1,2), \\ z_2 + (z_3 - z_2)(t-2), & t \in [2,3), \\ \ldots & \ldots \\ z_i + (z_{i+1} - z_i)(t-i), & t \in [i, i+1), \\ \ldots & \ldots \\ z_{n-1} + (z_n - z_{n-1})(t-n+1), & t \in [n-1, n], \end{cases} \tag{2.118}$$

(see [159]).

Note that

$$\varphi_{\mathbf{z}}(r) = z_r, \quad \text{for } r = 1, 2, \ldots, n.$$

In the following theorem, we can see that

$$\varphi_{\mathbf{z}}(p_j) \;=\; z_{p_j}, \quad \text{for } j = 1, 2, \ldots, k,$$
$$\varphi_{\mathbf{z}}(q_i) \;=\; z_{q_i}, \quad \text{for } i = 1, 2, \ldots, m,$$

since the entries of $\mathbf{p}$ and $\mathbf{q}$ are positive integers from the set $\{1, 2, \ldots, n\}$.

By employing Majorization inequality (2.4), Wu and Debnath [159] proved the following result.

Theorem 2.7.24 ***([159, Theorems 1 and 2]):*** *Let $(z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence, and $\psi : I \to R$ be a continuous increasing convex function on I.*

Then for any $(p_1, p_2, \ldots, p_k) \prec (q_1, q_2, \ldots, q_k)$ ($1 \le p_i \le n$, $1 \le q_i \le n$, $p_i, q_i \in N$, $i = 1, 2, \ldots, k$, $k \ge 2$), the following inequality holds

$$\sum_{i=1}^{k} \psi(z_{p_i}) \le \sum_{i=1}^{k} \psi(z_{q_i}). \tag{2.119}$$

In particular, if ψ is the identity function $\psi(t) = t$, $t \in I = R$, then (2.119) reduces to

$$\sum_{i=1}^{k} z_{p_i} \le \sum_{i=1}^{k} z_{q_i}. \tag{2.120}$$

Latreuch and Belaïdi [102, Theorem 1.4] proved discrete counterpart of Fejér and Hermite-Hadamard inequalities (2.116)-(2.117), as follows.

Theorem 2.7.25 ***([102, Theorem 1.4]):*** *Let $(z_1, z_2, \ldots, z_n)$ be a convex sequence of real numbers and $b = (b_1, b_2, \ldots, b_n)$ be a positive sequence symmetric about $\frac{n+1}{2}$.*

Then the following inequalities hold

$$\frac{z_N + z_{n+1-N}}{2} \sum_{i=1}^{n} b_i \le \sum_{i=1}^{n} b_i z_i \le \frac{z_1 + z_n}{2} \sum_{i=1}^{n} b_i, \tag{2.121}$$

where $N = \left[\frac{n+1}{2}\right]$ is the integer part of $\frac{n+1}{2}$.

In particular, if $b = \frac{1}{n}(1, 1, \ldots, 1)$ *then (2.121) becomes*

$$\frac{z_N + z_{n+1-N}}{2} \le \frac{1}{n}\sum_{i=1}^{n} z_i \le \frac{z_1 + z_n}{2}.$$

Remark 2.7.18.: For odd n the above inequality means that

$$z_{\frac{n+1}{2}} \le \frac{1}{n}\sum_{i=1}^{n} z_i \le \frac{z_1 + z_n}{2}.$$

However, for even n the index $t = \frac{n+1}{2}$ does not belong to the index set $\{1, 2, \ldots, n\}$. So, we must interpolate the symbol z_t with $t \in [1, n] \setminus \{1, 2, \ldots, n\}$. Therefore, by (2.118),

$$z_t = \psi_{\mathbf{z}}(t) = z_i + (z_{i+1} - z_i)(t - i),$$

whenever $t \in [i, i+1)$ for some $i \in \{1, 2, \ldots, n\}$.

Thus for even n we obtain $z_{\frac{n+1}{2}} = \frac{z_N + z_{n+1-N}}{2}$ (cf. (2.121)).

In the forthcoming Subsection 2.7.2., we prove a version of Theorem 2.2.4 for convex sequences (see Theorem 2.7.26). To this end, we use the method of Wu and Debnath mentioned above.

Subsequently, by applying Theorem 2.7.26 and Corollary 2.7.10., we generalize Theorem 2.7.25 by relaxing the majorization assumption on sequences $\mathbf{p}$ and $\mathbf{q}$. In consequence, we derive some discrete versions for convex sequences of Petrović and Giaccardi's inequalities [137].

Lastly, we give some applications of the obtained results (see Subsection 2.7.3.). In particular, we show an extension for convex sequences of Fejér and Hermite-Hadamard inequalities. Thus we develop some recent results by Latreuch and Belaïdi (see Theorem 2.7.24).

In our further studies we utilize column stochastic matrices as in Theorem 2.2.4.

Our interest is in double estimations (as in (2.116), (2.117) and (2.121)). However, for the sake of simplicity, we consider upper and lower bounds separately.

2.7.2. Sherman theorem for convex sequences

We now present a discrete counterpart of Theorem 2.2.4 for convex sequences.

Theorem 2.7.26*([131])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Let $p_j, q_i \in \{1, 2, \ldots, n\}$, $j = 1, 2, \ldots, k$, $i = 1, 2, \ldots, m$. Assume that $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, $\mathbf{q} = (q_1, q_2, \ldots, q_m)$, $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m$ and $\mathbf{b} = (b_1, b_2, \ldots, b_k) \in \mathbb{R}_+^k$.*

If

$$\mathbf{p} = \mathbf{q}\,\mathbf{S} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\,\mathbf{S}^T \tag{2.122}$$

for some $m \times k$ column stochastic matrix $\mathbf{S}$, then the following inequality holds

$$\sum_{j=1}^{k} b_j \psi(z_{p_j}) \leq \sum_{i=1}^{m} a_i \psi(z_{q_i}). \tag{2.123}$$

In particular, if $\psi(t) = t$ for $t \in I = \mathbb{R}$, then (5.97) takes the form

$$\sum_{j=1}^{k} b_j z_{p_j} \leq \sum_{i=1}^{m} a_i z_{q_i}. \tag{2.124}$$

Proof: *We shall proceed with this proof by passing from a convex sequence to a convex function via a method given in [159]. Subsequently we shall apply Theorem 2.2.4.*

Namely, take the sequences $\mathbf{x} = \mathbf{q}$ and $\mathbf{y} = \mathbf{p}$. By virtue of (2.122) we find that Sherman's conditions (2.9) are satisfied.

We now make use of the continuous convex function $\varphi_{\mathbf{z}} : [1, n] \to \mathbb{R}$ generated by the sequence $\mathbf{z}$ via (2.118).

According to Theorem 2.2.4 employed for the convex function $f(t) = \psi(\varphi_{\mathbf{z}}(t))$ for $t \in [1, n]$, we establish the inequality

$$\sum_{j=1}^{k} b_j \psi(z_{p_j}) = \sum_{j=1}^{k} b_j \psi\varphi_{\mathbf{z}}(p_j) \leq \sum_{i=1}^{m} a_i \psi\varphi_{\mathbf{z}}(q_i) = \sum_{i=1}^{m} a_i \psi(z_{q_i}),$$

as was to be proved. $\square$

Remark 2.7.19.: It follows from Theorem *2.2.4* that if $\mathbf{b} = (b_1, b_2, \ldots, b_n) = (1, 1, \ldots, 1)$, and $\mathbf{S}$ is doubly stochastic, then Theorem 2.7.26 reduces to Theorem 2.7.25 in Section 2.7.1. due to Wu and Debnath [159].

Corollary 2.7.10.*([131])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Let $p_j, q_i \in \{1, 2, \ldots, n\}$, $j = 1, 2, \ldots, k$, $i = 1, 2, \ldots, m$. Assume that $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ and $\mathbf{q} = (q_1, q_2, \ldots, q_m)$.*

If

$$\mathbf{p} = \mathbf{q}\,\mathbf{S} \tag{2.125}$$

for some $m \times k$ column stochastic matrix $\mathbf{S}$, then the following inequality holds

$$\sum_{j=1}^{k} \psi(z_{p_j}) \leq \sum_{i=1}^{m} a_i \psi(z_{q_i}), \tag{2.126}$$

where a_i is the sum of all entries of the ith row of $\mathbf{S}$, $i = 1, \ldots, m$.

Proof: *We put*

$$\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m_+ \quad \text{and} \quad \mathbf{b} = (1, 1, \ldots, 1) \in \mathbb{R}^n_+.$$

It follows that

$$(a_1, a_2, \ldots, a_m) = (1, 1, \ldots, 1)\mathbf{S}^T.$$

Therefore (2.125) gives (2.122). It is now enough to apply Theorem 2.7.26. □

Remark 2.7.20.: If the matrix $\mathbf{S}$ is doubly stochastic, then Corollary 2.7.10. becomes Theorem 2.7.25. The reason is that (2.125) with double stochasticity of $\mathbf{S}$ means $\mathbf{p} \prec \mathbf{q}$ and $a_i = 1$, $i = 1, \ldots, n$.

In order to illustrate inequalities (2.123)-(2.126) for sequences, we provide the following example.

Example 7.*(Petrović' inequality, [26, p. 123], [131])*: *Let f :[0,∞)$\to \mathbb{R}$ be a convex* *function. Then*

$$\sum_{i=1}^{k} f(x_i) \le f\left(\sum_{i=1}^{k} x_i\right) + (k-1)f(0) \tag{2.127}$$

for all $x_i \in [0,\infty)$, $i = 1,2,\dots,k$.

This inequality follows from (2.4), because

$$(x_1, x_2, \dots, x_k) \prec \left(\sum_{i=1}^{k} x_i, \underbrace{0,\dots,0}_{k-1 \ \ times}\right).$$

To prove a corresponding result for a convex sequence we shall use Corollary 2.7.10..

Namely, assume that $\psi : I \to \mathbb{R}$ *is a nondecreasing convex function with an interval* $I \subset \mathbb{R}$. *Consider a convex sequence* $\mathbf{z} = (z_0, z_1, \dots, z_n) \in I^{n+1}$. *Let* $p_j \in \{0,1,\dots,n\}$ *for* $j = 1,2,\dots,k$ *with* $\sum\limits_{j=1}^{k} p_j \le n$.

We also put $\mathbf{p} = (p_1, p_2, \dots, p_k)$ *and* $m = 2$, $\mathbf{q} = (q_1, q_2)$ *with* $q_1 = \sum\limits_{j=1}^{k} p_j$, $q_2 = 0$.

Because $p_j \in [q_2, q_1]$, $j = 1,2,\dots,k$, *for* $\lambda_j = p_j / \sum\limits_{j=1}^{k} p_j \in [0,1]$ *we get*

$$p_j = \lambda_j q_1 + (1-\lambda_j) q_2. \tag{2.128}$$

We now introduce the $2 \times k$ *column stochastic matrix*

$$\mathbf{S} = \begin{pmatrix} \lambda_1 & , & \lambda_2 & , & \dots, & \lambda_k \\ 1-\lambda_1 & , & 1-\lambda_2 & , & \dots, & 1-\lambda_k \end{pmatrix}.$$

Hence $\mathbf{p} = \mathbf{q}\,\mathbf{S}$ *by (2.128).*

On account of Corollary 2.7.10. we obtain the inequality

$$\sum_{j=1}^{k} \psi(z_{p_j}) \le a_1 \psi(z_{q_1}) + a_2 \psi(z_{q_2}), \tag{2.129}$$

where a_i *is the sum of all entries of the ith row of* $\mathbf{S}$, $i = 1,2$. *That is*

$$a_1 = \sum_{j=1}^{k} \lambda_j = \sum_{j=1}^{k} \frac{p_j}{\sum\limits_{j=1}^{k} p_j} = 1$$

and

$$a_2 = \sum_{j=1}^{k}(1-\lambda_j) = k - \sum_{j=1}^{k} \frac{p_j}{\sum\limits_{j=1}^{k} p_j} = k-1.$$

In summary, we deduce from (2.129) that

$$\sum_{j=1}^{k} \psi(z_{p_j}) \leq \psi(z_{\sum\limits_{j=1}^{k} p_j}) + (k-1)\psi(z_0).$$

This is the mentioned discrete counterpart for convex sequences of Petrovi´c's inequality (2.127).

We shall now present a more general result. To this end, we utilize Theorem 2.7.26 with $\mathbf{b} = (b_1, b_2, \ldots, b_k) \in \mathbb{R}_+^k$. *In consequence, we get the inequality*

$$\sum_{j=1}^{k} b_j\psi(z_{p_j}) \leq a_1\psi(z_{\sum\limits_{j=1}^{k} p_j}) + a_2\psi(z_0), \tag{2.130}$$

where $\mathbf{a} = (a_1, a_2)$ *is given by* $\mathbf{a} = \mathbf{b}\,\mathbf{S}^T$. *Equivalently,*

$$a_1 = \sum_{j=1}^{k} b_j\lambda_j = \frac{\sum\limits_{j=1}^{k} b_j p_j}{\sum\limits_{j=1}^{k} p_j}$$

and

$$a_2 = \sum_{j=1}^{k} b_j(1-\lambda_j) = \sum_{j=1}^{k} b_j - \frac{\sum\limits_{j=1}^{k} b_j p_j}{\sum\limits_{j=1}^{k} p_j}.$$

Thus we have established a discrete version (2.130) for convex sequences of the Giaccardi inequality (see [139]).

2.7.3. Inequalities of Fejér and Hermite-Hadamard types

In this subsection we discuss some generalizations of Fejér inequality. In doing so, we invoke to Theorem 2.7.26 to derive some inequalities with $m = 2$ terms on their right-hand sides.

Theorem 2.7.27*([131])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Assume that $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n_+$.*

Then the following inequality holds

$$\sum_{j=1}^{n} b_j \psi(z_j) \leq a_1 \psi(z_1) + a_2 \psi(z_n), \tag{2.131}$$

where

$$a_1 = \frac{1}{n-1}[(n-1)b_1 + (n-2)b_2 + \ldots + 1 \cdot b_{n-1} + 0 \cdot b_n], \tag{2.132}$$

$$a_2 = \frac{1}{n-1}[0 \cdot b_1 + 1 \cdot b_2 + \ldots + (n-2)b_{n-1} + (n-1)b_n]. \tag{2.133}$$

In particular, if $\psi(t) = t$ for $t \in I = \mathbb{R}$, then (2.131) reduces to

$$\sum_{j=1}^{n} b_j z_j \leq a_1 z_1 + a_2 z_n, \tag{2.134}$$

where a_1 and a_2 are given by (2.132)-(2.133).

Proof: *Let $m = 2$ and $k = n$. We denote*

$$\mathbf{p} = (p_1, p_2, \ldots, p_n) = (1, 2, \ldots, n) \quad \text{and} \quad \mathbf{q} = (q_1, q_2) = (1, n).$$

In addition, we take $\mathbf{a} = (a_1, a_2)$ with a_1 and a_2 satisfying (2.132)-(2.133).

We shall construct an $m \times k$ column stochastic matrix $\mathbf{S}$ so that $\mathbf{p} = \mathbf{qS}$ is met as in (2.122).

It is obvious that for each $i \in \{1, 2, \ldots, n\}$, i can be viewed as a convex combination of 1 and n, i.e.,

$$i = \alpha_i \cdot 1 + \beta_i \cdot n, \tag{2.135}$$

where

$$\alpha_i = \frac{(n-1)-(i-1)}{n-1} \quad \text{and} \quad \beta_i = \frac{i-1}{n-1}. \tag{2.136}$$

Evidently, $\alpha_i \geq 0$ and $\beta_i \geq 0$ and $\alpha_i + \beta_i = 1$ by $1 \leq i \leq n$.

To show (2.135) with (2.136), notice that

$$\begin{aligned}\alpha_i \cdot 1 + \beta_i \cdot n &= \frac{(n-1)-(i-1)}{n-1} \cdot 1 + \frac{i-1}{n-1} \cdot n \\ &= \frac{(n-1)+(i-1)(n-1)}{n-1} = \frac{(i-1+1)(n-1)}{n-1} = i.\end{aligned}$$

It follows from (2.135)–(2.136) that

for $i = 1$,

$$1 = \frac{n-1}{n-1} \cdot 1 + \frac{0}{n-1} \cdot n,$$

for $i = 2$,

$$2 = \frac{n-2}{n-1} \cdot 1 + \frac{1}{n-1} \cdot n,$$

$\ldots,$

for $i = n-1$,

$$n-1 = \frac{1}{n-1} \cdot 1 + \frac{n-2}{n-1} \cdot n,$$

for $i = n$,

$$n = \frac{0}{n-1} \cdot 1 + \frac{n-1}{n-1} \cdot n.$$

We define the required $2 \times n$ *column stochastic matrix* $\mathbf{S}$ *as follows:*

$$\mathbf{S} = \frac{1}{n-1} \begin{pmatrix} n-1, & n-2, & \ldots, & 1, & 0 \\ 0, & 1, & \ldots, & n-2, & n-1 \end{pmatrix}.$$

It is not hard to check that

$$\mathbf{p} = \mathbf{q}\,\mathbf{S},$$

that is

$$(1, 2, \ldots, n) = (1, n) \cdot \frac{1}{n-1} \begin{pmatrix} n-1, & n-2, & \ldots & 1, & 0 \\ 0, & 1, & \ldots & n-2, & n-1 \end{pmatrix}.$$

Moreover, according to (2.132)-(2.133) we have

$$(a_1, a_2) = (b_1, \ldots, b_n) \cdot \frac{1}{n-1} \begin{pmatrix} n-1, & 0 \\ n-2, & 1 \\ \vdots & \vdots \\ 1, & n-2 \\ 0, & n-1 \end{pmatrix}.$$

Therefore we find that

$$\mathbf{a} = \mathbf{b}\,\mathbf{S}^T.$$

We are now ready to apply Theorem 2.7.26 to $m = 2$, $k = n$, $p_j = j$ *for* $j = 1, \ldots, k$, $q_1 = 1$ *and* $q_1 = n$. *Then we obtain inequality (2.131), as desired.*

Furthermore, by letting $\psi(t) = t$ *for* $t \in I = \mathbb{R}$, *and using (2.131), we get (2.134) completing the proof.* □

From now on, for any positive integer n we use the following notation.

If n is odd, then

$$N = N_1 = N_2 = \frac{n+1}{2} = \left[\frac{n+1}{2}\right].$$

If n is even, then

$$N = N_1 = \frac{n}{2} = \left[\frac{n+1}{2}\right] \quad \text{and} \quad N_2 = \frac{n}{2} + 1.$$

A sequence $\mathbf{b} = (b_1, \ldots, b_n)$ is said to be **symmetric** about $\frac{n+1}{2}$ if

$$b_1 = b_n, \quad b_2 = b_{n-1}, \quad \ldots, \quad b_{N_1 - 1} = b_{N_2 + 1}, \quad b_{N_1} = b_{N_2}.$$

Additionally, we define

$$\epsilon_n = \begin{cases} \frac{1}{2}, & \text{when } n \text{ is odd,} \\ 1, & \text{when } n \text{ is even.} \end{cases}$$

In Theorem 2.7.28 we assume that the sequence $\mathbf{b} = (b_1, \ldots, b_n)$ is symmetric about

$\frac{n+1}{2}$. This allows to simplify the formulas (2.132)-(2.133) to the form

$$a_1 = a_2 = \frac{1}{2}\sum_{i=1}^{n} b_i. \tag{2.137}$$

Theorem 2.7.28*([131])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Assume that $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n_+$ is symmetric about $\frac{n+1}{2}$.*

Then the following Fejér type inequality holds

$$\sum_{j=1}^{n} b_j \psi(z_j) \leq \frac{\psi(z_1) + \psi(z_n)}{2} \sum_{i=1}^{n} b_i. \tag{2.138}$$

Proof: *It follows from Theorem 2.7.27 that*

$$\sum_{j=1}^{n} b_j \psi(z_j) \leq a_1\psi(z_1) + a_2\psi(z_n), \tag{2.139}$$

where a_1 and a_2 are given by (2.132)-(2.133).

By the symmetry of $\mathbf{b} = (b_1, \ldots, b_n)$ about $\frac{n+1}{2}$, we deduce that (2.132)-(2.133) can be rewritten as (2.137).

Indeed, we see that

$$\begin{aligned} a_1 &= \frac{1}{n-1}[(n-1)b_1 + (n-2)b_2 + \ldots + 1 \cdot b_{n-1} + 0 \cdot b_n] \\ &= \frac{1}{n-1}[(n-1)b_1 + \ldots + (n-N_1+1)b_{N_1-1} + (n-N_1)\epsilon_n b_{N_1} \\ &\quad + (n-N_2)\epsilon_n b_{N_2} + (n-N_2-1)b_{N_2+1} + \ldots + 0 \cdot b_n], \end{aligned}$$

because

$$\begin{aligned} &(n-N_1)\epsilon_n b_{N_1} + (n-N_2)\epsilon_n b_{N_2} \\ &= \begin{cases} (n-\frac{n}{2})b_{\frac{n}{2}} + (n-\frac{n}{2}-1)b_{\frac{n}{2}+1} & \text{for even } n \\ (n-\frac{n+1}{2})\frac{1}{2}b_{\frac{n+1}{2}} + (n-\frac{n+1}{2})\frac{1}{2}b_{\frac{n+1}{2}} & \text{for odd } n \end{cases} \\ &= \begin{cases} \frac{n}{2}b_{\frac{n}{2}} + (\frac{n}{2}-1)b_{\frac{n}{2}+1} & \text{for even } n \\ \frac{n-1}{2}b_{\frac{n+1}{2}} & \text{for odd } n. \end{cases} \end{aligned}$$

Therefore,

$$
\begin{aligned}
& a_1 \\
= & \frac{1}{n-1}[(n-1+0)b_1+\ldots+(n-N_1+1+n-N_2-1)b_{N_1-1}+(n-N_1+n-N_2)\epsilon_n b_{N_1}] \\
= & \frac{1}{n-1}\cdot(n-1)\left(b_1+\ldots+b_{N_1-1}+\epsilon_n b_{N_1}\right) \\
= & \frac{1}{2}\left(b_1+b_2+\ldots+b_{n-1}+b_n\right)=\frac{1}{2}\sum_{i=1}^{n} b_i,
\end{aligned}
$$

as wanted.

In a similar way we obtain from the symmetry of $\mathbf{b}=(b_1,\ldots,b_n)$ *about* $\frac{n+1}{2}$ *that*

$$
\begin{aligned}
a_2 &= \frac{1}{n-1}[0\cdot b_1+1\cdot b_2+\ldots+(n-2)\cdot b_{n-1}+(n-1)\cdot b_n] \\
&= \frac{1}{n-1}[0\cdot b_n+1\cdot b_{n-1}+\ldots+(n-2)\cdot b_2+(n-1)\cdot b_1] \\
&= a_1=\frac{1}{2}\sum_{i=1}^{n} b_i.
\end{aligned}
$$

So, by (2.137), we infer that inequality (2.139) gives (2.138). This completes the proof.
□

By making use of Theorem 2.7.28 with the identity function $\psi(t)=t$ for $t\in I=\mathbb{R}$, we get the following right-hand side of Fejér type inequality for convex sequences due to Latreuch and Belaïdi [102, Theorem 1.4] (cf. Theorem 2.7.24):

$$
\sum_{j=1}^{n} b_j z_j \leq \frac{z_1+z_n}{2}\sum_{i=1}^{n} b_i.
$$

A next specification of Theorem 2.7.28 leads to the Hermite–Hadamard type inequality for a convex sequence.

Corollary 2.7.11.*([131])*: *Let* $\psi: I\to\mathbb{R}$ *be a nondecreasing convex function defined on an interval* $I\subset\mathbb{R}$*. Let* $\mathbf{z}=(z_1,z_2,\ldots,z_n)\in I^n$ *be a convex sequence.*

Then the following Hermite-Hadamard type inequality holds

$$\frac{1}{n}\sum_{j=1}^{n}\psi(z_j) \leq \frac{\psi(z_1)+\psi(z_n)}{2}. \tag{2.140}$$

Proof: *We put*

$$\mathbf{b} = (b_1, \ldots, b_n) = \frac{1}{n}(1, \ldots, 1) \in \mathbb{R}_+^n.$$

It is clear that $\mathbf{b}$ *is symmetric about* $\frac{n+1}{2}$.

In light of Theorem 2.7.28 we see that

$$\sum_{j=1}^{n} b_j\psi(z_j) \leq \frac{\psi(z_1)+\psi(z_n)}{2}\sum_{i=1}^{n} b_i. \tag{2.141}$$

We also have

$$\sum_{i=1}^{n} b_i = n \cdot \frac{1}{n} = 1.$$

So, it follows from (2.141) that (2.140) is fulfilled. □

By Corollary 2.7.11. for the identity function $\psi(t) = t$ with $t \in I = \mathbb{R}$ we get the following right-hand side of Hermite-Hadamard type inequality due to Latreuch and Belaïdi [102, Section 4] (cf. Theorem 2.7.24):

$$\frac{1}{n}\sum_{j=1}^{n} z_j \leq \frac{z_1+z_n}{2}.$$

We now focus on some inequalities with $k = 2$ terms on their left-hand sides. To do so, we interpret Theorem 2.7.26 in the following way.

Theorem 2.7.29*([131])*: *Let* $\psi : I \to \mathbb{R}$ *be a nondecreasing convex function defined on an interval* $I \subset \mathbb{R}$. *Let* $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ *be a convex sequence. Let* $b_1, b_2 \geq 0$ *and* $p_1, p_2 \in \{1, 2, \ldots, n\}$, *where*

$$p_1 = 1 \cdot \lambda_1 + 2 \cdot \lambda_2 + \ldots + n \cdot \lambda_n, \tag{2.142}$$

$$p_2 = 1 \cdot \mu_1 + 2 \cdot \mu_2 + \ldots + n \cdot \mu_n \tag{2.143}$$

for some $\lambda_i \geq 0$, $\mu_i \geq 0$, $i = 1, 2, \ldots, n$, $\sum_{i=1}^{n} \lambda_i = 1 = \sum_{i=1}^{n} \mu_i$.

Then the following inequality holds

$$b_1\psi(z_{p_1}) + b_2\psi(z_{p_2}) \leq \sum_{i=1}^{n} a_i\psi(z_i), \tag{2.144}$$

where

$$a_i = b_1 \cdot \lambda_i + b_2 \cdot \mu_i \quad \textit{for} \quad i = 1, 2, \ldots, n. \tag{2.145}$$

Proof: *Under the notation*

$$\mathbf{p} = (p_1, p_2) \quad \textit{and} \quad \mathbf{q} = (q_1, q_2, \ldots, q_n) = (1, 2, \ldots, n),$$

it follows from (2.142)-(2.143) that

$$\mathbf{p} = \mathbf{q}\,\mathbf{S},$$

where $\mathbf{S}$ *is the* $n \times 2$ *column stochastic matrix given by*

$$\mathbf{S} = \begin{pmatrix} \lambda_1, & \mu_1 \\ \lambda_2, & \mu_2 \\ \vdots & \vdots \\ \lambda_n, & \mu_n \end{pmatrix}.$$

Moreover, from (2.145) we also obtain

$$\mathbf{a} = \mathbf{b}\,\mathbf{S}^T \quad \textit{with} \quad \mathbf{a} = (a_1, a_2, \ldots, a_n) \quad \textit{and} \quad \mathbf{b} = (b_1, b_2).$$

In summary, the usage of Theorem 2.7.26 for $m = n$ *and* $k = 2$ *yields inequality (2.144) provided that (2.145) is satisfied.* □

To give a next result, we now study the case of majorization $(p_1, p_2) \prec (1, n)$.

Theorem 2.7.30*([131])*: *Let* $\psi : I \to \mathbb{R}$ *be a nondecreasing convex function defined on an interval* $I \subset \mathbb{R}$. *Let* $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ *be a convex sequence. Let* $b_1, b_2 \geq 0$ *and*

$p_1, p_2 \in \{1, 2, \ldots, n\}$, *where*

$$p_1 = 1 \cdot \lambda_1 + 2 \cdot \lambda_2 + \ldots + n \cdot \lambda_n \tag{2.146}$$

for some $0 \leq \lambda_i \leq \frac{2}{n}$, $i = 1, 2, \ldots, n$, $\sum\limits_{i=1}^{n} \lambda_i = 1$.

If $(p_1, p_2) \prec (1, n)$, *then*

$$b_1 \psi(z_{p_1}) + b_2 \psi(z_{p_2}) \leq \sum_{i=1}^{n} a_i \psi(z_i), \tag{2.147}$$

where

$$a_i = b_1 \cdot \lambda_i + b_2 \cdot \left(\frac{2}{n} - \lambda_i\right) \quad \textit{for} \quad i = 1, 2, \ldots, n. \tag{2.148}$$

Proof: *We denote*

$$\mathbf{p} = (p_1, p_2) \quad \textit{and} \quad \mathbf{q} = (q_1, q_2, \ldots, q_n) = (1, 2, \ldots, n).$$

We have

$$p_1 + p_2 = n + 1, \tag{2.149}$$

because the pair $(1, n)$ *majorizes the pair* (p_1, p_2).

It is easily seen that

$$n + 1 = \frac{2}{n}(1 + 2 + \ldots + n). \tag{2.150}$$

By virtue of (2.146), (2.149) and (2.150), we have

$$p_2 = 1 \cdot \left(\frac{2}{n} - \lambda_1\right) + 2 \cdot \left(\frac{2}{n} - \lambda_2\right) + \ldots + n \cdot \left(\frac{2}{n} - \lambda_n\right). \tag{2.151}$$

In addition, $\frac{2}{n} - \lambda_i \geq 0$ *for* $i = 1, 2, \ldots, n$, *and*

$$\sum_{i=1}^{n} \left(\frac{2}{n} - \lambda_i\right) = n \cdot \frac{2}{n} - \sum_{i=1}^{n} \lambda_i = 2 - 1 = 1.$$

By introducing

$$\mu_i = \frac{2}{n} - \lambda_i \quad \textit{for } i = 1, 2, \ldots, n,$$

we find that (2.151) implies (2.143), and (2.148) implies (2.145).

So, we are permitted to utilize Theorem 2.7.29. Therefore we get (2.144). Further, we also obtain inequality (2.147) as a consequence of condition (2.148). Thus the proof is finished. □

Corollary 2.7.12. ***([131])*:** *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \dots, z_n) \in I^n$ be a convex sequence. Let $p_1, p_2 \in \{1, 2, \dots, n\}$, where*

$$p_1 = 1 \cdot \lambda_1 + 2 \cdot \lambda_2 + \ldots + n \cdot \lambda_n$$

for some $0 \le \lambda_i \le \frac{2}{n}$, $i = 1, 2, \dots, n$, $\sum\limits_{i=1}^{n} \lambda_i = 1$.

Assume that the pair (p_1, p_2) is majorized by $(1, n)$.

Then the following inequality holds

$$\frac{1}{2}\psi(z_{p_1}) + \frac{1}{2}\psi(z_{p_2}) \le \frac{1}{n}\sum_{i=1}^{n}\psi(z_i). \tag{2.152}$$

In particular, if $p_1 = p_2 = \frac{n+1}{2}$ (with odd n) then

$$\psi\left(z_{\frac{n+1}{2}}\right) \le \frac{1}{n}\sum_{i=1}^{n}\psi(z_i). \tag{2.153}$$

If $p_1 = \frac{n}{2}$ and $p_2 = \frac{n}{2} + 1$ (with even n) then

$$\frac{1}{2}\psi\left(z_{\frac{n}{2}}\right) + \frac{1}{2}\psi\left(z_{\frac{n}{2}+1}\right) \le \frac{1}{n}\sum_{i=1}^{n}\psi(z_i). \tag{2.154}$$

Proof: *By letting $b_1 = b_2 = \frac{1}{2}$ we find from Theorem 2.7.30 that*

$$a_i = \frac{1}{2} \cdot \lambda_i + \frac{1}{2} \cdot \left(\frac{2}{n} - \lambda_i\right) = \frac{1}{n} \quad \textit{for} \quad i = 1, 2, \dots, n$$

by (2.148).

Therefore inequality (2.147) can be restated as (2.152), which was to be proved.

Inequalities (2.153)-(2.154) are simple consequences of (2.152). □

Remark 2.7.21.: Let $p_1 = N_1 - c$ and $p_2 = N_2 + c$, where $N_1 = N_2 = \frac{n+1}{2}$ for odd n, and $N_1 = \frac{n}{2}$ and $N_2 = \frac{n}{2} + 1$ for even n.

It is not hard to verify that the majorization $(p_1, p_2) \prec (1, n)$ is met. So, inequality (2.152) can be rewritten in the form

$$\frac{1}{2}\psi(z_{N_1-c}) + \frac{1}{2}\psi(z_{N_2+c}) \leq \frac{1}{n}\sum_{i=1}^{n}\psi(z_i) \tag{2.155}$$

(see [118, p. 56]).

By taking $\psi(t) = t$ with $t \in I = \mathbb{R}$, we infer from inequalities (2.153)-(2.154) that

$$z_{\frac{n+1}{2}} \leq \frac{1}{n}\sum_{i=1}^{n} z_i, \quad \text{when } n \text{ is odd,}$$

$$\frac{1}{2}z_{\frac{n}{2}} + \frac{1}{2}z_{\frac{n}{2}+1} \leq \frac{1}{n}\sum_{i=1}^{n} z_i, \quad \text{when } n \text{ is even,}$$

(see Theorem 2.7.24). This can be though of as the left-hand side of Hermite-Hadamard inequality for a convex sequence $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$ (cf. (2.117)).

2.8. Further inequalities for convex sequences and nondecreasing convex functions

In this section, we derive inequalities of the Hammer-Bullen type for convex sequences and nondecreasing convex functions. We demonstrate an extension of a recent Farissi's inequality. We also provide an inequality of Fejér-Hammer-Bullen type for convex sequences. Finally we give some results related to an Hermite-Hadamard like inequality by Latreuch and Belaïdi. The results of this section are cited from [122].

2.8.1. Introduction and summary

We begin this expository subsection with the Hammer-Bullen inequality [28, 58], which provides a refinement of the Hermite-Hadamard inequality (2.117).

Theorem 2.8.31 ***(Hammer-Bullen inequality, [28, 58])*:** *Assume that $f : I \to \mathbb{R}$ is a convex function on an interval $I = [a, b]$. Then we have*

$$\frac{1}{b-a}\int_a^b f(x)\,dx \le \frac{1}{4}f(a) + \frac{1}{2}f\left(\frac{a+b}{2}\right) + \frac{1}{4}f(b) \le \frac{f(a)+f(b)}{2}. \tag{2.156}$$

In [46] Farissi showed the following refinement of (2.117).

Theorem 2.8.32 ***([46])*:** *Assume that $f : I \to \mathbb{R}$ is a convex function on an interval $I = [a, b]$. Then for all $\mu \in [0, 1]$, we have*

$$f\left(\frac{a+b}{2}\right) \le l(\mu) \le \frac{1}{b-a}\int_a^b f(x)\,dx \le L(\mu) \le \frac{f(a)+f(b)}{2}, \tag{2.157}$$

where

$$l(\mu) := \mu f\left(\frac{\mu b + (2-\mu)a}{2}\right) + (1-\mu) f\left(\frac{(1+\mu)b + (1-\mu)a}{2}\right),$$

$$L(\mu) := \frac{1}{2}[f(\mu b + (1-\mu)a) + \mu f(a) + (1-\mu)f(b)].$$

Some discrete inequalities of the Hermite-Hadamard-Fejér type with *two terms* on their right-hand sides similarly as (2.116)-(2.117) are presented in [131]. In this section we deal with discrete inequalities with *three terms* on their right-hand sides like in Theorems 2.8.31-2.8.32 (see (2.156)-(2.157)).

In Subsection 2.8.2., by using Theorem 2.7.26, we establish Hammer-Bullen type inequalities for convex sequences and nondecreasing convex functions. A main result of this type is Theorem 2.8.33. To prove it, we construct a required column stochastic matrix with the property that the relevant set of the three indices $\{1, \lambda \cdot 1 + (1-\lambda)n, n\}$ on the right-hand side of inequality is sent onto the set $\{1, 2, \ldots, n\}$ on the left-hand side of inequality.

In particular, Theorem 2.8.33 employed for the uniform distribution $b_i = \frac{1}{n}$, $i = 1, 2, \ldots, n$, leads to Corollary 2.8.13.. Its special case yields an extension of Farissi's inequality [46] with convex sequences in place of convex functions (cf. Corollary 2.8.14. and Theorem 2.8.32).

Subsection 2.8.3. is devoted to our problem applied to a symmetric sequence $(b_1, b_2, \ldots, b_n)$. In this context we derive an inequality of Fejér-Hammer-Bullen type (see Theorem 2.8.34). As an application, we provide a discrete version of the classical Hammer-Bullen inequality (cf. Theorem 2.8.31 and Corollary 2.8.15.). Next, for a symmetric probability sequence b_i, we get a discrete Hermite-Hadamard inequality (see Theorem 2.8.35).

In Subsection 2.8.4. we deal with some generalizations of an Hermite-Hadamard like inequality due to Latreuch and Belaïdi [102] (see left-hand side inequality (2.121) and Theorem 2.8.36). The key of our considerations is a construction of a column stochastic matrix suitable to indices labeling elements of a convex sequence $\mathbf{z}$ involved in inequalities.

2.8.2. Hammer-Bullen type inequalities for convex sequences

In this subsection we are going to establish some generalizations of the right-hand side of the Latruech-Belaïdi inequality (see Theorem 2.7.24). Our results concern a nondecreasing convex function ψ as well as coefficients b_i, $i = 1, 2, \ldots, n$, in place of $\frac{1}{n}$. As announced previously, we prefer to study inequalities with $m = 3$ terms on the right-hand side.

The first inequality (2.160) is of Sherman type with parameters $y(i)$, $i = 1, 2, \ldots, n$, running over a corresponding set as in (2.159).

We denote

$$\alpha(i) = \frac{(n-1)-(i-1)}{n-1} \quad \text{and} \quad \beta(i) = \frac{i-1}{n-1}, \qquad i = 1, 2, \ldots, n. \tag{2.158}$$

Theorem 2.8.33*([122])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Assume that $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n_+$. Let $\lambda \in (0, 1)$.*

If for each $i = 1, 2, \ldots, n$,

$$y(i) \in \left[0, \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1-\lambda}\right\}\right], \tag{2.159}$$

then the following inequality holds

$$\sum_{j=1}^{n} b_j\psi(z_j) \le a_1\psi(z_1) + a_2\psi\left(z_{\lambda\cdot 1+(1-\lambda)n}\right) + a_3\psi(z_n), \tag{2.160}$$

where

$$a_1 = \sum_{i=1}^{n} b_i[\alpha(i) - \lambda y(i)], \tag{2.161}$$

$$a_2 = \sum_{i=1}^{n} b_i y(i), \tag{2.162}$$

$$a_3 = \sum_{i=1}^{n} b_i[\beta(i) - \mu y(i)] \tag{2.163}$$

with $\mu = 1 - \lambda$.

In particular, if $\psi(t) = t$, $t \in I = \mathbb{R}$, *then (2.160) reduces to*

$$\sum_{j=1}^{n} b_j z_j \le a_1 z_1 + a_2 z_{\lambda\cdot 1+(1-\lambda)n} + a_3 z_n. \tag{2.164}$$

Proof: *We define*

$$\mathbf{p} = (p_1, p_2, \ldots, p_n) = (1, 2, \ldots, n) \quad \textit{and} \quad \mathbf{q} = (q_1, q_2, q_3) = (1, \lambda\cdot 1 + (1-\lambda)n, n).$$

We also set $\mathbf{a} = (a_1, a_2, a_3)$.

By Theorem 2.7.26 with $m = 3$ *and* $k = n$, *it is enough to construct a* $3 \times n$ *column stochastic matrix* $\mathbf{S}$ *such that* $\mathbf{p} = \mathbf{qS}$ *and* $\mathbf{a} = \mathbf{bS}^T$ *as in (2.122). That is, we need to find matrix* $\mathbf{S}$ *satisfying*

$$(1, 2, \ldots, n) = (1, \lambda\cdot 1 + (1-\lambda)n, n)\cdot \mathbf{S}, \tag{2.165}$$

and

$$(a_1, a_2, a_3) = (b_1, b_2, \ldots, b_n)\cdot \mathbf{S}^T, \tag{2.166}$$

where

$$\mathbf{S} = \begin{pmatrix} x(1), & x(2), & \ldots, & x(n) \\ y(1), & y(2), & \ldots, & y(n) \\ z(1), & z(2), & \ldots, & z(n) \end{pmatrix} \tag{2.167}$$

and $(x,y,z)^T = (x(i),y(i),z(i))^T$ *is the ith column of* $\mathbf{S}$, $i=1,2,\ldots,n$.

To do so, we shall build and solve a linear system with unknowns $x(i),y(i),z(i)$ *for* $i=1,2,\ldots,n$, *fulfilling some adequate restrictions.*

Namely, it follows from $\mathbf{p}=\mathbf{qS}$ *that*

$$i = p_i = x(i)\cdot 1 + y(i)[\lambda\cdot 1 + (1-\lambda)n] + z(i)\cdot n, \tag{2.168}$$

which gives

$$i = (x+\lambda y)\cdot 1 + [(1-\lambda)y + z]n,$$

and further

$$i = \alpha\cdot 1 + \beta\cdot n \quad \text{for each } i\in\{1,2,\ldots,n\}, \tag{2.169}$$

with $\alpha=\alpha(i)$ *and* $\beta=\beta(i)$ *given by (2.158). Obviously,* $\alpha\geq 0$, $\beta\geq 0$ *and* $\alpha+\beta=1$ *by* $1\leq i\leq n$.

In consequence, we get the linear system

$$\begin{cases} x & + & \lambda y & & & = \alpha \\ & & (1-\lambda)y & + & z & = \beta \\ x & + & y & + & z & = 1 \end{cases} \quad \text{subject to} \quad \begin{cases} 0\leq x\leq 1 \\ 0\leq y\leq 1 \\ 0\leq z\leq 1 \end{cases}. \tag{2.170}$$

By the column stochasticity of $\mathbf{S}$, *it is solved by*

$$\begin{cases} x & = \alpha-\lambda y \\ z & = \beta-(1-\lambda)y \\ & \end{cases} \quad \text{subject to} \quad \begin{cases} \frac{\alpha-1}{\lambda}\leq y\leq \frac{\alpha}{\lambda} \\ 0\leq y\leq 1 \\ \frac{\beta-1}{1-\lambda}\leq y\leq \frac{\beta}{1-\lambda} \end{cases}, \tag{2.171}$$

where the parameter y *runs over the set (2.172).*

To prove that the system (2.171) is consistent, we need to show that

$$[0,1]\cap\left[\frac{\alpha(i)-1}{\lambda},\frac{\alpha(i)}{\lambda}\right]\cap\left[\frac{\beta(i)-1}{1-\lambda},\frac{\beta(i)}{1-\lambda}\right]\neq\emptyset \tag{2.172}$$

for each $i=1,2,\ldots,n$.

In fact, by (2.158), for $i = 1, 2, \ldots, n$, we have

$$\alpha(i) - 1 \le 0 \le \alpha(i) \le 1$$

and

$$\beta(i) - 1 \le 0 \le \beta(i) \le 1.$$

Therefore

$$\frac{\alpha(i) - 1}{\lambda} \le 0 \le \frac{\alpha(i)}{\lambda} \quad \textit{for } i = 1, 2, \ldots, n, \tag{2.173}$$

$$\frac{\beta(i) - 1}{1 - \lambda} \le 0 \le \frac{\beta(i)}{1 - \lambda} \quad \textit{for } i = 1, 2, \ldots, n. \tag{2.174}$$

Hence

$$0 \in [0, 1] \cap \left[\frac{\alpha(i) - 1}{\lambda}, \frac{\alpha(i)}{\lambda}\right] \cap \left[\frac{\beta(i) - 1}{1 - \lambda}, \frac{\beta(i)}{1 - \lambda}\right] \ne \emptyset,$$

which completes the proof of (2.172).

Furthermore, it follows that

$$\frac{\alpha(i)}{\lambda} \le 1, \quad \textit{whenever} \quad \frac{\alpha(i)}{\lambda} \le \frac{\beta(i)}{1 - \lambda},$$

and

$$\frac{\beta(i)}{1 - \lambda} \le 1, \quad \textit{whenever} \quad \frac{\alpha(i)}{\lambda} \ge \frac{\beta(i)}{1 - \lambda}.$$

To prove the former, suppose not. Then $1 < \frac{\alpha(i)}{\lambda}$. Since $\frac{\alpha(i)}{\lambda} \le \frac{\beta(i)}{1-\lambda}$, we get $1 < \frac{\beta(i)}{1-\lambda}$. In consequence, $\lambda < \alpha(i)$ and $1 - \lambda < \beta(i)$. For this reason, $1 = \lambda + (1 - \lambda) < \alpha(i) + \beta(i) = 1$, which is a contradiction.

To see the latter, use a similar manner of the proof.

Summarizing all of this, we obtain

$$0 \le \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1 - \lambda}\right\} \le 1.$$

On the other hand, (2.173) and (2.174) ensure that

$$\left[0, \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1 - \lambda}\right\}\right] = [0, 1] \cap \left[\frac{\alpha(i) - 1}{\lambda}, \frac{\alpha(i)}{\lambda}\right] \cap \left[\frac{\beta(i) - 1}{1 - \lambda}, \frac{\beta(i)}{1 - \lambda}\right]. \tag{2.175}$$

The statement (2.172) guarantees the existence of $y = y(i)$ *for* $i = 1, 2, \ldots, n$ *satisfying (2.171). This implies the existence of matrix* $\mathbf{S}$ *determined by (2.167) with properties (2.165)-(2.166).*

In light of (2.167), (2.171) and (2.175), it holds that

$$\mathbf{S} = \begin{pmatrix} \alpha(1) - \lambda y(1), & \alpha(2) - \lambda y(2), & \ldots, & \alpha(n) - \lambda y(n) \\ y(1), & y(2), & \ldots, & y(n) \\ \beta(1) - (1-\lambda)y(1), & \beta(2) - (1-\lambda)y(2), & \ldots, & \beta(n) - (1-\lambda)y(n) \end{pmatrix}, \tag{2.176}$$

where

$$y(i) \in \left[0, \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1-\lambda}\right\}\right], \qquad i = 1, 2, \ldots, n.$$

In addition, $\alpha(i) + \beta(i) = 1$ *for* $i = 1, 2, \ldots, n$*, so* $\mathbf{S}$ *is column-stochastic, as required.*

It follows from equalities (2.161)-(2.163) and (2.176) that the condition (2.166) is satisfied. Moreover, (2.168) is also met by (2.169) and (2.170), so (2.165) holds true.

On account of Theorem 2.7.26, inequality (2.160) is valid, as claimed.

Finally, (2.164) follows from (2.160). □

Corollary 2.8.13. *([122])*: *Let* $\psi : I \to \mathbb{R}$ *be a nondecreasing convex function defined on an interval* $I \subset \mathbb{R}$*. Let* $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ *be a convex sequence. Let* $\lambda \in (0, 1)$*.*

If for each $i = 1, 2, \ldots, n$*,*

$$y(i) \in \left[0, \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1-\lambda}\right\}\right],$$

then the following inequality holds

$$\frac{1}{n}\sum_{j=1}^{n} \psi(z_j) \leq a_1\psi(z_1) + a_2\psi\left(z_{\lambda\cdot 1 + (1-\lambda)n}\right) + a_3\psi(z_n), \tag{2.177}$$

where

$$a_1 = \frac{1}{2} - \lambda a_2, \qquad a_2 = \frac{1}{n}\sum_{i=1}^{n} y(i), \quad \textit{and} \quad a_3 = \frac{1}{2} - (1-\lambda)a_2.$$

Proof. We set

$$\mathbf{b} = (b_1, b_2, \ldots, b_n) = \frac{1}{n}(1, 1, \ldots, 1).$$

By virtue of Theorem 2.8.33, inequality (2.177) holds with

$$a_1 = \frac{1}{n}\left(\sum_{i=1}^{n}\alpha(i) - \lambda\sum_{i=1}^{n}y(i)\right), \tag{2.178}$$

$$a_2 = \frac{1}{n}\sum_{i=1}^{n}y(i), \tag{2.179}$$

$$a_3 = \frac{1}{n}\left(\sum_{i=1}^{n}\beta(i) - (1-\lambda)\sum_{i=1}^{n}y(i)\right). \tag{2.180}$$

Simultaneously, by (2.158) we get

$$\sum_{i=1}^{n}\beta(i) = \sum_{i=1}^{n}\frac{i-1}{n-1} = \frac{n}{2}.$$

Likewise,

$$\sum_{i=1}^{n}\alpha(i) = \sum_{i=1}^{n}\left(1 - \frac{i-1}{n-1}\right) = \frac{n}{2}.$$

Thus conditions (2.178)-(2.180) become

$$\begin{aligned} a_1 &= \frac{1}{n}\left(\frac{n}{2} - \lambda\sum_{i=1}^{n}y(i)\right) = \frac{1}{2} - \lambda a_2, \\ a_2 &= \frac{1}{n}\sum_{i=1}^{n}y(i), \\ a_3 &= \frac{1}{n}\left(\frac{n}{2} - (1-\lambda)\sum_{i=1}^{n}y(i)\right) = \frac{1}{2} - (1-\lambda)a_2. \end{aligned}$$

So, the proof of Corollary 2.8.13. is finished. □

As a corollary to the last result we now demonstrate a discrete version (of right-hand side) of Farissi's inequality (see Theorem 2.8.32).

Corollary 2.8.14.*([122])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Let $\lambda \in (0,1)$.*

Then the following inequality holds

$$\frac{1}{n}\sum_{j=1}^{n}\psi(z_j) \le a_1\psi(z_1) + a_2\psi\left(z_{\lambda\cdot 1+(1-\lambda)n}\right) + a_3\psi(z_n), \tag{2.181}$$

with

$$\begin{aligned} a_1 &= \frac{1}{2} - \lambda a_2, && (2.182)\\ a_2 &= \frac{([M]-1)[M]}{2\mu(n-1)n} + \frac{(n-[M]-1)(n-[M])}{2\lambda(n-1)n}, && (2.183)\\ a_3 &= \frac{1}{2} - \mu a_2, && (2.184) \end{aligned}$$

where $\mu = 1-\lambda$, $M = \mu(n-1)+1$ and $[M]$ stands for the entire part of M.

Proof: *For each $i = 1, 2, \ldots, n$, we take*

$$y(i) = \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1-\lambda}\right\}, \tag{2.185}$$

and

$$\mathbf{b} = (b_1, b_2, \ldots, b_n) = \frac{1}{n}(1, 1, \ldots, 1).$$

We deduce from Corollary 2.8.13. that (2.181) holds with

$$\begin{aligned} a_1 &= \frac{1}{2} - \lambda a_2, && (2.186)\\ a_2 &= \frac{1}{n}\sum_{i=1}^{n} y(i), && (2.187)\\ a_3 &= \frac{1}{2} - (1-\lambda)a_2. && (2.188) \end{aligned}$$

We need to show that under the help of (2.185) condition (2.187) reduces to (2.182).

It is not hard to verify that

$$\frac{\alpha(i)}{\lambda} < \frac{\beta(i)}{\mu} \quad \text{iff} \quad \alpha(i) < \lambda \quad \text{iff} \quad \beta(i) > \mu \quad \text{iff} \quad i > \mu(n-1)+1$$

and

$$\frac{\alpha(i)}{\lambda} \geq \frac{\beta(i)}{\mu} \quad \text{iff} \quad \alpha(i) \geq \lambda \quad \text{iff} \quad \beta(i) \leq \mu \quad \text{iff} \quad i \leq \mu(n-1)+1.$$

So, we have

$$y(i) = \begin{cases} \frac{\alpha(i)}{\lambda} & \text{for } \ i > \mu(n-1)+1 = M,\\ \frac{\beta(i)}{\mu} & \text{for } \ i \leq \mu(n-1)+1 = M. \end{cases}$$

For this reason a_2 in (2.187) can be restated as follows:

$$\begin{aligned} a_2 &= \frac{1}{n}\sum_{i=1}^{n} y(i) = \frac{1}{n}\sum_{1\le i\le M}\frac{\beta(i)}{\mu} + \frac{1}{n}\sum_{M<i\le n}\frac{\alpha(i)}{\lambda} \\ &= \frac{1}{\mu n}\sum_{1\le i\le [M]}\beta(i) + \frac{1}{\lambda n}\sum_{[M]+1\le i\le n}\alpha(i). \end{aligned}$$

Now, by employing (2.158) we find that

$$\begin{aligned} a_2 &= \frac{1}{\mu n}\sum_{1\le i\le [M]}\frac{i-1}{n-1} + \frac{1}{\lambda n}\sum_{[M]+1\le i\le n}\frac{(n-1)-(i-1)}{n-1} \\ &= \frac{1}{\mu n(n-1)}\frac{([M]-1)[M]}{2} + \frac{1}{\lambda n(n-1)}\frac{(n-[M]-1)(n-[M])}{2}. \end{aligned}$$

Thus (2.182) follows, which completes the proof of Corollary 2.8.14.. □

Remark 2.8.22.: In the previous corollary, we have $M = \mu(n-1)+1$ and $n - M = \lambda(n-1)$. Moreover,

$$0 \le \frac{M-1}{\mu(n-1)} - \frac{[M]-1}{\mu(n-1)} \le \frac{1}{\mu(n-1)} \to 0 \quad \text{when } n\to\infty.$$

Consequently,

$$\frac{[M]-1}{\mu(n-1)} \to 1 \quad \text{when } n\to\infty,$$

and

$$0 \le \frac{n-[M]}{\lambda(n-1)} - \frac{n-M}{\lambda(n-1)} \le \frac{1}{\lambda(n-1)} \to 0 \quad \text{when } n\to\infty.$$

Hence

$$\frac{n-[M]}{\lambda(n-1)} \to 1 \quad \text{when } n\to\infty.$$

Therefore for large n we get (see (2.183))

$$a_2 \approx \frac{[M]}{2n} + \frac{(n-[M]-1)}{2n} = \frac{1}{2}\frac{n-1}{n} \approx \frac{1}{2}.$$

So, by (2.182) and (2.184),

$$a_1 \approx \frac{1}{2} - \lambda\frac{1}{2} = \frac{1}{2}\mu$$

and

$$a_3 \approx \frac{1}{2} - \mu\frac{1}{2} = \frac{1}{2}\lambda.$$

This is approximately in accordance with continuous case (see Theorem 2.8.32).

2.8.3. Fejèr-Hammer-Bullen inequalities with symmetric sequence b

Throughout this subsection, if n is odd, then we set

$$N = N_1 = N_2 = \frac{n+1}{2} = \left[\frac{n+1}{2}\right],$$

and, if n is even, then we set

$$N = N_1 = \frac{n}{2} = \left[\frac{n+1}{2}\right] \quad \text{and} \quad N_2 = \frac{n}{2} + 1,$$

where the symbol $\left[\frac{n+1}{2}\right]$ stands for the entire part of $\frac{n+1}{2}$.

A sequence $\mathbf{b} = (b_1, \ldots, b_n)$ is said to be *symmetric about* $\frac{n+1}{2}$ if

$$b_1 = b_n, \quad b_2 = b_{n-1}, \quad \ldots, \quad b_{N_1-1} = b_{N_2+1}, \quad b_{N_1} = b_{N_2}.$$

In order to abbreviate some notation, we also introduce

$$\epsilon_n = \begin{cases} \frac{1}{2}, & \text{when } n \text{ is odd,} \\ 1, & \text{when } n \text{ is even.} \end{cases}$$

For instance, the following identity is of interest

$$\begin{aligned} & b_1c_1 + b_2c_2 + \ldots + b_nc_n \\ = \ & b_1(c_1 + c_n) + b_2(c_2 + c_{n-1}) + \ldots \\ & + b_{N_1-1}(c_{N_1-1} + c_{N_2+1}) + \epsilon_n b_{N_1}(c_{N_1} + c_{N_2}) \end{aligned} \tag{2.189}$$

for any real symmetric sequence $b_1, b_2, \ldots, b_n$ and real sequence $c_1, c_2, \ldots, c_n$.

If in addition

$$c_1 + c_n = 2K, \quad c_2 + c_{n-1} = 2K, \quad \ldots, \quad c_{N_1-1} + c_{N_2+1} = 2K, \quad c_{N_1} + c_{N_2} = 2K,$$

then the identity (2.189) can be rewritten in the form:

$$\begin{aligned} b_1c_1 + b_2c_2 + \ldots + b_nc_n &= 2K(b_1 + b_2 + \ldots + b_{N_1-1} + \epsilon_n b_{N_1}) \\ &= K(b_1 + b_2 + \ldots + b_n). \end{aligned}$$

We also define

$$\widetilde{b_i} = \begin{cases} b_i & \text{for } i = 1, 2, \ldots, N_1 - 1, \\ \epsilon_n b_{N_1} & \text{for } i = N_1. \end{cases}$$

Theorem 2.8.34*([122])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Assume that $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$ is symmetric about $\frac{n+1}{2}$.*

Then the following inequality holds

$$\sum_{j=1}^{n} b_j \psi(z_j) \le a_1 \psi(z_1) + a_2 \psi\left(z_{\frac{n+1}{2}}\right) + a_3 \psi(z_n), \tag{2.190}$$

where

$$a_1 = a_3 = \sum_{i=1}^{N_1} \widetilde{b_i} \left(1 - 2\frac{i-1}{n-1}\right) = \begin{cases} \sum\limits_{i=1}^{n/2} b_i \left(1 - 2\frac{i-1}{n-1}\right) & \textit{for even } n, \\ \sum\limits_{i=1}^{(n-1)/2} b_i \left(1 - 2\frac{i-1}{n-1}\right) & \textit{for odd } n, \end{cases} \tag{2.191}$$

$$a_2 = 4\sum_{i=1}^{N_1} \widetilde{b_i} \frac{i-1}{n-1} = \begin{cases} 4\sum\limits_{i=1}^{n/2} b_i \frac{i-1}{n-1} & \textit{for even } n, \\ 4\sum\limits_{i=1}^{(n-1)/2} b_i \frac{i-1}{n-1} + b_{\frac{n+1}{2}} & \textit{for odd } n. \end{cases} \tag{2.192}$$

Proof: *Let $\lambda = \frac{1}{2}$ and $\mu = 1 - \lambda = \frac{1}{2}$. For each $i = 1, 2, \ldots, n$, we take*

$$y(i) = \min\left\{\frac{\alpha(i)}{\lambda}, \frac{\beta(i)}{1-\lambda}\right\} = \begin{cases} 2\beta(i), & i = 1, 2 \ldots, N_1, \\ 2\alpha(i), & i = N_2, \ldots, n. \end{cases} \tag{2.193}$$

In light of Theorem 2.8.33 one sees that inequality (2.190) holds whenever conditions (2.161)-(2.163) are satisfied. Therefore, by using (2.193), we shall show that conditions (2.161)-(2.163) become (2.191)-(2.192).

Namely, it follows from (2.161) that

$$\begin{aligned} a_1 &= \sum_{i=1}^{n} b_i \left(\alpha(i) - \frac{1}{2}y(i)\right) \\ &= \sum_{i=1}^{N_1-1} b_i\alpha(i) + \epsilon_n b_{N_1}\alpha(N_1) + \epsilon_n b_{N_2}\alpha(N_2) + \sum_{i=N_2+1}^{n} b_i\alpha(i) \\ &\quad -\frac{1}{2}\left(\sum_{i=1}^{N_1-1} b_i y(i) + \epsilon_n b_{N_1} y(N_1) + \epsilon_n b_{N_2} y(N_2) + \sum_{i=N_2+1}^{n} b_i y(i)\right) \\ &= \sum_{i=1}^{N_1} \widetilde{b_i}\left(\alpha(i) + \alpha(n+1-i)\right) - \frac{1}{2}\sum_{i=1}^{N_1}\widetilde{b_i}\left(y(i) + y(n+1-i)\right) \\ &= \sum_{i=1}^{N_1} \widetilde{b_i}\left(\alpha(i) - \beta(i)\right) = \sum_{i=1}^{N_1}\widetilde{b_i}\left(1 - 2\frac{i-1}{n-1}\right). \end{aligned}$$

So, for even n we have

$$a_1 = \sum_{i=1}^{\frac{n}{2}} b_i \left(1 - 2\frac{i-1}{n-1}\right)$$

and, for odd n,

$$a_1 = \sum_{i=1}^{N_1-1}\widetilde{b_i}\left(1 - 2\frac{i-1}{n-1}\right) + \widetilde{b}_{N_1}\left(1 - 2\frac{N_1-1}{n-1}\right) = \sum_{i=1}^{\frac{n-1}{2}} b_i\left(1 - 2\frac{i-1}{n-1}\right),$$

completing the proof of (2.191) for a_1.

Similarly, condition (2.162) in Theorem 2.8.33 gives

$$\begin{aligned} a_2 &= \sum_{i=1}^{n} b_i y(i) = \sum_{i=1}^{N_1-1} b_i y(i) + \epsilon b_{N_1} y(N_1) + \epsilon b_{N_2} y(N_2) + \sum_{i=N_2+1}^{n} b_i y(i) \\ &= \sum_{i=1}^{N_1-1} b_i\left(y(i) + y(n+1-i)\right) + \epsilon b_{N_1}\left(y(N_1) + y(N_2)\right) \\ &= 2\sum_{i=1}^{N_1}\widetilde{b_i}\left(\beta(i) + \alpha(n+1-i)\right) = 4\sum_{i=1}^{N_1}\widetilde{b_i}\frac{i-1}{n-1}. \end{aligned}$$

For example, for even n,

$$a_2 = 4\sum_{i=1}^{\frac{n}{2}} b_i \frac{i-1}{n-1},$$

and for odd n,

$$a_2 = 4\sum_{i=1}^{N_1-1} b_i \frac{i-1}{n-1} + 2b_{N_1}\frac{N_1-1}{n-1} = 4\sum_{i=1}^{\frac{n-1}{2}} b_i \frac{i-1}{n-1} + b_{\frac{n+1}{2}},$$

which proves (2.192).

Thanks to condition (2.163) in Theorem 2.8.33 we infer that

$$\begin{aligned}
a_3 &= \sum_{i=1}^{n} b_i \left(\beta(i) - \frac{1}{2}y(i)\right) \\
&= \sum_{i=1}^{N_1-1} b_i\beta(i) + \epsilon_n b_{N_1}\beta(N_1) + \epsilon_n b_{N_2}\beta(N_2) + \sum_{i=N_2+1}^{n} b_i\beta(i) \\
&\quad -\frac{1}{2}\left(\sum_{i=1}^{N_1-1} b_i y(i) + \epsilon_n b_{N_1} y(N_1) + \epsilon_n b_{N_2} y(N_2) + \sum_{i=N_2+1}^{n} b_i y(i)\right) \\
&= \sum_{i=1}^{N_1} \widetilde{b_i}\left(\beta(i) + \beta(n+1-i)\right) - \frac{1}{2}\sum_{i=1}^{N_1} \widetilde{b_i}\left(y(i) + y(n+1-i)\right) \\
&= \sum_{i=1}^{N_1} \widetilde{b_i}\left(\alpha(i) - \beta(i)\right) = \sum_{i=1}^{N_1} \widetilde{b_i}\left(1 - 2\frac{i-1}{n-1}\right) = a_1.
\end{aligned}$$

Thus the proof is completed. □

Theorem 2.8.35 *([122])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in I^n$ be a convex sequence. Assume that $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n$ is a probability sequence symmetric about $\frac{n+1}{2}$.*

Then the following inequality holds

$$\sum_{j=1}^{n} b_j\psi(z_j) \le a_1\psi(z_1) + a_2\psi\left(z_{\frac{n+1}{2}}\right) + a_3\psi(z_n) \le \frac{\psi(z_1)+\psi(z_n)}{2}, \tag{2.194}$$

where $a_1 = a_3$ and a_2 are given by (2.191)-(2.192).

Proof: *By Theorem 2.8.34 we need to show that*

$$a_1\psi(z_1) + a_2\psi\left(z_{\frac{n+1}{2}}\right) + a_3\psi(z_n) \leq \frac{\psi(z_1)+\psi(z_n)}{2}. \tag{2.195}$$

It is easily seen that

$$\psi\left(z_{\frac{n+1}{2}}\right) \leq \frac{\psi(z_1)+\psi(z_n)}{2}. \tag{2.196}$$

In fact, we have $\left(\frac{n+1}{2}, \frac{n+1}{2}\right) \prec (1, n)$*, where* $\prec$ *denotes the majorization preorder on* $\mathbb{R}^2$*. So, from Majorization Theorem we obtain*

$$\psi\varphi_{\mathbf{z}}\left(\frac{n+1}{2}\right) + \psi\varphi_{\mathbf{z}}\left(\frac{n+1}{2}\right) \leq \psi\varphi_{\mathbf{z}}(1) + \psi\varphi_{\mathbf{z}}(n),$$

where $\varphi_{\mathbf{z}}$ *is given by (2.118). Equivalently,*

$$\psi(z_{\frac{n+1}{2}}) + \psi(z_{\frac{n+1}{2}}) \leq \psi(z_1) + \psi(z_n),$$

which forces (2.196).

In consequence, the vector

$$\mathbf{d} = \left(\frac{\psi(z_1)+\psi(z_n)}{2}, \psi\left(z_{\frac{n+1}{2}}\right)\right) \in \mathbb{R}^2$$

is nonincreasing. Hence the rearrangement inequality ensures that

$$\langle \mathbf{d}, \mathbf{x}\rangle \leq \langle \mathbf{d}, \mathbf{y}\rangle, \tag{2.197}$$

whenever $\mathbf{x} = (x_1, x_2)$*,* $\mathbf{y} = (y_1, y_2)$ *are two vectors in* $\mathbb{R}^2$ *such that* $\mathbf{x} \prec \mathbf{y}$ *and* $y_1 \geq y_2$*. Here the symbol* $\langle\cdot,\cdot\rangle$ *denotes the standard inner product on* $\mathbb{R}^2$*.*

It remains to prove that

$$\left(\frac{1}{2}, \frac{1}{2}\right) \prec (2a_1, a_2) \prec (1, 0). \tag{2.198}$$

According to (2.191)-(2.192) and (2.189) (with $c_i = 1$, $i = 1, 2, \dots, n$), we deduce that

$$2a_1 + a_2 = 2\sum_{i=1}^{N_1} \widetilde{b_i}\left(1 - 2\frac{i-1}{n-1}\right) + 4\sum_{i=1}^{N_1} \widetilde{b_i}\frac{i-1}{n-1} = 2\sum_{i=1}^{N_1} \widetilde{b_i} = \sum_{i=1}^{n} b_i = 1. \tag{2.199}$$

Additionally, $2a_1, a_2 \in [0,1]$. Indeed, by $N_1 \le \frac{n+1}{2}$, we get $2(N_1 - 1) \le n - 1$. On the other hand, we have $2(i-1) \le 2(N_1 - 1)$ for $i = 1, 2, \dots, N_1$. By combining the last two inequalities we derive $2(i-1) \le n-1$ for $i = 1, 2, \dots, N_1$. Thus we obtain $1 - 2\frac{i-1}{n-1} \ge 0$ for $i = 1, 2, \dots, N_1$, which implies $\sum\limits_{i=1}^{N_1} \widetilde{b_i}\left(1 - 2\frac{i-1}{n-1}\right) \ge 0$, since $\widetilde{b_i} \ge 0$. Now, on account of (2.191) we establish that $a_1 \ge 0$.

Furthermore, we have $a_2 \ge 0$ by (2.192) with $\widetilde{b_i} \ge 0$. Simultaneously, $2a_1 + a_2 = 1$, so $2a_1 \le 2a_1 + a_2 = 1$ and $a_2 \le 2a_1 + a_2 = 1$. Hence $2a_1 \le 1$ and $a_2 \le 1$. Therefore $2a_1, a_2 \in [0,1]$ with $2a_1 + a_2 = 1$ and $\frac{1}{2} \le \max\{2a_1, a_2\} \le 1$. Thus we get (2.198), as claimed.

Finally, it is a consequence of (2.197) and (2.198) that

$$2a_1\frac{\psi(z_1) + \psi(z_n)}{2} + a_2\psi\left(z_{\frac{n+1}{2}}\right) \le 1 \cdot \frac{\psi(z_1) + \psi(z_n)}{2} + 0 \cdot \psi\left(z_{\frac{n+1}{2}}\right),$$

which implies (2.195), as wanted. □

Remark 2.8.23.: By (2.197)-(2.198), the last inequality in the above proof can be updated in the following manner provided that $2a_1 \ge a_2$ (e.g., see Corollary 2.8.15.):

$$\frac{1}{4}\psi(z_1) + \frac{1}{2}\psi\left(z_{\frac{n+1}{2}}\right) + \frac{1}{4}\psi(z_n) \le a_1\psi(z_1) + a_2\psi\left(z_{\frac{n+1}{2}}\right) + a_3\psi(z_n).$$

From Theorems 2.8.34-2.8.35 we obtain a discrete version of Hammer-Bullen inequality (cf. Theorem 2.8.31), as follows.

Corollary 2.8.15.*([122])*: *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \dots, z_n) \in I^n$ be a convex sequence.*

Then the following inequality holds

$$\frac{1}{n}\sum_{j=1}^{n} \psi(z_j) \le a_1\psi(z_1) + a_2\psi\left(z_{\frac{n+1}{2}}\right) + a_3\psi(z_n) \le \frac{\psi(z_1) + \psi(z_n)}{2}, \tag{2.200}$$

where

$$a_1 = a_3 = \begin{cases} \frac{1}{4} + \frac{1}{4(n-1)} & \text{if } n \text{ is even,} \\ \frac{1}{4} + \frac{1}{4n} & \text{if } n \text{ is odd,} \end{cases} \tag{2.201}$$

$$a_2 = \begin{cases} \frac{n-2}{2(n-1)} & \text{if } n \text{ is even,} \\ \frac{n-1}{2n} & \text{if } n \text{ is odd.} \end{cases} \tag{2.202}$$

Proof. With the setting $\mathbf{b} = \frac{1}{n}(1, \ldots, n)$ in Theorem 2.8.35, we reduce the statements (2.191)-(2.192) in the following way.

For even n,

$$a_1 = a_3 = \sum_{i=1}^{\frac{n}{2}} b_i \left(1 - 2\frac{i-1}{n-1}\right) = \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} \left(1 - 2\frac{i-1}{n-1}\right) = \frac{1}{4} + \frac{1}{4(n-1)}.$$

For odd n,

$$a_1 = a_3 = \sum_{i=1}^{\frac{n-1}{2}} b_i \left(1 - 2\frac{i-1}{n-1}\right) = \frac{1}{n}\sum_{i=1}^{\frac{n-1}{2}} \left(1 - 2\frac{i-1}{n-1}\right) = \frac{1}{4} + \frac{1}{4n}.$$

By utilizing (2.192) with even n, we get

$$a_2 = 4\sum_{i=1}^{\frac{n}{2}} b_i \frac{i-1}{n-1} = 4\frac{1}{n}\sum_{i=1}^{\frac{n}{2}} \frac{i-1}{n-1} = \frac{n-2}{2(n-1)}.$$

According to (2.192) with odd n, we have

$$a_2 = 4\sum_{i=1}^{\frac{n-1}{2}} b_i \frac{i-1}{n-1} + b_{\frac{n+1}{2}} = 4\frac{1}{n}\sum_{i=1}^{\frac{n-1}{2}} \frac{i-1}{n-1} + \frac{1}{n} = \frac{n-1}{2n}.$$

This finishes the proof of Corollary 2.8.15.. □

2.8.4. Further inequalities for convex sequences

This subsection is devoted to deriving some results related to the left-hand side inequality (2.121). We emphasize the usage of some indices other than $\frac{n+1}{2}$.

Theorem 2.8.36 ***([122])*:** *Let $\psi : I \to \mathbb{R}$ be a nondecreasing convex function defined on* *an interval $I \subset \mathbb{R}$. Let $\mathbf{z} = (z_1, z_2, \dots, z_n) \in I^n$ be a convex sequence.*

Then for arbitrary nonnegative b_1, b_2, b_3,

$$b_1\psi(z_{\frac{n+1}{2}}) + b_2\psi(z_{\frac{2n+1}{3}}) + b_3\psi(z_{\frac{3n(n+1)}{4n+2}}) \leq \sum_{i=1}^{n} a_i\psi(z_i), \tag{2.203}$$

where

$$a_i = \frac{1}{n}b_1 + \frac{2\cdot i}{n(n+1)}b_2 + \frac{6\cdot i^2}{n(n+1)(2n+1)}b_3 \quad \text{for } i = 1, 2, \dots, n. \tag{2.204}$$

Proof. Evidently,

$$\frac{1}{n}(1 + 2 + \ldots + n) = \frac{n+1}{2}, \tag{2.205}$$

$$\frac{2}{n(n+1)}(1^2 + 2^2 + \ldots + n^2) = \frac{2n+1}{3}, \tag{2.206}$$

$$\frac{6}{n(n+1)(2n+1)}(1^3 + 2^3 + \ldots + n^3) = \frac{3n(n+1)}{4n+2}. \tag{2.207}$$

We set

$$\mathbf{p} = (p_1, p_2, p_3) = \left(\frac{n+1}{2}, \frac{2n+1}{3}, \frac{3n(n+1)}{4n+2}\right)$$

and

$$\mathbf{q} = (q_1, q_2, \dots, q_n) = (1, 2, \dots, n).$$

By virtue of (2.205)-(2.207) we have

$$p_1 = \frac{1}{n}q_1 + \frac{1}{n}q_2 + \ldots + \frac{1}{n}q_n. \tag{2.208}$$

$$p_2 = \frac{2\cdot 1}{n(n+1)}q_1 + \frac{2\cdot 2}{n(n+1)}q_2 + \ldots + \frac{2\cdot n}{n(n+1)}q_n. \tag{2.209}$$

$$p_3 = \frac{6\cdot 1^2}{n(n+1)(2n+1)}q_1 + \frac{6\cdot 2^2}{n(n+1)(2n+1)}q_2 + \ldots + \frac{6\cdot n^2}{n(n+1)(2n+1)}q_n. \tag{2.210}$$

We now define $n \times 3$ matrix $\mathbf{S}$ by

$$\mathbf{S} = \begin{pmatrix} \frac{1}{n}, & \frac{2\cdot 1}{n(n+1)}, & \frac{6\cdot 1^2}{n(n+1)(2n+1)} \\ \frac{1}{n}, & \frac{2\cdot 2}{n(n+1)}, & \frac{6\cdot 2^2}{n(n+1)(2n+1)} \\ \vdots & \vdots & \vdots \\ \frac{1}{n}, & \frac{2\cdot n}{n(n+1)}, & \frac{6\cdot n^2}{n(n+1)(2n+1)} \end{pmatrix}.$$

It is clear that the matrix $\mathbf{S}$ is column-stochastic, and that $\mathbf{p} = \mathbf{qS}$ (see (2.208)-(2.210)).

So, by Theorem 2.7.26 (for $m = 3$) we can write

$$b_1\psi(z_{p_1}) + b_2\psi(z_{p_2}) + b_3\psi(z_{p_3}) \leq \sum_{i=1}^{n} a_i\psi(z_i) \tag{2.211}$$

where

$$\mathbf{a} = \mathbf{bS}^T \quad \text{with} \quad \mathbf{a} = (a_1, a_2, \ldots, a_n) \quad \text{and} \quad \mathbf{b} = (b_1, b_2, b_3) \tag{2.212}$$

with $b_j \geq 0$, $j = 1, 2, 3$.

By (2.212) we see that

$$(a_1, a_2, \ldots, a_n) = (b_1, b_2, b_3) \cdot \begin{pmatrix} \frac{1}{n}, & \frac{2\cdot 1}{n(n+1)}, & \frac{6\cdot 1^2}{n(n+1)(2n+1)} \\ \frac{1}{n}, & \frac{2\cdot 2}{n(n+1)}, & \frac{6\cdot 2^2}{n(n+1)(2n+1)} \\ \vdots & \vdots & \vdots \\ \frac{1}{n}, & \frac{2\cdot n}{n(n+1)}, & \frac{6\cdot n^2}{n(n+1)(2n+1)} \end{pmatrix}^T.$$

Therefore

$$a_1 = \frac{1}{n}b_1 + \frac{2\cdot 1}{n(n+1)}b_2 + \frac{6\cdot 1^2}{n(n+1)(2n+1)}b_3, \tag{2.213}$$

$$a_2 = \frac{1}{n}b_1 + \frac{2\cdot 2}{n(n+1)}b_2 + \frac{6\cdot 2^2}{n(n+1)(2n+1)}b_3, \tag{2.214}$$

$$\vdots$$

$$a_n = \frac{1}{n}b_1 + \frac{2\cdot n}{n(n+1)}b_2 + \frac{6\cdot n^2}{n(n+1)(2n+1)}b_3. \tag{2.215}$$

In addition, (2.211) is met under conditions (2.213)-(2.215), and (2.213)-(2.215) amount to (2.204). So, the proof of (2.203) is completed. □

2.9. Sherman like inequalities for strongly convex functions

In this section we deal with strongly convex functions, a subclass of convex functions with stronger versions of analogous properties. We present versions of the Lah-Ribarič inequality, Sherman's inequality and its converse for strongly convex function. As easy consequences, we get Jensen's and Majorization inequality and their conversions for strongly convex functions. We also extend Sherman's theorem to the class of strongly convex functions of higher order. Results of this section are cited from [65].

2.9.1. Strongly convex and n-strongly convex functions

For the purposes of the following results, we give definition of strongly convexity, the concept introduced by Polyak [140], which has a wide appearance in many different fields of applications, particular in many branches of mathematics as well as optimization theory, mathematical economics and approximation theory.

Definition 2.9.22.: A function $f : I \to \mathbb{R}$ is called **strongly convex** with modulus $c > 0$ if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - c\lambda(1-\lambda)(x-y)^2 \tag{2.216}$$

for all $x, y \in I$ and $\lambda \in [0,1]$. A function f is called **strongly concave** with modulus c if $-f$ is strongly convex with modulus c.

Remark 2.9.24.: Note that a function f that satisfies (2.216) with $c = 0$, i.e.

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \tag{2.217}$$

for all $x, y \in I$ and $\lambda \in [0,1]$, is convex in a usual sense. Specially, if the inequality in (2.217) is strict, then f is called **strictly convex.**

It is well known that following implications hold:

$$\text{strongly convex} \quad \Rightarrow \quad \text{strictly convex} \quad \Rightarrow \quad \text{convex.}$$

But the reverse implications are not true, in general. For example, the function $f(x) = x^2$ is strongly convex and also strictly convex and convex on $\mathbb{R}$. The function $g(x) = e^x$ is strictly convex and convex but not strongly convex on $\mathbb{R}$. The function $h(x) = x$ is convex but neither strictly nor strongly convex on $\mathbb{R}$.

Here are some examples of strongly convex functions.

Example 8.: *A square function $f(t) = t^2$ is strongly convex on $(0, \infty]$ with $c = 1$.*

Example 9.: *Let consider the functions defined on $(0, a]$ for some $a > 0$. Then:*

a) $f(t) = -\ln t$ is strongly convex with $c = \frac{1}{2a^2}$;

b) $f(t) = -\ln t$ is strongly convex with $c = \frac{1}{2a^2}$;

c) $f(t) = t \ln t$ is strongly convex with $c = \frac{1}{2a}$;

d) $f(t) = \left(\sqrt{t} - 1\right)^2$ is strongly convex with $c = \frac{1}{4\sqrt{a^3}}$.

Following R. Ger and K. Nikodem [56], we give definition of higher order strongly convexity in terms of divided differences.

Definition 2.9.23.: A function $f : I \to \mathbb{R}$ is **strongly convex of order** n with modulus $c > 0$ (or n-**strongly convex** with modulus $c > 0$) if

$$[z_0, ..., z_n; f] \geq c \tag{2.218}$$

for all $z_0, ..., z_n \in I$.

The concept of strong convexity is a strengthening of the concept of convexity and some properties of strongly convex functions are just stronger versions of analogous properties of convex functions.

We quote few useful characterizations through the following remark.

Remark 2.9.25.: a) Note that 2-strongly convex function with modulus c is just strongly convex function with modulus c as given by (2.216).

For $n = 2$, the condition (2.218) is equivalent to

$$\frac{f(z_0)}{(z_0 - z_1)(z_0 - z_2)} + \frac{f(z_1)}{(z_1 - z_2)(z_1 - z_0)} + \frac{f(z_2)}{(z_2 - z_1)(z_2 - z_0)} \geq c$$

or

$$f(z_1) \leq \frac{z_2 - z_1}{z_2 - z_0} f(z_0) + \frac{z_1 - z_0}{z_2 - z_0} f(z_2) - c(z_2 - z_1)(z_1 - z_0).$$

b) A function $f : I \to \mathbb{R}$ is strongly n-convex with modulus c iff the function $g(x) = f(x) - cx^n$ is n-convex.

c) A function $f : I \to \mathbb{R}$ is strongly n-convex with modulus c iff $f^{(n)} \geq cn!$.

d) A function $f : I \to \mathbb{R}$ is strongly convex with modulus $c > 0$ iff for every $x_0 \in IntI$, there exists $l \in \mathbb{R}$, such that

$$f(x) \geq f(x_0) + l(x - x_0) + c(x - x_0)^2, \quad x \in I, \tag{2.219}$$

i.e. f admits a quadratic support at every interior point of its domain.

For more information see [56, 134, 135].

The following Jensen's type inequality for strongly convex function is valid (see [126]).

Theorem 2.9.37: *Let* $\mathbf{x} = (x_1, ..., x_n) \in I^n$, $\mathbf{p} = (p_1, ..., p_n) \in [0, \infty)^n$ *with* $\sum\limits_{i=1}^{n} p_i = 1$ *and* $\bar{x} = \sum\limits_{i=1}^{n} p_i x_i$. *For a function* $f : I \to \mathbb{R}$ *strongly convex with modulus* $c > 0$, *the following inequality holds*

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1}^{n} p_i f(x_i) - c \sum_{i=1}^{n} p_i (x_i - \bar{x})^2. \tag{2.220}$$

If we compare (2.220) with the classical Jensen's inequality (1.6) for a convex function f, we can see that the inequality (2.220) includes a better upper bound for the term $f\left(\sum\limits_{i=1}^{n} p_i x_i\right)$ since $c \sum\limits_{i=1}^{n} p_i (x_i - \bar{x})^2 \geq 0$. Specially, for $c = 0$ the strongly convexity reduces to ordinary convexity, and (2.220) becomes (1.6).

2.9.2. Sherman's type inequalities and conversions

We start this subsection with the Lah-Ribarič type inequality for strongly convex functions (see [65, 87]).

Theorem 2.9.38: *Let $f : I \to \mathbb{R}$ be strongly convex with modulus $c > 0$. Let $\alpha, \beta \in I$, $\alpha < \beta$, $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$ and $\mathbf{a} = (a_1, ..., a_l) \in [0, \infty)^l$ with $\sum_{j=1}^{l} a_j = 1$ and $\bar{x} = \sum_{j=1}^{l} a_j x_j$. Then*

$$\sum_{j=1}^{l} a_j f(x_j) \leq \frac{\beta - \bar{x}}{\beta - \alpha} f(\alpha) + \frac{\bar{x} - \alpha}{\beta - \alpha} f(\beta) - c \sum_{j=1}^{l} a_j(\beta - x_j)(x_j - \alpha). \tag{2.221}$$

Proof: *Since for strongly convex function we have*

$$f(z_1) \leq \frac{z_2 - z_1}{z_2 - z_0} f(z_0) + \frac{z_1 - z_0}{z_2 - z_0} f(z_2) - c(z_2 - z_1)(z_1 - z_0),$$

by substituting $z_1 = x_j$, $z_2 = \beta$ and $z_1 = \alpha$, we get

$$f(x_j) \leq \frac{\beta - x_j}{\beta - \alpha} f(\alpha) + \frac{x_j - \alpha}{\beta - \alpha} f(\beta) - c(\beta - x_j)(x_j - \alpha).$$

Now, multiplying with a_j and summing over j we have

$$\sum_{j=1}^{l} a_j f(x_j) \leq \frac{\beta - \sum_{j=1}^{l} a_j x_j}{\beta - \alpha} f(\alpha) + \frac{\sum_{j=1}^{l} a_j x_j - \alpha}{\beta - \alpha} f(\beta) - c \sum_{j=1}^{l} a_j(\beta - x_j)(x_j - \alpha),$$

as claimed. □

Now we give Sherman's type inequality for strongly convex functions.

Theorem 2.9.39: *Let $f : [\alpha, \beta] \to \mathbb{R}$ be strongly convex with modulus $c > 0$. Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in [0, \infty)^l$, $\mathbf{b} = (b_1, ..., b_m) \in [0, \infty)^m$ and $\mathbf{A} = (a_{ij})$ be $l \times m$ row stochastic matrix such that*

$$\mathbf{a} = \mathbf{b}\mathbf{A} \quad \text{and} \quad \mathbf{y} = \mathbf{x}\mathbf{A}^T. \tag{2.222}$$

Then

$$\sum_{i=1}^{m} b_i f(y_i) \leq \sum_{j=1}^{l} a_j f(x_j) - c\left(\sum_{j=1}^{l} a_j x_j^2 - \sum_{i=1}^{m} b_i y_i^2\right). \tag{2.223}$$

Proof: *Since Sherman's conditions (2.222) hold, i.e.* $y_i = \sum_{j=1}^{l} x_j a_{ij}$, $i = 1, ..., m$ *and* $a_j = \sum_{i=1}^{m} b_i a_{ij}$, $j = 1, ..., l$, *applying (2.220), we have*

$$\begin{aligned}\sum_{i=1}^{m} b_i f(y_i) &= \sum_{i=1}^{m} b_i f\left(\sum_{j=1}^{l} x_j a_{ij}\right) \\ &\leq \sum_{i=1}^{m} b_i \left(\sum_{j=1}^{l} a_{ij} f(x_j) - c\sum_{j=1}^{l} a_{ij}(x_j - y_i)^2\right) \\ &= \sum_{j=1}^{l} a_j f(x_j) - c\sum_{i=1}^{m} b_i \sum_{j=1}^{l} a_{ij}(x_j - y_i)^2.\end{aligned} \tag{2.224}$$

By an easy calculation, we get

$$\begin{aligned}&\sum_{j=1}^{l} a_j f(x_j) - c\sum_{i=1}^{m} b_i \sum_{j=1}^{l} a_{ij}(x_j - y_i)^2 \\ &= \sum_{j=1}^{l} a_j f(x_j) - c\sum_{i=1}^{m} b_i \sum_{j=1}^{l} a_{ij}(x_j^2 - 2x_j y_i + y_i^2) \\ &= \sum_{j=1}^{l} a_j f(x_j) - c\left(\sum_{j=1}^{l} a_j x_j^2 - \sum_{i=1}^{m} b_i y_i^2\right).\end{aligned} \tag{2.225}$$

Now, combining (2.224) and (2.225), we get (2.223). □

Remark 2.9.26.: a) If we compare (2.223) with Sherman's inequality (2.10) for convex functions, we can see that the inequality (2.223) includes a better upper bound for $\sum_{i=1}^{m} b_i f(y_i)$ since $c\left(\sum_{j=1}^{l} a_j x_j^2 - \sum_{i=1}^{m} b_i y_i^2\right) \geq 0$ because $t \mapsto t^2$ is convex function and then by Sherman's inequality we have $\sum_{j=1}^{l} a_j x_j^2 - \sum_{i=1}^{m} b_i y_i^2 \geq 0$. Moreover, we get the double inequality

$$\sum_{i=1}^{m} b_i f(y_i) \leq \sum_{j=1}^{l} a_j f(x_j) - c\left(\sum_{j=1}^{l} a_j x_j^2 - \sum_{i=1}^{m} b_i y_i^2\right) \leq \sum_{j=1}^{l} a_j f(x_j). \tag{2.226}$$

b) A classical majorization is a special case of the previous consideration. Namely, if we put $m = l$ and take $\mathbf{a} = \mathbf{b} = \mathbf{e} = (1, ..., 1) \in [0, \infty)^m$, then (2.222) becomes

$$\mathbf{e} = \mathbf{eA} \quad \text{and} \quad \mathbf{y} = \mathbf{xA}^T.$$

The first condition $\mathbf{e} = \mathbf{eA}$ implies that $\mathbf{A}$ is also column stochastic. Moreover, $\mathbf{A}$ is doubly stochastic matrix. Then, the second condition $\mathbf{y} = \mathbf{xA}^T$ implies a classical majorization $\mathbf{y} \prec \mathbf{x}$. Moreover, the inequality (2.223) reduces to

$$\sum_{i=1}^{m} f(y_i) \leq \sum_{i=1}^{m} f(x_i) - c\left(\sum_{i=1}^{m} x_i^2 - \sum_{i=1}^{m} y_i^2\right) \leq \sum_{i=1}^{m} f(x_i),$$

which presents Majorization inequality for strongly convex functions. This inequality improves Majorization inequiality for convex functions (2.4).

c) Specially, for $m = 1$ and $\mathbf{b} = (1)$, the inequality (2.226) becomes

$$f\left(\sum_{j=1}^{l} a_j x_j\right) \leq \sum_{j=1}^{l} a_j f(x_j) - c\left(\sum_{j=1}^{l} a_j x_j^2 - \bar{x}^2\right) = \sum_{j=1}^{l} a_j f(x_j) - c\sum_{j=1}^{l} a_j (x_j - \bar{x})^2 \leq \sum_{j=1}^{l} a_j f(x_j),$$

where $\bar{x} = \sum_{j=1}^{l} a_j x_j$, i.e. we get Jensen's inequality for strongly convex functions (2.220).

Next we give conversion to Sherman's inequality for strongly convex functions.

Theorem 2.9.40: *Let $f:[\alpha,\beta]\to\mathbb{R}$ be strongly convex with modulus $c>0$. Let $\mathbf{x}=(x_1,...,x_l)\in[\alpha,\beta]^l$, $\mathbf{y}=(y_1,...,y_m)\in[\alpha,\beta]^m$, $\mathbf{a}=(a_1,...,a_l)\in[0,\infty)^l$, $\mathbf{b}=(b_1,...,b_m)\in[0,\infty)^m$ and $\mathbf{A}=(a_{ij})$ be $l\times m$ row stochastic matrix such that (2.222) holds. Denote $\sum_{j=1}^m b_j=B_m$. Then*

$$\sum_{j=1}^{l} a_j f(x_j) \le \frac{B_m\beta-\sum_{j=1}^{l} a_j x_j}{\beta-\alpha} f(\alpha)+\frac{\sum_{j=1}^{l} a_j x_j-B_m\alpha}{\beta-\alpha} f(\beta) \tag{2.227}$$
$$-c\sum_{j=1}^{l} a_j(\beta-x_j)(x_j-\alpha).$$

Proof: *Using (2.222) we have*

$$\sum_{j=1}^{l} a_j f(x_j)=\sum_{j=1}^{l}\left(\sum_{i=1}^{m} b_j a_{ij}\right) f(x_j)=\sum_{i=1}^{m} b_j\left(\sum_{j=1}^{l} a_{ij} f(x_j)\right). \tag{2.228}$$

Applying (2.221)-(2.222) we get

$$\sum_{j=1}^{l} a_{ij} f(x_j) \le \frac{\beta-\sum_{j=1}^{l} a_{ij} x_j}{\beta-\alpha} f(\alpha)+\frac{\sum_{j=1}^{l} a_{ij} x_j-\alpha}{\beta-\alpha} f(\beta) \tag{2.229}$$
$$-c\sum_{j=1}^{l} a_{ij}(\beta-x_j)(x_j-\alpha).$$

Now, combining (2.228) and (2.229), we get (2.227). □

Remark 2.9.27.: a) Specially, if $m=1$ and $\mathbf{b}=(1)$, then (2.226) and (2.227) gives the following series of inequalities

$$\begin{aligned} f\left(\sum_{j=1}^{l} a_j x_j\right) &\le \sum_{j=1}^{l} a_j f(x_j)-c\sum_{j=1}^{l} a_j\left(x_j-\bar{x}\right)^2\\ &\le \sum_{j=1}^{l} a_j f(x_j)\\ &\le \frac{\beta-\sum_{j=1}^{l} a_j x_j}{\beta-\alpha} f(\alpha)+\frac{\sum_{j=1}^{l} a_j x_j-\alpha}{\beta-\alpha} f(\beta)\\ &\quad -c\sum_{j=1}^{l} a_j(\beta-x_j)(x_j-\alpha). \end{aligned}$$

i.e. we get Jensen's inequality and its conversion for strongly convex functions.

b) If $m = l$ and $\mathbf{b} = \mathbf{e} = (1, ..., 1)$, then (2.226) and (2.227) gives

$$\begin{aligned}\sum_{i=1}^{m} f(y_i) &\le \sum_{j=1}^{l} f(x_j) - c\left(\sum_{j=1}^{l} x_j^2 - \sum_{i=1}^{m} y_i^2\right)\\ &\le \sum_{j=1}^{l} f(x_j)\\ &\le \frac{\bar{\beta} - \sum_{j=1}^{l} x_j}{\beta - \alpha} f(\alpha) + \frac{\sum_{j=1}^{l} x_j - \bar{\alpha}}{\beta - \alpha} f(\beta)\\ &- c\sum_{j=1}^{l} (\beta - x_j)(x_j - \alpha),\end{aligned}$$

where $\bar{\beta} = m\beta$ and $\bar{\alpha} = m\alpha$, i.e. we get Majorization inequality and its conversion for strongly convex functions. This result improves (2.4) and (2.12).

2.9.3. Generalization of Sherman's inequality for n-strongly convex function

In this subsection we present generalization of Sherman's inequality to the class of strongly convex functions of higher order. To prove that result we need the following identity, known as Fink's identity (see [47]).

Theorem 2.9.41: *[47] Let $n \in \mathbb{N}$, $n \ge 1$, and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$. Then*

$$\begin{aligned}f(s) = \frac{n}{\beta - \alpha}\int_{\alpha}^{\beta} f(t)dt - \sum_{w=1}^{n-1} \frac{n - w}{w!} \cdot \frac{f^{(w-1)}(\alpha)(s - \alpha)^w - f^{(w-1)}(\beta)(s - \beta)^w}{\beta - \alpha}\\ + \frac{1}{(n-1)!(\beta - \alpha)}\int_{\alpha}^{\beta} (s - t)^{n-1} P(t, s) f^{(n)}(t)dt, \qquad (2.230)\end{aligned}$$

where

$$P(t, s) = \begin{cases} t - \alpha, & \alpha \le t \le s \le \beta \\ t - \beta, & \alpha \le s < t \le \beta \end{cases}. \qquad (2.231)$$

The sum in (2.230) is zero when $n = 1$.

Now we prove new identity that involves Sherman's difference.

Theorem 2.9.42: *Let $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$. Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in [0, \infty)^l$, $\mathbf{b} = (b_1, ..., b_m) \in [0, \infty)^m$ and $\mathbf{A} = (a_{ij})$ be $l \times m$ row stochastic matrix such that (2.222) holds. Then*

$$
\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (2.232)\\
&= \frac{1}{\beta - \alpha} \sum_{w=2}^{n-1} \frac{n-w}{w!} \cdot f^{(w-1)}(\beta) \left(\sum_{j=1}^{l} a_j (x_j - \beta)^w - \sum_{i=1}^{m} b_i (y_i - \beta)^w \right)\\
&- \frac{1}{\beta - \alpha} \sum_{w=2}^{n-1} \frac{n-w}{w!} \cdot f^{(w-1)}(\alpha) \left(\sum_{j=1}^{l} a_j (x_j - \alpha)^w - \sum_{i=1}^{m} b_i (y_i - \alpha)^w \right)\\
&+ \frac{1}{(n-1)!(\beta - \alpha)} \times\\
&\int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j (x_j - t)^{n-1} P(t, x_j) - \sum_{i=1}^{m} b_i (y_i - t)^{n-1} P(t, y_i) \right] f^{(n)}(t) dt.
\end{aligned}
$$

Proof: *Applying (2.230) to the Sherman difference $\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)$, we get (2.232).*

□

The following generalization of Sherman's inequality is valid.

Theorem 2.9.43: *Let all the assumptions of Theorem 2.9.42 be satisfied. Additionally, let f be n-strongly convex with modulus $c > 0$.*

If

$$
\sum_{j=1}^{l} a_j (x_j - t)^{n-1} P(t, x_j) - \sum_{i=1}^{m} b_i (y_i - t)^{n-1} P(t, y_i) \geq 0, \quad t \in [\alpha, \beta], \qquad (2.233)
$$

then

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) - c\left(\sum_{j=1}^{l} a_j x_j^n - \sum_{i=1}^{m} b_i y_i^n\right) \tag{2.234}$$
$$\geq \frac{1}{\beta-\alpha}\sum_{w=1}^{n-1}\frac{n-w}{w!}$$
$$\cdot\left[f^{(w-1)}(\beta) - cn(n-1)...(n-w+2)\beta^{n-w+2}\right]\left(\sum_{j=1}^{l} a_j (x_j-\beta)^w - \sum_{i=1}^{m} b_i (y_i-\beta)^w\right)$$
$$-\frac{1}{\beta-\alpha}\sum_{w=1}^{n-1}\frac{n-w}{w!}$$
$$\cdot\left[f^{(w-1)}(\alpha) - cn(n-1)...(n-w+2)\alpha^{n-w+2}\right]\left(\sum_{j=1}^{l} a_j (x_j-\alpha)^w - \sum_{i=1}^{m} b_i (y_i-\alpha)^w\right).$$

If the reverse inequality in (2.233) holds, then the reverse inequality in (2.234) holds.

Proof: *Let us consider the function $g(x) = f(x) - cx^n$. Since f is strongly n-convex with modulus c, then g is n-convex. We may assume without loss of generality that f and g are n-times differentiable and $g^{(n)} \geq 0$ on $[\alpha, \beta]$ (see [138, p. 16]).*

Applying (2.232) to g, we have

$$\sum_{j=1}^{l} a_j g(x_j) - \sum_{i=1}^{m} b_i g(y_i) \tag{2.235}$$
$$= \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot g^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j (x_j-\beta)^w - \sum_{i=1}^{m} b_i (y_i-\beta)^w\right)$$
$$-\frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot g^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j (x_j-\alpha)^w - \sum_{i=1}^{m} b_i (y_i-\alpha)^w\right)$$
$$+\frac{1}{(n-1)!(\beta-\alpha)}\times$$
$$\int_{\alpha}^{\beta}\left[\sum_{j=1}^{l} a_j (x_j-t)^{n-1}P(t,x_j) - \sum_{i=1}^{m} b_i (y_i-t)^{n-1}P(t,y_i)\right] g^{(n)}(t)dt.$$

Moreover, if (2.233) holds, then

$$\begin{aligned}
&\sum_{j=1}^{l} a_j g(x_j) - \sum_{i=1}^{m} b_i g(y_i) \qquad (2.236)\\
&\geq \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot g^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j (x_j-\beta)^w - \sum_{i=1}^{m} b_i (y_i-\beta)^w\right)\\
&- \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot g^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j (x_j-\alpha)^w - \sum_{i=1}^{m} b_i (y_i-\alpha)^w\right)
\end{aligned}$$

which is equivalent to (2.234).

If the reverse inequality in (2.233) holds, then the last term in (2.235) is nonpositive and then the reverse inequality in (2.236) holds. This ends the proof. □

Remark 2.9.28.: Consider the function $\sigma : [\alpha,\beta] \to \mathbb{R}$ defined by

$$\sigma(x) = (x-t)^{n-1}P(t,x) = \begin{cases} (x-t)^{n-1}(t-\alpha), & \alpha \leq t \leq x \leq \beta \\ (x-t)^{n-1}(t-\beta), & \alpha \leq x < t \leq \beta \end{cases}.$$

We have

$$\sigma''(x) = \begin{cases} (n-1)(n-2)(x-t)^{n-3}(t-\alpha), & \alpha \leq t \leq x \leq \beta \\ (n-1)(n-2)(x-t)^{n-3}(t-\beta), & \alpha \leq x < t \leq \beta \end{cases}.$$

Then for even n, the function σ is convex and by Sherman's theorem, we have

$$\sum_{j=1}^{l} a_j (x_j-t)^{n-1}P(t,x_j) - \sum_{i=1}^{m} b_i (y_i-t)^{n-1}P(t,y_i) \geq 0,$$

i.e. the assumption (2.233) is immediately satisfied. Therefore, by Theorem 2.9.43, the inequality (2.234) holds.

Specially, for $n=2$, the inequality (2.234) reduces to (2.223).

2.10. Sherman like inequalities for superquadratic functions

In this section we consider superquadratic functions, the class of functions closely connecting with the class of convex functions. It is important to emphasize that nonnegative superquadratic functions have the property that they are also convex. This property enables us to obtain improvements of inequalities for convex functions. We extend Sherman's theorem to superquadratic functions. More precisely, for superquadratic functions which are not convex, we obtain results analog ones for convex functions. For nonnegative superquadratic functions which are also convex, we get improvements. Also, we present related conversions, as well as Jensen's and Majorization type inequalities and their conversions. The results from this section is cited from [64].

2.10.1. Superquadratic functions

A class of superquadratic functions was introduced and initially studied by S. Abramovich and G. Jameson [3, 4]. Name "superquadratic functions" is justified by the fact that nonnegative superquadratic functions, which are shown to be convex, grow equally fast or even faster than the quadratic function (property of superquadratic growth). The role played by affine function in the case of convex functions, in this class of functions there is exactly a quadratic function, so superquadratic functions are somehow more "convex" than quadratic functions.

We introduce the definition of superquadratic functions and quote some basic properties. For more informations see [3, 4].

Definition 2.10.24.: A function $\varphi : [0, \infty) \to \mathbb{R}$ is **superquadratic** provided that for all $x \in [0, \infty)$, there exists a constant $C_x \in \mathbb{R}$ such that

$$\varphi(y) \geq \varphi(x) + C_x(y - x) + \varphi(|y - x|) \tag{2.237}$$

for all $y \in [0, \infty)$. We say that φ is **subquadratic** if $-\varphi$ is a superquadratic function.

The following theorem gives Jensen's inequality for superquadratic functions (see [4]).

Theorem 2.10.44: *Let* $\mathbf{x} = (x_1, \ldots, x_k) \in [0, \infty)$ *and* $\mathbf{u} = (u_1, \ldots, u_k)$ *be a nonnegative* k*-tuple such that* $\sum_{i=1}^{k} u_i = 1$ *and* $\overline{x} = \sum_{i=1}^{k} u_i x_i$*. Then for every superquadratic function* $\varphi : [0, \infty) \to \mathbb{R}$*, we have*

$$\varphi(\overline{x}) + \sum_{i=1}^{k} u_i \varphi(|x_i - \overline{x}|) \le \sum_{i=1}^{k} u_i \varphi(x_i). \tag{2.238}$$

Moreover, if φ *is subquadratic, then the reverse of (2.238) holds.*

Definition 2.10.25.: A function $\varphi : [0, \infty) \to \mathbb{R}$ is **superadditive** if

$$\varphi(x + y) \ge \varphi(x) + \varphi(y), \tag{2.239}$$

holds for all $x, y \ge 0$. If the reverse of (2.239) holds, then φ is **subadditive**.

The next lemma gives sufficient conditions for superquadracity.

Lemma 2.10.5.: *[4, Lemma 3.1] Suppose that* $\varphi : [0, \infty) \to \mathbb{R}$ *is a continuously differentiable function and* $\varphi(0) \le 0$ $(\varphi(0) = 0)$*. If* φ' *is superadditive (subadditive) or* $\frac{\varphi'(x)}{x}$ *is nondecreasing (nonincreasing), then* φ *is superquadratic (subquadratic).*

The following lemma gives same basic properties of superquadratic functions.

Lemma 2.10.6.: *[4, Lemma 2.2] Let* $\varphi : [0, \infty) \to \mathbb{R}$ *be a superquadratic function with* C_x *as in (2.237).*

(i) Then $\varphi(0) \le 0$.

(ii) If $\varphi(0) = \varphi'(0) = 0$*, then* $C_x = \varphi'(x)$ *whenever* φ *is differentiable at* $x > 0$.

(iii) If $\varphi \ge 0$*, then* φ *is convex and* $\varphi(0) = \varphi'(0) = 0$.

The property (iii) from the previous lemma enables us to obtain improvements of various inequalities that apply to convex functions. Thus, nonnegative superquadratic functions have the property that they are also convex. However, as can be seen from the below examples, within the class of superquadratic functions we can find negative functions as well as functions that change sign, then concave functions, as well as those that are neither convex nor concave at the considered interval within the domain.

Example 10.: *a) The function $\varphi(x) = x^p$, is superquadratic for $p \geq 2$ and subquadratic for $0 < p \leq 2$. Further, if φ is defined by $\varphi(x) = \operatorname{sgn}(p-2)x^p$, then φ is superquadratic for all $p > 0$.*

b) The function

$$\varphi(x) = \begin{cases} xe^{-\frac{1}{x}}, & x > 0 \\ 0, & x = 0 \end{cases}$$

is convex on $[0, \infty)$, but superquadratic only on $[0, L]$, for $0.8955 \leq L \leq 1$.

c) Superquadratic function is

$$\varphi(x) = \begin{cases} x^2 \ln x, & x > 0 \\ 0, & x = 0 \end{cases}.$$

d) The functions $f(t) = \ln(1+t) - t$ is superquadratic.

e) Any function $\varphi : [0, \infty) \to [-2, -1]$ is superquadratic (satisfies (2.237) for $C_x = 0$).

For more informations see [1, 3, 4, 19].

2.10.2. Improvements of majorization theorems

To prove our main results we need the following technical lemma.

Lemma 2.10.7.*([64])*: *Let $\varphi : [0, \infty) \to \mathbb{R}$ be a continuously differentiable and $\varphi' : [0, \infty) \to \mathbb{R}$ be a superadditive function. Then the function $\psi : [0, \infty) \to \mathbb{R}$, defined by*

$$\psi(y) = \varphi(y) - \varphi(z) - \varphi'(z)(y-z) - \varphi(|y-z|) + \varphi(0),$$

is nonnegative for all $z \in [0, \infty)$.

If φ' is subadditive, then ψ is nonpositive for all $z \in [0, \infty)$.

Proof: Since φ' is superadditive, then for $0 \leq z \leq y$ we have

$$0 \leq \int_z^y \left(\varphi'(t) - \varphi'(z) - \varphi'(t-z)\right) \mathrm{d}t$$
$$= \varphi(y) - \varphi(z) - \varphi'(z)(y-z) - \varphi(y-z) + \varphi(0)$$

and for $0 \leq y \leq z$,

$$0 \leq \int_y^z \left(\varphi'(z) - \varphi'(t) - \varphi'(z-t)\right) \mathrm{d}t$$
$$= \varphi'(z)(z-y) - \varphi(z) + \varphi(y) + \varphi(0) - \varphi(z-y).$$

Together these show that for any $y, z \in [0, \infty)$ the inequality

$$\varphi(y) - \varphi(z) - \varphi'(z)(y-z) - \varphi(|y-z|) + \varphi(0) \geq 0$$

holds. So we conclude that the function ψ is nonnegative on $[0, \infty)$.

The case when φ' is subadditive we can prove analogously. $\square$

In the sequel we use the following notation.

Let $\mathbf{x} = (x_1, ..., x_l)$ and $\mathbf{y} = (y_1, ..., y_k)$ be vectors with nonnegative entries and $\mathbf{u} = (u_1, ..., u_l)$ and $\mathbf{v} = (v_1, ..., v_k)$ be nonnegative weights. We write

$$(\mathbf{y}, \mathbf{v}) \prec (\mathbf{x}, \mathbf{u})$$

if $\mathbf{y} = \mathbf{x}\mathbf{T}^T$ and $\mathbf{u} = \mathbf{v}\mathbf{T}$, i.e. $y_i = \sum_{j=1}^{l} x_j t_{ij}$, $i = 1, ..., k$, and $u_j = \sum\limits_{i=1}^{k} v_i t_{ij}$, $j = 1, ..., l$, holds for some row stochastic matrix $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{kl}(\mathbb{R})$.

One can write $(\mathbf{y}, \mathbf{v}) \prec_T (\mathbf{x}, \mathbf{u})$ instead $(\mathbf{y}, \mathbf{v}) \prec (\mathbf{x}, \mathbf{u})$.

Main results of this section contain Sherman's type inequalities as follows.

Theorem 2.10.45 ***([64])*:** *Let $\mathbf{x} = (x_1, ..., x_l)$, $\mathbf{y} = (y_1, ..., y_k)$ be vectors with nonnegative entries, $\mathbf{u} = (u_1, ..., u_l)$, $\mathbf{v} = (v_1, ..., v_k)$ be nonnegative weights and $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{kl}(\mathbb{R})$ be row stochastic matrix such that $(\mathbf{y}, \mathbf{v}) \prec_T (\mathbf{x}, \mathbf{u})$.*

If $\varphi : [0, \infty) \to \mathbb{R}$ is a continuously differentiable and $\varphi' : [0, \infty) \to \mathbb{R}$ is a superadditive *function, then*

$$\sum_{i=1}^{k} v_i\varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{l} v_i t_{ij}\varphi\left(|x_j - y_i|\right) \leq \sum_{j=1}^{l} u_j\varphi\left(x_j\right) + \sum_{i=1}^{k} v_i\varphi(0). \tag{2.240}$$

a) If in addition $\varphi\left(0\right) \leq 0$, *then* φ *is superquadratic and*

$$\sum_{i=1}^{k} v_i\varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{l} v_i t_{ij}\varphi\left(|x_j - y_i|\right) \leq \sum_{j=1}^{l} u_j\varphi\left(x_j\right). \tag{2.241}$$

b) If in addition $\varphi \geq 0$ *and* $\varphi(0) = \varphi^{'}(0) = 0$, *then* φ *is superquadratic and convex increasing and*

$$\sum_{i=1}^{k} v_i\varphi(y_i) \leq \sum_{i=1}^{k} v_i\varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{l} v_i t_{ij}\varphi\left(|x_j - y_i|\right) \leq \sum_{j=1}^{l} u_j\varphi\left(x_j\right). \tag{2.242}$$

Proof: By Lemma 2.10.7., choosing $y = x_j$ we have

$$\psi\left(x_j\right) = \varphi\left(x_j\right) - \varphi\left(z\right) - \varphi'\left(z\right)\left(x_j - z\right) - \varphi\left(|x_j - z|\right) + \varphi\left(0\right) \geq 0.$$

Now, multiplying with t_{ij} and summing over $j = 1, ..., l$, we get

$$\begin{aligned}&\sum_{j=1}^{l} t_{ij}\psi\left(x_j\right)\\ &= \sum_{j=1}^{l} t_{ij}\varphi\left(x_j\right) - \varphi\left(z\right) - \varphi'\left(z\right)\left(\sum_{j=1}^{l} t_{ij}x_j - z\right) - \sum_{j=1}^{l} t_{ij}\varphi\left(|x_j - z|\right) + \varphi\left(0\right) \geq 0.\end{aligned}$$

Inserting $z = y_i = \sum_{j=1}^{l} t_{ij}x_j$ we get

$$\sum_{j=1}^{l} t_{ij}\varphi\left(x_j\right) - \varphi\left(y_i\right) - \sum_{j=1}^{l} t_{ij}\varphi\left(|x_j - y_i|\right) + \varphi\left(0\right) \geq 0,$$

i.e.

$$\varphi\left(y_i\right) + \sum_{j=1}^{l} t_{ij}\varphi\left(|x_j - y_i|\right) \leq \sum_{j=1}^{l} t_{ij}\varphi\left(x_j\right) + \varphi\left(0\right), \quad \text{for } i = 1, ..., k.$$

Multiplying with v_i and summing over $i = 1, ..., k$, we get

$$\sum_{i=1}^{k} v_i \varphi(y_i) + \sum_{i=1}^{k} v_i \sum_{j=1}^{l} t_{ij} \varphi(|x_j - y_i|) \leq \sum_{i=1}^{k} v_i \sum_{j=1}^{l} t_{ij} \varphi(x_j) + \sum_{i=1}^{k} v_i \varphi(0). \tag{2.243}$$

Since $\sum_{i=1}^{k} v_i t_{ij} = u_j$, then (2.243) is equivalent to (2.240), as required.

a) Since $\varphi(0) \leq 0$, then by Lemma 2.10.5. φ is superquadratic. The inequality (2.241) follows from (2.240).

b) Since $\varphi(0) = 0$, then by Lemma 2.10.5. φ is superquadratic. Therefore, the convexity follows from Lemma 2.10.6.. Because of $\varphi \geq 0$ and $\varphi(0) = 0$, it follows that φ is increasing. Moreover, since φ is nonnegative, then

$$\sum_{i=1}^{k} \sum_{j=1}^{l} v_i t_{ij} \varphi(|x_j - y_i|) \geq 0. \tag{2.244}$$

Therefore, together with

$$\sum_{i=1}^{k} v_i \varphi(y_i) + \sum_{i=1}^{k} \sum_{j=1}^{l} v_i t_{ij} \varphi(|x_j - y_i|) \leq \sum_{j=1}^{l} u_j \varphi(x_j),$$

it gives (2.242). This completes the proof. □

Note that our results for convex functions (2.242) improves Sherman's inequality (2.10). An analogous inequality to (2.241) obtained by a different approach can be found in [123].

Remark 2.10.29.: Choosing $k = 1$, $v_1 = 1$, $t_{1j} = u_j$, $j = 1, ..., l$, $y_1 = \sum_{j=1}^{l} u_j x_j$ with $\bar{x} = \sum_{j=1}^{l} u_j x_j$, from Theorem 2.10.45 we obtain analogous Jensen's type inequalities. The inequality (2.240) reduces to

$$\varphi(\bar{x}) + \sum_{j=1}^{l} u_j \varphi(|x_j - y_i|) \leq \sum_{j=1}^{l} u_j \varphi(x_j) + \varphi(0).$$

a) If in addition $\varphi(0) \leq 0$, then φ is superquadratic and

$$\varphi(\bar{x}) + \sum_{j=1}^{l} u_j \varphi(|x_j - \bar{x}|) \leq \sum_{j=1}^{l} u_j \varphi(x_j).$$

b) If in addition $\varphi \geq 0$ and $\varphi(0) = \varphi^{'}(0) = 0$, then φ is superquadratic and convex increasing and

$$\varphi\left(\bar{x}\right) \leq \varphi\left(\bar{x}\right) + \sum_{j=1}^{l} u_j \varphi\left(\left|x_j - \bar{x}\right|\right) \leq \sum_{j=1}^{l} u_j \varphi\left(x_j\right).$$

Choosing $k = l$ and $\mathbf{v} = \mathbf{e} = (1, ..., 1)$ in Theorem 2.10.45, the condition $(\mathbf{y}, \mathbf{v}) \prec (\mathbf{x}, \mathbf{u})$ implies $\mathbf{y} = \mathbf{x}\mathbf{T}^T$ with some doubly stochastic matrix $\mathbf{T}$ and $\mathbf{u} = \mathbf{v}\mathbf{T}=\mathbf{e} = (1, ..., 1)$, and as direct consequence of Theorem 2.10.45 we get the following corollary which gives analogous results for Majorization inequality (2.5).

Corollary 2.10.16.: *Let $\mathbf{x} = (x_1, ..., x_k)$, $\mathbf{y} = (y_1, ..., y_k)$ be vectors with nonnegative entries and $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{kk}(\mathbb{R})$ be doubly stochastic matrix such that $\mathbf{y} = \mathbf{x}\mathbf{T}^\top$.*

If $\varphi : [0, \infty) \to \mathbb{R}$ is a continuously differentiable and $\varphi' : [0, \infty) \to \mathbb{R}$ is a superadditive function, then

$$\sum_{i=1}^{k} \varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{k} t_{ij} \varphi\left(\left|x_j - y_i\right|\right) \leq \sum_{i=1}^{k} \varphi\left(x_i\right) + k\varphi(0).$$

a) If in addition $\varphi\left(0\right) \leq 0$, then φ is superquadratic and

$$\sum_{i=1}^{k} \varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{k} t_{ij} \varphi\left(\left|x_j - y_i\right|\right) \leq \sum_{i=1}^{k} \varphi\left(x_i\right).$$

b) If in addition $\varphi \geq 0$ and $\varphi(0) = \varphi^{'}(0) = 0$, then φ is superquadratic and convex increasing and

$$\sum_{i=1}^{k} \varphi(y_i) \leq \sum_{i=1}^{k} \varphi(y_i) + \sum_{i=1}^{k}\sum_{j=1}^{k} t_{ij} \varphi\left(\left|x_j - y_i\right|\right) \leq \sum_{i=1}^{k} \varphi\left(x_i\right). \tag{2.245}$$

Remark 2.10.30.: Our result (2.245) applied to convex functions improves Majorization inequality (2.5).

Theorem 2.10.46 ***([64])*****:** *Let $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v}$ and $\mathbf{T}$ be as in Theorem 2.10.45. If $\varphi : [0, \infty) \to \mathbb{R}$ is a continuously differentiable and $\varphi' : [0, \infty) \to \mathbb{R}$ is a subadditive function, then the reverse of (2.240) holds. If in addition $\varphi\left(0\right) = 0$, then φ is subquadratic and the reverse of (2.241) holds.*

Proof: *We use the same technique as in the proof of (2.240) and (2.241). We omit the details.* □

2.10.3. Conversions via superquadracity

The next remark, which is direct consequence of Theorem 2.10.44, will be very useful for us in proving the following conversions of results from the previous subsection.

Remark 2.10.31.: From (2.238), specially for $k=2$, $u_1=t\in[0,1]$, $u_2=1-t$ and $x_1=z_1\geq 0$, $x_2=z_2\geq 0$, we get

$$\begin{aligned}&\varphi\left(tz_1+(1-t)z_2\right)+t\varphi\left((1-t)\left|z_1-z_2\right|\right)+(1-t)\varphi\left(t\left|z_1-z_2\right|\right)\\&\leq t\varphi(z_1)+(1-t)\varphi(z_2)\end{aligned}\tag{2.246}$$

Further, let $x,z_1,z_2\geq 0$ be such that $z_1<x<z_2$. Then there exists $t\in[0,1]$ such that $x=tz_1+(1-t)\,z_2$. It follows

$$t=\frac{z_2-x}{z_2-z_1}\quad\text{and}\quad 1-t=\frac{x-z_1}{z_2-z_1}.$$

Replacing that in (2.246), we get

$$\varphi\left(x\right)+\frac{z_2-x}{z_2-z_1}\varphi\left(x-z_1\right)+\frac{x-z_1}{z_2-z_1}\varphi\left(z_2-x\right)\leq\frac{z_2-x}{z_2-z_1}\varphi(z_1)+\frac{x-z_1}{z_2-z_1}\varphi(z_2),$$

which is equivalent to

$$\varphi\left(x\right)\leq\frac{z_2-x}{z_2-z_1}\left[\varphi(z_1)-\varphi\left(x-z_1\right)\right]+\frac{x-z_1}{z_2-z_1}\left[\varphi(z_2)-\varphi\left(z_2-x\right)\right].\tag{2.247}$$

Theorem 2.10.47*([64])*: *Let* $\mathbf{x}=(x_1,...,x_l)$, $\mathbf{y}=(y_1,...,y_k)$ *be vectors with entries from* $[m,M]$ $(0\leq m<M)$, $\mathbf{u}=(u_1,...,u_l)$, $\mathbf{v}=(v_1,...,v_k)$ *be nonnegative weights such that* $(\mathbf{y},\mathbf{v})\prec(\mathbf{x},\mathbf{u})$.

If $\varphi : [0, \infty) \to \mathbb{R}$ is a superquadratic function, then

$$\sum_{j=1}^{l} u_j \varphi(x_j) + \Omega \leq \frac{M \sum\limits_{i=1}^{k} v_i - \sum\limits_{i=1}^{k} v_i y_i}{M-m} \varphi(m) + \frac{\sum\limits_{i=1}^{k} v_i y_i - m \sum\limits_{i=1}^{k} v_i}{M-m} \varphi(M), \tag{2.248}$$

where

$$\Omega = \frac{1}{M-m} \sum_{j=1}^{l} u_j \left[(M - x_j) \varphi(x_j - m) + (x_j - m) \varphi(M - x_j) \right]. \tag{2.249}$$

If in addition $\varphi \geq 0$, then φ is convex and the following series of inequalities hold

$$\begin{aligned} \sum_{i=1}^{k} v_i \varphi(y_i) &\leq \sum_{j=1}^{l} u_j \varphi(x_j) \\ &\leq \frac{M \sum\limits_{i=1}^{k} v_i - \sum\limits_{i=1}^{k} v_i y_i}{M-m} \varphi(m) + \frac{\sum\limits_{i=1}^{k} v_i y_i - m \sum\limits_{i=1}^{k} v_i}{M-m} \varphi(M) - \Omega \\ &\leq \frac{M \sum\limits_{i=1}^{k} v_i - \sum\limits_{i=1}^{k} v_i y_i}{M-m} \varphi(m) + \frac{\sum\limits_{i=1}^{k} v_i y_i - m \sum\limits_{i=1}^{k} v_i}{M-m} \varphi(M). \end{aligned} \tag{2.250}$$

Proof: *Substituting $z_1 = m, z_2 = M$ and $x = x_j$ in (2.247) we get*

$$\varphi(x_j) \leq \frac{M - x_j}{M-m} \left[\varphi(m) - \varphi(x_j - m) \right] + \frac{x_j - m}{M-m} \left[\varphi(M) - \varphi(M - x_j) \right],$$

which is equivalent to

$$\varphi(x_j) + \frac{M - x_j}{M-m} \varphi(x_j - m) + \frac{x_j - m}{M-m} \varphi(M - x_j) \leq \frac{M - x_j}{M-m} \varphi(m) + \frac{x_j - m}{M-m} \varphi(M).$$

Since $(\mathbf{x}, \mathbf{u}) \prec (\mathbf{y}, \mathbf{v})$, there exists a row stochastic natrix $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{kl}(\mathbb{R})$ such that $(\mathbf{x}, \mathbf{u}) \prec_T (\mathbf{y}, \mathbf{v})$.

Now, multiplying with t_{ij} and summing over $j=1,...,l$, we get

$$\sum_{j=1}^{l} t_{ij}\varphi(x_j) + \sum_{j=1}^{l} t_{ij}\left(\frac{M-x_j}{M-m}\varphi(x_j-m) + \frac{x_j-m}{M-m}\varphi(M-x_j)\right)$$
$$\leq \frac{M-\sum\limits_{j=1}^{l} t_{ij}x_j}{M-m}\varphi(m) + \frac{\sum\limits_{j=1}^{l} t_{ij}x_j - m}{M-m}\varphi(M).$$

Further, multiplying with v_i and summing over $i=1,...,k$, we get

$$\sum_{i=1}^{k}\sum_{j=1}^{l} v_i t_{ij}\varphi(x_j)$$
$$+\frac{1}{M-m}\sum_{i=1}^{k}\sum_{j=1}^{l} v_i t_{ij}\left[(M-x_j)\varphi(x_j-m)+(x_j-m)\varphi(M-x_j)\right]$$
$$\leq \frac{M\sum\limits_{i=1}^{k} v_i - \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{l} v_i t_{ij}x_j}{M-m}\varphi(m) + \frac{\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{l} v_i t_{ij}x_j - m\sum\limits_{i=1}^{k} v_i}{M-m}\varphi(M),$$

which is equivalent to (2.248).

Therefore, if $\varphi \geq 0$, then from Lemma 2.10.6. it follows that φ is convex. The first inequality in (2.250) is Sherman's inequality which holds under assumptions of theorem. Others inequalities in (2.250) follow from (2.248) since $\Omega \geq 0$. □

Remark 2.10.32.: In the previous theorem we got series of inequalities that include Sherman's inequality and its converse. This result (2.250) improves (2.11).

Remark 2.10.33.: If φ is subquadratic, then the reverse inequality in (2.248) holds.

Choosing $k=l$ and $\mathbf{v}=\mathbf{e}=(1,...,1)$ in Theorem 2.10.47, we get the following corollary.

Corollary 2.10.17. *([64])*: *Let $\mathbf{x}=(x_1,...,x_k)$, $\mathbf{y}=(y_1,...,y_k)$ be vectors with entries from $[m,M]$ $(0 \leq m < M)$ such that $\mathbf{y} \prec \mathbf{x}$.*

If $\varphi:[0,\infty)\to\mathbb{R}$ is a superquadratic function, then

$$\sum_{i=1}^{k}\varphi(x_i)+\Lambda \leq \frac{kM-\sum\limits_{i=1}^{k} y_i}{M-m}\varphi(m)+\frac{\sum\limits_{i=1}^{k} y_i - km}{M-m}\varphi(M)$$

where

$$\Lambda = \frac{1}{M-m}\sum_{i=1}^{k}\left[(M-x_i)\,\varphi\,(x_i-m)+(x_i-m)\,\varphi\,(M-x_i)\right].$$

If in addition $\varphi \geq 0$, then φ is convex and the following series of inequalities hold

$$\begin{aligned}\sum_{i=1}^{k}\varphi\,(y_i) &\leq \sum_{i=1}^{k}\varphi\,(x_i) \qquad (2.251)\\ &\leq \frac{kM-\sum\limits_{i=1}^{k}y_i}{M-m}\varphi(m)+\frac{\sum\limits_{i=1}^{k}y_i-km}{M-m}\varphi(M)-\Lambda\\ &\leq \frac{kM-\sum\limits_{i=1}^{k}y_i}{M-m}\varphi(m)+\frac{\sum\limits_{i=1}^{k}y_i-km}{M-m}\varphi(M).\end{aligned}$$

Remark 2.10.34.: In the previous corollary we obtained series of inequalities that include Majorization inequality and its converse.. This result improves (2.12).

CHAPTER 3.

On Sherman's inequality for n-Convex Functions

Through this chapter, by using different classes of interpolation polynomials in combination with nice properties of Green's functions, which we introduce in the sequel, we obtain extensions of Sherman's inequality (2.10) on the class of convex functions of the higher order which are in special case convex in usual sense (see Definition 1.1.6.). We also obtain extension of Majorization inequality (see Theorem 2.1.1 and Theorem 2.1.3). Using results for the Čebyšev functionals (Theorem 1.4.12 and Theorem 1.4.13) we derive related generalized inequalities of the Grüss type. We also present generalized Ostrowski type inequalities. Obtained generalizations we use for constructing new linear functionals which enables us to present results in the form of mean value theorems. Applying the exponential convexity method, established in [77], we present several families of exponentially convex functions that are an essential ingredient of new Cauchy type means closely related to Stolarsky means. Constructed exponentially convex functions are added as new non-trivial examples that sparse examples of exponentially convex functions since invention of exponential convexity back to 1929. Material of this chapter is quoted mainly from [6], [7], [8], [9], [10], [66], [67], [68], [69], [73], [74].

3.1. Sherman's inequality and Taylor's formula

The techniques that are used in this section are based on the classical real analysis and an application of Taylor's formula with the integral remainder which we introduce in the sequel. The results from this section are quoted from [67].

Through this section we always assume that $[\alpha, \beta] \subset \mathbb{R}$ unless otherwise stated.

Theorem 3.1.1: *Let $n \in \mathbb{N}$, $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$ and $c \in [\alpha, \beta]$. Then for all $x \in [\alpha, \beta]$ we have*

$$f(x) = T_{n-1}(f; c, x) + R_{n-1}(f; c, x), \tag{3.1}$$

where $T_{n-1}(f; c, x)$ is a Taylor's polynomial of degree $n-1$, i.e.

$$T_{n-1}(f; c, x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k \tag{3.2}$$

and the remainder is given by

$$R_{n-1}(f; c, x) = \frac{1}{(n-1)!} \int_c^x f^{(n)}(t)(x-t)^{n-1} dt. \tag{3.3}$$

In literature the previous theorem is known as **Taylor's theorem** or **Taylor's formula** with the integral remainder.

Let us consider the real valued function

$$x \mapsto (x-t)_+ = \begin{cases} x-t, & t \leq x \\ 0, & t > x \end{cases}. \tag{3.4}$$

If we apply Taylor's formula (3.1), i.e.

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k + \frac{1}{(n-1)!} \int_c^x f^{(n)}(t)(x-t)^{n-1} dt \tag{3.5}$$

at the point $c = \alpha$, we have

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k + \frac{1}{(n-1)!}\int_\alpha^\beta ((x-t)_+)^{n-1} f^{(n)}(t)dt, \tag{3.6}$$

where

$$\int_\alpha^\beta ((x-t)_+)^{n-1} f^{(n)}(t)dt = \int_\alpha^x (x-t)^{n-1} f^{(n)}(t)dt + \int_x^\beta 0dt. \tag{3.7}$$

Similarly, applying Taylor's formula (3.5) to function f at the point $c = \beta$, we have

$$f(x) = \sum_{k=0}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!}(\beta - x)^k - \frac{(-1)^{n-1}}{(n-1)!}\int_\alpha^\beta ((t-x)_+)^{n-1} f^{(n)}(t)dt, \tag{3.8}$$

where

$$\int_\alpha^\beta ((t-x)_+)^{n-1} f^{(n)}(t)dt = \int_\alpha^x 0dt + \int_x^\beta (t-x)^{n-1} f^{(n)}(t)dt. \tag{3.9}$$

3.1.1. Generalizations of Sherman's inequality by Taylor's formula

Using Taylor's formula with the previous announcement we introduce new identities related to Sherman's inequality (2.10) which will be very helpful to obtain generalizations of Sherman's theorem (Theorem 2.2.4) to the class of n-convex functions.

Theorem 3.1.2: *Let $n \in \mathbb{N}$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$. Let $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \text{and} \quad \mathbf{a} = \mathbf{b}\mathbf{A} \tag{3.10}$$

holds for some matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^{l} a_{ij} = 1, i = 1, 2, ..., m$. Then the following identities hold:

(i)

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (3.11)\\
&= \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j-\alpha)^k - \sum_{i=1}^{m} b_i (y_i-\alpha)^k \right]\\
&+ \frac{1}{(n-1)!} \int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j \left((x_j-t)_+\right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i-t)_+\right)^{n-1} \right] f^{(n)}(t) d\ ;
\end{aligned}$$

(ii)

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (3.12)\\
&= \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} \left[\sum_{j=1}^{l} a_j (\beta-x_j)^k - \sum_{i=1}^{m} b_i (\beta-y_i)^k \right]\\
&- \frac{(-1)^{n-1}}{(n-1)!} \int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j ((t-x_j)_+)^{n-1} - \sum_{i=1}^{m} b_i \left((t-y_i)_+\right)^{n-1} \right] f^{(n)}(t) dt.
\end{aligned}$$

Proof: (i) Using (3.6) in the difference

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i), \qquad (3.13)$$

we get

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)\\
&= \sum_{j=1}^{l} a_j \left[\sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} (x_j-\alpha)^k + \frac{1}{(n-1)!} \int_{\alpha}^{\beta} ((x_j-t)_+)^{n-1} f^{(n)}(t) dt \right]\\
&- \sum_{i=1}^{m} b_i \left[\sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} (y_i-\alpha)^k + \frac{1}{(n-1)!} \int_{\alpha}^{\beta} ((y_i-t)_+)^{n-1n-1} f^{(n)}(t) dt \right].
\end{aligned}$$

Now, by interchanging the order of summation and integration, we get

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \right]$$
$$+ \frac{1}{(n-1)!} \int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j \left((x_j - t)_+\right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i - t)_+\right)^{n-1} \right] f^{(n)}(t) dt.$$

Since under the assumption (3.10) we have

$$\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k = 0 \quad \text{for} \quad k = 0, 1,$$

the identity (3.11) immediately follows.

(ii) Similarly, using (3.8) in Sherman's difference (3.13), with analogous argument as in the previous case, we obtain (3.12). □

Note that in the previous theorem we assume that weights $\mathbf{a}$, $\mathbf{b}$ and matrix $\mathbf{A}$ are with real entries, not necessary nonnegative as in Sherman's theorem (Theorem 2.2.4).

Using the previous identities we get the following extension of Sherman's theorem to n-convex functions.

Theorem 3.1.3: *Suppose that all the assumptions of Theorem 3.1.2 hold.*

(i) If f is n-convex and

$$\sum_{j=1}^{l} a_j \left((x_j - t)_+\right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i - t)_+\right)^{n-1} \geq 0, \quad t \in [\alpha, \beta], \tag{3.14}$$

then

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \geq \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \right]. \tag{3.15}$$

(ii) If f is n-convex and

$$(-1)^{n-1} \left[\sum_{j=1}^{l} a_j ((t - x_j)_+)^{n-1} - \sum_{i=1}^{m} b_i \left((t - y_i)_+\right)^{n-1} \right] \leq 0, \quad t \in [\alpha, \beta], \tag{3.16}$$

then

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \geq \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} \left[\sum_{j=1}^{l} a_j (\beta - x_j)^k - \sum_{i=1}^{m} b_i (\beta - y_i)^k \right]. \tag{3.17}$$

Proof: (i) By Theorem 3.1.2 the identity (3.11) holds. Since f is n-convex on $[\alpha, \beta]$, we may assume without loss of generality that f is n-times differentiable and $f^{(n)} \geq 0$ on $[\alpha, \beta]$ (see [138, p. 16]). Moreover, if (3.14) holds, it follows that the last term in (3.11) is nonnegative, i.e. the inequality (3.15) holds.

(ii) Analogous to (i). □

When we take into account Sherman's condition of nonnegativity, i.e. if we suppose that weights **a**, **b** and matrix **A** are nonnegative, then the assumption (3.14) is equivalent to the requirement that the function $x \mapsto (x-t)_+^{n-1}$, $t \in [\alpha, \beta]$, must be convex. Notice that $x \mapsto ((x-t)_+)^{n-1}$, $t \in [\alpha, \beta]$, is convex for every $n \in \mathbb{N}$ and then by Theorem 3.1.3, the inequality (3.15) holds. Specially, for $n = 2$, the sum on the right hand side of (3.15) is zero and we get

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \geq 0,$$

i.e. we obtain Sherman's inequality as direct consequence.

Moreover, the following generalization hold.

Theorem 3.1.4: *Suppose that all the assumptions of Theorem 3.1.2 hold. Additionally, let* **a**, **b** *and* **A** *be nonnegative.*

(i) If f is n-convex, then the inequality (3.15) holds. Moreover, if the function

$$\tilde{F}(\cdot) = \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} (\cdot - \alpha)^k \tag{3.18}$$

is convex on $[\alpha, \beta]$*, then*

$$0 \leq \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \right] \leq \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i). \tag{3.19}$$

(ii) If n *is even and* f *is* n*-convex, then the inequality (3.17) holds. Moreover, if the function*

$$\bar{F}(\cdot) = \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} (\beta - \cdot)^k \tag{3.20}$$

is convex on $[\alpha, \beta]$*, then*

$$0 \leq \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} \left[\sum_{j=1}^{l} a_j (\beta - x_j)^k - \sum_{i=1}^{m} b_i (\beta - y_i)^k \right] \leq \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i). \tag{3.21}$$

Proof: (i) Since the function $x \mapsto ((x-t)_+)^{n-1}$, $t \in [\alpha, \beta]$, is convex for every $n \in \mathbb{N}$, then by Sherman's theorem we have

$$\sum_{j=1}^{l} a_j \left((x_j - t)_+\right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i - t)_+\right)^{n-1} \geq 0, \quad t \in [\alpha, \beta],$$

i.e. the condition (3.14) is satisfied. Moreover, by Theorem 3.1.3 the inequality (3.15) holds. Further, the right hand side of (3.15) can be written in the form

$$\sum_{j=1}^{l} a_j \tilde{F}(x_j) - \sum_{i=1}^{m} b_i \tilde{F}(y_i),$$

where $\tilde{F}$ is defined as in (3.18). If $\tilde{F}$ is convex, then by Sherman's theorem we have

$$\sum_{j=1}^{l} a_j \tilde{F}(x_j) - \sum_{i=1}^{m} b_i \tilde{F}(y_i) \geq 0,$$

i.e. the right hand side of the inequality (3.15) is nonnegative, so (3.19) holds.

(ii) Since the function

$$y \mapsto (-1)^n \left((t-y)_+\right)^n = \begin{cases} 0, & t < y \\ (y-t)^n, & t \geq y \end{cases}, \quad t \in [\alpha, \beta],$$

is convex for even n and concave for odd n, then the function $y \mapsto (-1)^{n-1}\left((t-y)_+\right)^{n-1}$ is concave for even n and applying Sherman's theorem we have

$$(-1)^{n-1}\left[\sum_{j=1}^{l} a_j((t-x_j)_+)^{n-1} - \sum_{i=1}^{m} b_i\left((t-y_i)_+\right)^{n-1}\right] \leq 0, \quad t \in [\alpha, \beta],$$

i.e. the condition (3.16) is satisfied. Now, applying Theorem 3.1.3 we get (3.17). Further, the right hand side of (3.17) can be written in the form

$$\sum_{j=1}^{l} a_j \bar{F}(x_j) - \sum_{i=1}^{m} b_i \bar{F}(y_i),$$

where $\bar{F}$ is defined as in (3.20). The rest of proof is analogous to (i). □

By the previous theorem, under Sherman's condition of nonnegativity, the inequality (3.15) holds. Since the function $x \mapsto (x-\alpha)^k$ is convex on $[\alpha, \beta]$, then

$$\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \geq 0.$$

Further, if $f^{(k)}(\alpha) \geq 0$, $k = 2, ..., n-1$, then the right hand side of (3.15) is nonnegative and again double inequality (3.19) holds. Notice that (3.19) represents new lower bounds for Sherman's difference $\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)$.

If n is even, then the inequality (3.17) holds. So, if $f^{(k)}(\beta) \geq 0$ for even k and $f^{(k)}(\beta) \leq 0$ for odd k, then (3.21) holds, i.e. we get the lower bound in another form.

3.1.2. The Ostrowski and Grüss type inequalities

In the previous subsection we get the lower bounds for the difference (3.13) in the forms (3.19) and (3.21). In this subsection we give upper bounds related to obtained generalizations.

To simplify the notation of the following results, we define the functions $\mathcal{B}_1, \mathcal{B}_2 : [\alpha, \beta] \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{B}_1(t) &= \sum_{j=1}^{l} a_j \left((x_j - t)_+\right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i - t)_+\right)^{n-1}, \qquad (3.22)\\ \mathcal{B}_2(t) &= (-1)^{n-1}\left[\sum_{j=1}^{l} a_j((t - x_j)_+)^{n-1} - \sum_{i=1}^{m} b_i \left((t - y_i)_+\right)^{n-1}\right],\end{aligned}$$

where $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ are such as in Theorem 3.1.2, i.e. such that (3.10) holds for some matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^{l} a_{ij} = 1, i = 1, 2, ..., m$.

For such functions, we consider the Čebyšev functionals defined by

$$T(\mathcal{B}_p, \mathcal{B}_p) = \frac{1}{\beta - \alpha}\int_{\alpha}^{\beta} \mathcal{B}_p^2(t)dt - \left(\frac{1}{\beta - \alpha}\int_{\alpha}^{\beta} \mathcal{B}_p(t)dt\right)^2, \quad p = 1, 2.$$

Theorem 3.1.5: *Let $n \in \mathbb{N}$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n)} \in AC([\alpha, \beta])$ with $(\cdot - \alpha)(\beta - \cdot)(f^{(n+1)})^2 \in L_1[\alpha, \beta]$. Let $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that (3.10) holds for some matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^{l} a_{ij} = 1, i = 1, 2, ..., m$. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be defined as in (3.22), respectively.*

(i) The remainder $\kappa_n^1(f; \alpha, \beta)$, defined by

$$\begin{aligned}\kappa_n^1(f; \alpha, \beta) \ &= \ \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)\\ &\quad - \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!}\left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k\right]\\ &\quad - \frac{f^{(n-1)}(\beta) - f^{(n-1)}(\alpha)}{(\beta - \alpha)(n-1)!}\int_{\alpha}^{\beta} \mathcal{B}_1(t)dt \qquad (3.23)\end{aligned}$$

satisfies the estimation

$$\left|\kappa_n^1(f; \alpha, \beta)\right| \le \frac{\sqrt{\beta - \alpha}}{\sqrt{2}(n-1)!}[T(\mathcal{B}_1, \mathcal{B}_1)]^{\frac{1}{2}}\left(\int_{\alpha}^{\beta} (t - \alpha)(\beta - t)[f^{(n+1)}(t)]^2 dt\right)^{\frac{1}{2}}. \qquad (3.24)$$

(ii) The remainder $\kappa_n^2(f;\alpha,\beta)$, defined by

$$\begin{aligned}\kappa_n^2(f;\alpha,\beta) &= \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \\ &\quad - \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} \left[\sum_{j=1}^{l} a_j (\beta - x_j)^k - \sum_{i=1}^{m} b_i (\beta - y_i)^k \right] \\ &\quad - \frac{f^{(n-1)}(\beta) - f^{(n-1)}(\alpha)}{(\alpha-\beta)(n-1)!} \int_\alpha^\beta \mathcal{B}_2(t)dt \end{aligned} \tag{3.25}$$

satisfies the estimation

$$\left|\kappa_n^2(f;\alpha,\beta)\right| \leq \frac{\sqrt{\beta-\alpha}}{\sqrt{2}(n-1)!}[T(\mathcal{B}_2,\mathcal{B}_2)]^{\frac{1}{2}} \left(\int_\alpha^\beta (t-\alpha)(\beta-t)[f^{(n+1)}(t)]^2 dt \right)^{\frac{1}{2}}. \tag{3.26}$$

Proof: (i) Comparing (3.11) and (3.23) we have

$$\begin{aligned}\kappa_n^1(f;\alpha,\beta) &= \int_a^b \mathcal{B}_1(t) f^{(n)}(t)dt - \frac{f^{(n-1)}(\beta) - f^{(n-1)}(\alpha)}{\beta-\alpha} \int_a^b \mathcal{B}_1(t)dt \\ &= \int_a^b \mathcal{B}_1(t) f^{(n)}(t)dt - \frac{1}{\beta-\alpha} \int_a^b f^{(n)}(t)dt \int_a^b \mathcal{B}_1(t)dt \\ &= (\beta-\alpha)T(\mathcal{B}_1, f^{(n)}).\end{aligned}$$

Applying Theorem 1.4.12 on the function $\mathcal{B}_1$ and $f^{(n)}$ we obtain (3.24).

(ii) Similarly, we can prove this part. □

The following Grüss type inequalities can be obtained by using Theorem 1.4.13.

Theorem 3.1.6: *Let $n \in \mathbb{N}$ and $f : [\alpha,\beta] \to \mathbb{R}$ be such that $f^{(n)} \in AC([\alpha,\beta])$ and $f^{(n+1)} \geq 0$ on $[\alpha,\beta]$. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be defined as in (3.22), respectively.*

(i) Then the remainder $\kappa_n^1(f;\alpha,\beta)$ defined by (3.23) satisfies the estimation

$$\begin{aligned}&\left|\kappa_n^1(f;\alpha,\beta)\right| \\ &\leq \frac{\beta-\alpha}{(n-1)!} \left\|\mathcal{B}_1'\right\|_\infty \left[\frac{f^{(n-1)}(\beta) + f^{(n-1)}(\alpha)}{2} - \frac{f^{(n-2)}(\beta) - f^{(n-2)}(\alpha)}{\beta-\alpha} \right].\end{aligned} \tag{3.27}$$

(ii) Then the remainder $\kappa_n^2(f;\alpha,\beta)$ defined by (3.25) satisfies the estimation

$$\begin{aligned}&\left|\kappa_n^2(f;\alpha,\beta)\right| \\ &\leq \frac{\beta-\alpha}{(n-1)!}\left\|\mathcal{B}_2'\right\|_\infty\left[\frac{f^{(n-1)}(\beta)+f^{(n-1)}(\alpha)}{2}-\frac{f^{(n-2)}(\beta)-f^{(n-2)}(\alpha)}{\beta-\alpha}\right].\end{aligned} \tag{3.28}$$

Proof: (i) Since

$$\kappa_n^1(f;\alpha,\beta)=(\beta-\alpha)T(\mathcal{B}_1,f^{(n)}),$$

applying Theorem 1.4.13 on the function $\mathcal{B}_1$ and $f^{(n)}$ we obtain (3.27).

(ii) Similarly, we can prove this part. □

As direct consequences of the previous theorems, choosing $n=2$, we obtain the following two corollaries.

Corollary 3.1.1.: *Let $f:[\alpha,\beta]\to\mathbb{R}$ be such that $f''\in AC([\alpha,\beta])$ and $(\cdot-\alpha)(\beta-\cdot)(f''')^2\in L_1[\alpha,\beta]$. Let $\mathbf{x}\in[\alpha,\beta]^l$, $\mathbf{y}\in[\alpha,\beta]^m$, $\mathbf{a}\in\mathbb{R}^l$ and $\mathbf{b}\in\mathbb{R}^m$ be such that (3.10) holds for some matrix $\mathbf{A}\in M_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^l a_{ij}=1, i=1,2,...,m$.*

(i) Then the remainder $\kappa^1(f;\alpha,\beta)$ defined by

$$\begin{aligned}\kappa^1(f;\alpha,\beta)=&\sum_{j=1}^l a_jf(x_j)-\sum_{i=1}^m b_if(y_i)\\ &-\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta\left(\sum_{j=1}^l a_j(x_j-t)_+-\sum_{i=1}^m b_i(y_i-t)_+\right)dt\end{aligned} \tag{3.29}$$

satisfies the estimation

$$\left|\kappa^1(f;\alpha,\beta)\right|\leq\sqrt{\frac{\beta-\alpha}{2}}[T(\mathcal{B},\mathcal{B})]^{\frac{1}{2}}\left(\int_\alpha^\beta(t-\alpha)(\beta-t)[f'''(t)]^2dt\right)^{\frac{1}{2}},$$

where

$$\begin{aligned}T(\mathcal{B},\mathcal{B})=&\frac{1}{\beta-\alpha}\int_\alpha^\beta\left(\sum_{j=1}^l a_j(x_j-t)_+-\sum_{i=1}^m b_i(y_i-t)_+\right)^2dt\\ &-\left(\frac{1}{\beta-\alpha}\int_\alpha^\beta\left(\sum_{j=1}^l a_j(x_j-t)_+-\sum_{i=1}^m b_i(y_i-t)_+\right)dt\right)^2.\end{aligned}$$

(ii) Then the remainder $\kappa^2(f;\alpha,\beta)$ defined by

$$\kappa^2(f;\alpha,\beta) = \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) - \frac{f'(\beta)-f'(\alpha)}{\alpha-\beta}\int_\alpha^\beta \left(\sum_{i=1}^{m} b_i(t-y_i)_+ - \sum_{j=1}^{l} a_j(t-x_j)_+\right) dt \tag{3.30}$$

satisfies the estimation

$$\left|\kappa^2(f;\alpha,\beta)\right| \leq \sqrt{\frac{\beta-\alpha}{2}}[T(\mathcal{R},\mathcal{R})]^{\frac{1}{2}} \left(\int_\alpha^\beta (t-\alpha)(\beta-t)[f'''(t)]^2 dt\right)^{\frac{1}{2}},$$

where

$$T(\mathcal{R},\mathcal{R}) = \frac{1}{\beta-\alpha}\int_\alpha^\beta \left(\sum_{i=1}^{m} b_i(t-y_i)_+ - \sum_{j=1}^{l} a_j(t-x_j)_+\right)^2 dt - \left(\frac{1}{\beta-\alpha}\int_\alpha^\beta \left(\sum_{i=1}^{m} b_i(t-y_i)_+ - \sum_{j=1}^{l} a_j(t-x_j)_+\right) dt\right)^2.$$

Corollary 3.1.2.: *Let $f : [\alpha,\beta] \to \mathbb{R}$ be such that $f'' \in AC([\alpha,\beta])$ and $f''' \geq 0$ on $[\alpha,\beta]$. Let $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that (3.10) holds for some matrix $\mathbf{A} \in M_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^{l} a_{ij} = 1, i = 1,2,...,m$.*

(i) Then the remainder $\kappa^1(f;\alpha,\beta)$ defined by (3.29) satisfies the estimation

$$\left|\kappa^1(f;\alpha,\beta)\right| \leq (\beta-\alpha)\left\|\mathcal{B}'\right\|_\infty \left[\frac{f'(\beta)+f'(\alpha)}{2} - \frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right],$$

where

$$\mathcal{B}(t) = \sum_{j=1}^{l} a_j(x_j-t)_+ - \sum_{i=1}^{m} b_i(y_i-t)_+.$$

(ii) Then the remainder $\kappa^2(f;\alpha,\beta)$ defined by (3.30) satisfies the estimation

$$\left|\kappa^2(f;\alpha,\beta)\right| \leq (\beta-\alpha)\left\|\mathcal{R}'\right\|_\infty \left[\frac{f'(\beta)+f'(\alpha)}{2} - \frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right],$$

where

$$\mathcal{R}(t) = \sum_{i=1}^{m} b_i (t - y_i)_+ - \sum_{j=1}^{l} a_j (t - x_j)_+.$$

Now we give related inequalities of the Ostrowski type.

Theorem 3.1.7: *Suppose that all the assumptions of Theorem 3.1.2 hold. Let (p, q) be a pair of conjugate exponents, i.e. $1 \le p, q \le \infty$, $1/p + 1/q = 1$ and $f^{(n)} \in L_p[\alpha, \beta]$. Then*

(i)

$$\left| \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) - \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \right] \right| \le \frac{1}{(n-1)!} \left\| f^{(n)} \right\|_p \|\mathcal{B}_1\|_q ; \tag{3.31}$$

(ii)

$$\left| \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) - \sum_{k=2}^{n-1} \frac{(-1)^k f^{(k)}(\beta)}{k!} \left[\sum_{j=1}^{l} a_j (\beta - x_j)^k - \sum_{i=1}^{m} b_i (\beta - y_i)^k \right] \right| \le \frac{1}{(n-1)!} \left\| f^{(n)} \right\|_p \|\mathcal{B}_2\|_q . \tag{3.32}$$

The constants on the right-hand side of (3.31) and (3.32) are sharp for $1 < p \le \infty$ and the best possible for $p = 1$.

Proof: (i) Applying the well known Hölder inequality to (3.11), we obtain

$$\left| \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) - \sum_{k=2}^{n-1} \frac{f^{(k)}(\alpha)}{k!} \left[\sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k \right] \right| = \left| \int_\alpha^\beta \mathcal{S}(t) f^{(n)}(t) dt \right| \le \left\| f^{(n)} \right\|_p \left(\int_\alpha^\beta |\mathcal{S}(t)|^q \, dt \right)^{\frac{1}{q}},$$

where we denote

$$\mathcal{S}(t) = \frac{1}{(n-1)!} \left(\sum_{j=1}^{l} a_j \left((x_j - t)_+ \right)^{n-1} - \sum_{i=1}^{m} b_i \left((y_i - t)_+ \right)^{n-1} dt \right) = \frac{1}{(n-1)!} \mathcal{B}_1(t).$$

For the proof of the sharpness of the constant $\left(\int_\alpha^\beta |\mathcal{S}(t)|^q \, dt\right)^{\frac{1}{q}}$, let us find a function f for which the equality in (3.31) is obtained.

For $1 < p < \infty$ take f to be such that

$$f^{(n)}(t) = sgn\mathcal{S}(t)|\mathcal{S}(t)|^{\frac{1}{p-1}}.$$

For $p = \infty$ take $f^{(n)}(t) = sgn\mathcal{S}(t)$.

For $p = 1$ we prove that

$$\left|\int_\alpha^\beta \mathcal{S}(t) f^{(n)}(t) dt\right| \leq \max_{t\in[\alpha,\beta]} |\mathcal{S}(t)| \left(\int_\alpha^\beta \left|f^{(n)}(t)\right| dt\right) \tag{3.33}$$

is the best possible inequality.

Suppose that $|\mathcal{S}(t)|$ attains its maximum at $t_0 \in [\alpha, \beta]$.

First we assume $\mathcal{S}(t_0) > 0$. For ε small enough we define $f_\varepsilon(t)$ by

$$f_\varepsilon(t) = \begin{cases} 0, & \alpha \leq t \leq t_0 \\ \frac{1}{\varepsilon n!}(t - t_0)^n, & t_0 \leq t \leq t_0 + \varepsilon \\ \frac{1}{n!}(t - t_0)^{n-1}, & t_0 + \varepsilon \leq t \leq \beta \end{cases} .$$

Then for ε small enough

$$\left|\int_\alpha^\beta \mathcal{S}(t) f^{(n)}(t) dt\right| = \left|\int_{t_0}^{t_0+\varepsilon} \mathcal{S}(t) \frac{1}{\varepsilon} dt\right| = \frac{1}{\varepsilon} \int_{t_0}^{t_0+\varepsilon} \mathcal{S}(t) dt.$$

Now from the inequality (3.33) we have

$$\frac{1}{\varepsilon} \int_{t_0}^{t_0+\varepsilon} \mathcal{S}(t) dt \leq \mathcal{S}(t_0) \int_{t_0}^{t_0+\varepsilon} \frac{1}{\varepsilon} dt = \mathcal{S}(t_0).$$

Since

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_{t_0}^{t_0+\varepsilon} \mathcal{S}(t) dt = \mathcal{S}(t_0),$$

the statement follows.

In the case $\mathcal{S}(t_0) < 0$, we define $f_\varepsilon(t)$ by

$$f_\varepsilon(t) = \begin{cases} \frac{1}{n!}(t - t_0 - \varepsilon)^{n-1}, & \alpha \leq t \leq t \\ -\frac{1}{\varepsilon n!}(t - t_0 - \varepsilon)^n, & t_0 \leq t \leq t_0 + \varepsilon \\ 0, & t_0 + \varepsilon \leq t \leq \beta \end{cases},$$

and the rest of the proof is the same as above.

(ii) Similar to the first part. □

The following corollary is direct consequence of the previous theorem for $n = 2$.

Corollary 3.1.3.: *Let $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f' \in AC([\alpha, \beta])$. Let $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that (3.10) holds for some matrix $\mathbf{A} \in M_{ml}(\mathbb{R})$ satisfying the condition $\sum_{j=1}^{l} a_{ij} = 1, i = 1, 2, ..., m$. Let (p, q) be a pair of conjugate exponents, i.e. $1 \leq p, q \leq \infty$, $1/p + 1/q = 1$ and $f'' \in L_p[\alpha, \beta]$. Then*

(i)

$$\left| \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \right| \leq \|f''\|_p \left[\int_\alpha^\beta \left| \sum_{j=1}^{l} a_j (x_j - t)_+ - \sum_{i=1}^{m} b_i (y_i - t)_+ \right|^q dt \right]^{\frac{1}{q}}; \tag{3.34}$$

(ii)

$$\left| \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \right| \leq \|f''\|_p \left[\int_\alpha^\beta \left| \sum_{j=1}^{l} a_j (t - x_j)_+ - \sum_{i=1}^{m} b_i (t - y_i)_+ \right|^q dt \right]^{\frac{1}{q}}. \tag{3.35}$$

The constants on the right-hand side of (3.34) and (3.35) are sharp for $1 < p \leq \infty$ and the best possible for $p = 1$.

3.1.3. Mean value theorems and n-exponential convexity

Motivated by inequalities (3.15) and (3.17), we define some linear functionals as follows.

Definition 3.1.1.: Under the assumptions of Theorem 3.1.3, we define

$$A_1(f)=\sum_{j=1}^{l}a_jf(x_j)-\sum_{i=1}^{m}b_if(y_i)-\sum_{k=2}^{n-1}\frac{f^{(k)}(\alpha)}{k!}\left[\sum_{j=1}^{l}a_j(x_j-\alpha)^k-\sum_{i=1}^{m}b_i(y_i-\alpha)^k\right], \tag{3.36}$$

$$A_2(f)=\sum_{j=1}^{l}a_jf(x_j)-\sum_{i=1}^{m}b_if(y_i)-\sum_{k=2}^{n-1}\frac{(-1)^kf^{(k)}(\beta)}{k!}\left[\sum_{j=1}^{l}a_j(\beta-x_j)^k-\sum_{i=1}^{m}b_i(\beta-y_i)^k\right].$$

Remark 3.1.1.: It should be noticed that if $f:[\alpha,\beta]\to\mathbb{R}$ is n-convex, then by Theorem 3.1.3, it holds

$$A_p(f)\geq 0,\quad p=1,2.$$

Using given linear functionals we derive mean value theorems of the Lagrange and Cauchy type.

Theorem 3.1.8: *Let $f\in C^n([\alpha,\beta])$ and A_p, $p=1,2$, be the linear functionals defined as in (3.36). Then there exists $\xi_p\in[\alpha,\beta]$, $(p=1,2)$, such that*

$$A_p(f)=f^{(n)}(\xi_p)A_p(\varphi_0),\quad p=1,2, \tag{3.37}$$

where $\varphi_0(x)=x^n/n!$.

Proof: Since $f\in C^n([\alpha,\beta])$, then there exist $m=\min_{x\in[\alpha,\beta]}f^{(n)}(x)$ and $M=\max_{x\in[\alpha,\beta]}f^{(n)}(x)$ such that $m\leq f^{(n)}(x)\leq M$ for each $x\in[\alpha,\beta]$.

We define the functions $\varphi_1,\varphi_2:[\alpha,\beta]\to\mathbb{R}$ by

$$\varphi_1(x)=\tfrac{Mx^n}{n!}-f(x)=M\varphi_0(x)-f(x)\quad\text{and}\quad\varphi_2(x)=f(x)-\tfrac{mx^n}{n!}=f(x)-m\varphi_0(x).$$

Then

$$\varphi_1^{(n)}(x)=M-f^{(n)}(x)\geq 0\quad\text{and}\quad\varphi_2^{(n)}(x)=f^{(n)}(x)-m\geq 0,$$

so the functions φ_1,φ_2 are n-convex.

Hence, $A_i(\varphi_1) \geq 0$ and $A_i(\varphi_2) \geq 0$, i.e. we have

$$MA_i(\varphi_0) \geq A_i(f) \quad \text{and} \quad mA_i(\varphi_0) \leq A_i(f), \qquad \text{for} \quad i = 1, 2.$$

Combining the last two inequalities we obtain

$$mA_i(\varphi_0) \leq A_i(f) \leq MA_i(\varphi_0), \quad \text{for} \ \ i = 1, 2.$$

Then there exists $\xi_i \in [\alpha, \beta]$ such that $A_i(f) = f^{(n)}(\ \xi_i)A_i(\varphi_0)$, $i = 1, 2$, holds. $\square$

Theorem 3.1.9: *Let $f, g \in C^n([\alpha, \beta])$ and A_p, $(p = 1, 2)$, be the linear functionals defined as in (3.36). Then there exists $\xi_p \in [\alpha, \beta]$, $p = 1, 2$, such that*

$$\frac{f^{(n)}(\xi_p)}{g^{(n)}(\xi_p)} = \frac{A_p(f)}{A_p(g)}, \quad p = 1, 2, \tag{3.38}$$

assuming that both denominators are non-zero.

Proof: Define the functions $h_i : [\alpha, \beta] \to \mathbb{R}$, $i = 1, 2$, by

$$h_i(x) = A_i(g)f(x) - A_i(f)g(x).$$

Then $h_i \in C^n([\alpha, \beta])$ for each $i = 1, 2$.

By the previous theorem, it follows that there exists $\xi_i \in [\alpha, \beta]$ such that

$$A_i(h_i) = h_i^{(n)}(\xi_i)A_i(\varphi_0), \quad i = 1, 2,$$

where $\varphi_0(x) = x^n/n!$.

Since

$$A_i(h_i) = A_i(g)A_i(f) - A_i(f)A_i(g) = 0$$

and

$$h_i^{(n)}(\xi_i)A_i(\varphi_0) = \left[A_i(g)f^{(n)}(\xi_i) - A_i(f)g^{(n)}(\xi_i)\right] A_i(\varphi_0)$$

for each $i = 1, 2$, it follows that

$$A_i(g)f^{(n)}(\xi_i) - A_i(f)g^{(n)}(\xi_i) = 0,$$

i.e.

$$\frac{f^{(n)}(\xi_i)}{g^{(n)}(\xi_i)} = \frac{A_i(f)}{A_i(g)}, \quad i = 1, 2,$$

what we need to prove. □

Remark 3.1.2.: If $\frac{f^{(n)}}{g^{(n)}}$ is an invertible function, then from (3.38) we have

$$\xi_p = \left(\frac{f^{(n)}}{g^{(n)}}\right)^{-1} \left(\frac{A_p(f)}{A_p(g)}\right)$$

which presents a new mean of Cauchy's type of the interval $[\alpha, \beta]$.

Applying exponential convexity method, established in [77], we could previous results interpret in the form of exponentially convex functions or in a special case logarithmically convex functions. As outcome we define new classes of two-parameter Cauchy-type means.

In order to obtain results regarding the exponential convexity, we define the families of functions as follows.

Definition 3.1.2.: Let I be an open interval in $\mathbb{R}$. For every choice of $t+1$ mutually different points $z_0, z_1, ..., z_t \in [\alpha, \beta]$ we define:

- $\mathcal{F}_1 = \{f_v : [\alpha, \beta] \to \mathbb{R} : v \in I$ and $v \mapsto [z_0, z_1, ..., z_t; f_v]$ is n-exponentially convex in the Jensen sense on $I\}$
- $\mathcal{F}_2 = \{f_v : [\alpha, \beta] \to \mathbb{R} : v \in I$ and $v \mapsto [z_0, z_1, ..., z_t; f_v]$ is exponentially convex in the Jensen sense on $I\}$
- $\mathcal{F}_3 = \{f_v : [\alpha, \beta] \to \mathbb{R} : v \in I$ and $v \mapsto [z_0, z_1, ..., z_t; f_v]$ is 2-exponentially convex in the Jensen sense on $I\}$

Now we present theorem which gives non trivial examples of n-exponentially and exponentially convex functions including the positive linear functionals A_p $(p = 1, 2)$.

Theorem 3.1.10: *Let A_p, $(p = 1, 2)$, be the linear functionals defined by (3.36) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$).*

Then the following statements are valid for each $p = 1, 2$.

(i) The function $v \mapsto A_p(f_v)$ is n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) The matrix $\left[A_p\left(f_{\frac{t_j+t_l}{2}}\right)\right]_{j,l=1}^{m}$ is a positive semi-definite. Particularly,

$$\det\left[A_p\left(f_{\frac{t_j+t_l}{2}}\right)\right]_{j,l=1}^{m} \geq 0$$

for all $m \in \mathbb{N}$, $m \leq n$, and $t_1, .., t_m \in I$ (resp. for all $m \in \mathbb{N}$ and $t_1, .., t_m \in I$).

(iii) If the function $v \mapsto A_p(f_v)$ is continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I.

Proof: *(i) Let $p_j, s_j \in \mathbb{R}$, $j = 1, ..., n$, and $s_{jk} = \frac{s_j+s_k}{2}$, $1 \leq j, k \leq n$. We consider the function $h : [\alpha, \beta] \to \mathbb{R}$ defined by*

$$h(x) = \sum_{j,k=1}^{n} p_j p_k f_{s_{jk}}(x),$$

where $f_{s_{jk}} \in \mathcal{F}_1$. It follows by definition that $h \in C([\alpha, \beta])$.

Since the function $t \mapsto [x_0, x_1, ..., x_l; f_t]$ is n-exponentially convex in the Jensen sense on I, we have

$$[x_0, x_1, ..., x_l; h] = \sum_{j,k=1}^{n} p_j p_k \left[x_0, x_1, ..., x_l; f_{s_{jk}}\right] \geq 0.$$

Hence, the function h is n-convex. Therefore, we have

$$A_i(h) = \sum_{j,k=1}^{n} p_j p_k A_i\left(f_{s_{jk}}\right) \geq 0,$$

for each $i = 1, 2$ and all choices of $p_j, s_j \in \mathbb{R}$.

Now we conclude that the function $t \mapsto A_i(f_t)$ is n-exponentially convex in the Jensen sense on I what we need to prove.

(ii) It follows from (i) and Definition 1.2.8..

(iii) Since $v \mapsto A_p(f_v)$ is n-exponentially convex (resp. exponentially convex) in the Jensen

sense on I and in addition continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I. □

Corollary 3.1.4.: *Let A_p, $p = 1, 2$, be the linear functional defined as in (3.36) associated with the family $\mathcal{F}_3$.*

Then the following statements are valid for each $p = 1, 2$.

(i) If the function $v \mapsto A_p(f_v)$ is continuous on I, then it is 2-exponentially convex on I. If $v \mapsto A_p(f_v)$ is additionally positive, then it is also log-convex on I. Furthermore, for every choice $r, s, t \in I$, such that $r < s < t$, it holds

$$[A_p(f_s)]^{t-r} \leq [A_p(f_r)]^{t-s} [A_p(f_t)]^{s-r}.$$

(ii) If the function $v \mapsto A_p(f_v)$ is positive and differentiable on I, then for all $r, s, u, w \in I$ such that $r \leq u$, $s \leq w$, we have

$$M_{r,s}(A_p, \mathcal{F}_3) \leq M_{u,w}(A_p, \mathcal{F}_3), \tag{3.39}$$

where

$$M_{r,s}(A_p, \mathcal{F}_3) = \begin{cases} \left(\frac{A_p(f_r)}{A_p(f_s)}\right)^{\frac{1}{r-s}}, & r \neq s \\ \exp\left(\frac{\frac{d}{dr}(A_p(f_r))}{A_p(f_r)}\right), & r = s \end{cases}, \tag{3.40}$$

for $f_r, f_s \in \mathcal{F}_3$.

Proof: (i) The first part of statement is an easy consequence of Theorem 3.1.10 and the second one of Remark 1.2.10..

(ii) From (i) we have that the function $t \mapsto A_p(f_t)$ is log-convex on I, i.e. the function $t \mapsto \log A_p(f_t)$ is convex on I. Applying Remark 1.1.1. b) we get

$$\frac{\log A_p(f_r) - \log A_p(f_s)}{r-s} \leq \frac{\log A_p(f_u) - \log A_p(f_v)}{u-v} \tag{3.41}$$

for $r \leq u$, $s \leq v$, $r \neq u$, $s \neq v$. Therefore, we have

$$M_{r,s}(A_p, \mathcal{F}_3) \leq M_{u,v}(A_p, \mathcal{F}_3).$$

Case $r = s$, $u = v$ follows from (3.41) as limiting case. □

Remark 3.1.3.: The property (3.39) is **monotonicity property** over both parameters.

Note that the results from Theorem 3.1.10 and Corollary 3.1.4. still hold when two of the points $z_0, z_1, ..., z_t \in [\alpha, \beta]$ coincide, say $z_0 = z_1$, for the family of differentiable functions f_v such that the function $v \mapsto [z_0, z_1, ..., z_t; f_v]$ is an n-exponentially convex in the Jensen sense (exponentially convex in the Jensen sense, log-convex in the Jensen sense), and furthermore, they still hold when all $(t+1)$ points coincide for a family of t differentiable functions with the same property. The proofs are obtained by the formula

$$[\underbrace{x, ..., x}_{j\text{-times}}; f] = \frac{f^{(j-1)}(x)}{(j-1)!} \tag{3.42}$$

and by a suitable characterization of convexity.

Example 1.: *Consider the family of functions*

$$\Omega = \{\varphi_u : [0, \infty) \to \mathbb{R} : u > 0\},$$

defined by

$$\varphi_u(x) = \begin{cases} \frac{x^u}{u(u-1)\cdots(u-n+1)}, & u \notin \{1, ..., n-1\}, \\ \frac{x^j \log x}{(-1)^{n-1-j} j!(n-1-j)!}, & u = j \in \{1, ..., n-1\}, \end{cases}$$

with $0 \log 0 := 0$.

Since $\frac{d^n \varphi_u}{dx^n}(x) = x^{u-n} = e^{(u-n)\log x} > 0$, *it follows that* φ_u *is* n*-convex on* $[0, \infty)$ *for every* $u > 0$. *The function* $u \mapsto \frac{d^n \varphi_u}{dx^n}(x)$ *is exponentially convex. Therefore, using the same arguments as in proof of Theorem 3.1.10, we conclude that the function* $u \mapsto [z_0, z_1, ..., z_t; \varphi_u]$ *is exponentially convex (and so exponentially convex in the Jensen sense). Also, it follows that* $u \mapsto A_1(\varphi_u)$ *is exponentially convex since it is exponentially convex in the Jensen sense and continuous. If* $[\alpha, \beta] \subset [0, \infty)$ *and* $A_1(\varphi_u) > 0$, *for this family of functions,* *(3.40) becomes*

$$M_{r,s}(A_1,\Omega)=\begin{cases}\left(\frac{s(s-1)\dots(s-n+1)}{r(r-1)\dots(r-n+1)}\frac{\sum_{j=1}^{l}a_jx_j^r-\sum_{i=1}^{m}b_iy_i^r}{\sum_{j=1}^{l}a_jx_j^s-\sum_{i=1}^{m}b_iy_i^s}\right)^{\frac{1}{r-s}}, & r\neq s,\\ \exp\left(\frac{\sum_{j=1}^{l}a_jx_j^r\log x_j-\sum_{i=1}^{m}b_iy_i^r\log y_i}{\sum_{j=1}^{l}a_jx_j^r-\sum_{i=1}^{m}b_iy_i^r}+\sum_{k=0}^{n-1}\frac{1}{k-r}\right), \\ r=s\notin\{1,\dots,n-1\}, \\ \exp\left(\frac{\sum_{j=1}^{l}a_jx_j^r\log^2 x_j-\sum_{i=1}^{m}b_iy_i^r\log^2 y_i}{2(\sum_{j=1}^{l}a_jx_j^r\log x_j-\sum_{i=1}^{m}b_iy_i^r\log y_i)}+\sum_{k=0,i\neq r}^{n-1}\frac{1}{k-r}\right), \\ r=s\in\{1,\dots,n-1\}.\end{cases} \tag{3.43}$$

For a different choice of parameters $r,s\in\mathbb{R}$, extensions are obtained by continuity.

To show that $M_{r,s}(A_1,\Omega)$ presents a mean of an interval $[\alpha,\beta]$, we apply Theorem 3.1.9 to $f=\varphi_r$ and $g=\varphi_s$, where $r,s\in\mathbb{R}$, $r\neq s$, $r,s\neq 0,1$. Setting $\alpha=\min_{i,j}\{x_j,y_i\}$ and $\beta=\max_{i,j}\{x_j,y_i\}$, by Theorem 3.1.9, there exists unique $\xi\in[\alpha,\beta]$ such that

$$\alpha\le\xi=\left(\frac{A_1(\varphi_r)}{A_1(\varphi_s)}\right)^{\frac{1}{r-s}}\le\beta.$$

Also, $M_{r,s}(A_1,\Omega)$ is obviously symmetric. The monotonicity properties over both parameters follows from (ii)-part of Corollary 3.1.4..

We obtain a new class of two-parameter Cauchy-type means.

Using the linear functional A_2, defined by (3.36), we can on analogous way derive another class of two-parameter Cauchy-type means with the same properties.

Example 2.: *Consider the family of functions*

$$\Omega=\{\varphi_u:[0,\infty)\to\mathbb{R}:u>0\},$$

defined by

$$\varphi_u(x)=\begin{cases}\frac{x^u}{u(u-1)\cdots(u-n+1)}, & u\notin\{1,\dots,n-1\},\\ \frac{x^j\log x}{(-1)^{n-1-j}j!(n-1-j)!}, & u=j\in\{1,\dots,n-1\},\end{cases}$$

with $0\log 0:=0$.

Since $\frac{d^n\varphi_u}{dx^n}(x)=x^{u-n}=e^{(u-n)\log x}>0$, $u\in\mathbb{R}$, it follows that φ_u is n-convex on $[0,\infty)$ for every $u\in\mathbb{R}$. The function $u\mapsto\frac{d^n\varphi_u}{dx^n}(x)$ is exponentially convex. Therefore, using *the same arguments as in proof of Theorem 3.1.10, we conclude that the function* $u \mapsto [z_0, z_1, ..., z_t; \varphi_u]$ *is exponentially convex (and so exponentially convex in the Jensen sense). It also follows that* $u \mapsto A_1(\varphi_u)$ *is exponentially convex since it is exponentially convex in the Jensen sense and continuous. If* $[\alpha, \beta] \subset [0, \infty)$ *and* $A_1(\varphi_u) > 0$, *for this family of functions, (3.40) becomes*

$$\begin{aligned}
M_{r,s} &= \left(\frac{s(s-1)...(s-n+1)}{r(r-1)...(r-n+1)} \frac{\sum_{j=1}^{l} a_j x_j^r - \sum_{i=1}^{m} b_i y_i^r - \sum_{k=2}^{n-1} \frac{r(r-1)...(r-k+1)}{k!} \alpha^{r-k} A_k}{\sum_{j=1}^{l} a_j x_j^s - \sum_{i=1}^{m} b_i y_i^s - \sum_{k=2}^{n-1} \frac{s(s-1)...(s-k+1)}{k!} \alpha^{s-k} A_k} \right)^{\frac{1}{r-s}}, \\
r &\neq s; \\
M_{r,r} &= exp\left(\frac{\sum_{j=1}^{l} a_j x_j^r \log x_j - \sum_{i=1}^{m} b_i y_i^r \log y_i - B_1}{\sum_{j=1}^{l} a_j x_j^r - \sum_{i=1}^{m} b_i y_i^r - B_2} + \sum_{k=0}^{n-1} \frac{1}{k-r} \right), \\
r &= s \notin \{1, ..., n-1\}; \\
M_{r,r} &= exp\left(\frac{\sum_{j=1}^{l} a_j x_j^r \log^2 x_j - \sum_{i=1}^{m} b_i y_i^r \log^2 y_i - B_3}{2\left(\sum_{j=1}^{l} a_j x_j^r \log x_j - \sum_{i=1}^{m} b_i y_i^r \log y_i - B_1\right)} + \sum_{k=0, k\neq r}^{n-1} \frac{1}{k-r} \right), \\
r &= s \in \{1, ..., n-1\},
\end{aligned}$$

where

$$\begin{aligned}
B_1 &= \sum_{k=2}^{n-1} \frac{1}{k!} A_k \frac{d^k}{dx^k} |x^r \log x|_{x=\alpha}, \\
B_2 &= \sum_{k=2}^{n-1} \frac{r(r-1)...(r-k+1)}{k!} \alpha^{r-k} A_k, \\
B_2 &= \sum_{k=2}^{n-1} \frac{r(r-1)...(r-k+1)}{k!} \alpha^{r-k} A_k, \\
A_k &= \sum_{j=1}^{l} a_j (x_j - \alpha)^k - \sum_{i=1}^{m} b_i (y_i - \alpha)^k.
\end{aligned}$$

For a different choice of parameters $r, s \in \mathbb{R}$, *extensions are obtained by continuity.*

As we explane in the previous example, in an analogous way we can show that we obtain a new class of two-parameter Cauchy-type means which are monotonic over both parameters. Using the linear functional A_2, *defined by (3.36), we can on analog way derive a new class of two-parameter Cauchy-type means* $M_{r,s}(A_2, \Omega)$ *with the same properties.*

3.2. Sherman's inequality and Taylor's formula with Green's functions

In this section, by combining Taylor's formula with Green's functions, which we define below, we derive new extensions of Sherman's inequality (2.10) for convex functions of a higher order. We obtain new generalized inequalities of the Grüss and Ostrowski type. We use the obtained generalizations to give results in the form of mean value theorems and to construct new nontrivial families of exponentially convex functions that enable us to derive new means of the Cauchy type. The results from this section are quoted from [73].

In the following results we consider the function $G_1 : [\alpha, \beta] \times [\alpha, \beta] \to \mathbb{R}$ defined by

$$G_1(u,s) = \begin{cases} \alpha - s, & s \le u, \\ \alpha - u, & u \le s. \end{cases} \tag{3.44}$$

It is Green's function of the boundary value problem

$$z'' = 0, \quad z(\alpha) = z'(\beta) = 0.$$

The function G_1 is continuous and convex in u, since it is symmetric, i.e. $G_1(u,s) = G_1(s,u)$, then also in s.

We also introduce four other type of Green's functions defined on $[\alpha, \beta] \times [\alpha, \beta]$ as follows:

$$G_2(u,s)=\begin{cases} u-\beta, & s\le u \\ s-\beta, & u\le s \end{cases}, \tag{3.45}$$

$$G_3(u,s)=\begin{cases} u-\alpha, & s\le u \\ s-\alpha, & u\le s \end{cases}, \tag{3.46}$$

$$G_4(u,s)=\begin{cases} \beta-s, & s\le u \\ \beta-u, & u\le s \end{cases}, \tag{3.47}$$

$$G_5(u,s)=\begin{cases} \frac{(u-\beta)(s-\alpha)}{\beta-\alpha}, & s\le u, \\ \frac{(s-\beta)(u-\alpha)}{\beta-\alpha}, & u\le s. \end{cases} \tag{3.48}$$

All Green's functions are continuous, symmetric and convex with respect to the both variables u and s. For more details see [157].

Next we present one technical lemma which gives us new identities involving defined Green's functions.

Lemma 3.2.1.: *Let $G_z(\cdot,s)$, $s\in[\alpha,\beta]$, $z=1,2,3,4,5$ be defined as in (3.44)-(3.48). Then for every $f\in C^2([\alpha,\beta])$, it holds:*

$$f(u)=f(\alpha)+(u-\alpha)f'(\beta)+\int_\alpha^\beta G_1(u,s)f''(s)ds, \tag{3.49}$$

$$f(u)=f(\beta)+(u-\beta)f'(\alpha)+\int_\alpha^\beta G_2(u,s)f''(s)ds, \tag{3.50}$$

$$f(u)=f(\beta)-(\beta-\alpha)f'(\beta)+(u-\alpha)f'(\alpha)+\int_\alpha^\beta G_3(u,s)f''(s)ds, \tag{3.51}$$

$$f(u)=f(\alpha)+(\beta-\alpha)f'(\alpha)-(\beta-u)f'(\beta)+\int_\alpha^\beta G_4(u,s)f''(s)ds, \tag{3.52}$$

$$f(u)=\frac{\beta-u}{\beta-\alpha}f(\alpha)+\frac{u-\beta}{\beta-\alpha}f(\beta)+\int_\alpha^\beta G_5(u,s)f''(s)ds. \tag{3.53}$$

Proof: Let us consider Green's function $G_2(\cdot,s)$ defined by (3.45).

Utilizing integration by parts we have

$$
\begin{aligned}
\int_\alpha^\beta G_2(u,s)f''(s)ds &= \int_\alpha^u G_2(u,s)f''(s)ds + \int_u^\beta G_2(u,s)f''(s)ds \\
&= \int_\alpha^u (u-\beta)f''(s)ds + \int_u^\beta (s-\beta)f''(s)ds \\
&= (u-\beta)f'(s)|_\alpha^u + (s-\beta)f'(s)|_u^\beta - \int_u^\beta f'(s)ds \\
&= (u-\beta)f'(u) - (u-\beta)f'(\alpha) \\
&+ (\beta-\beta)f'(\beta) - (u-\beta)f'(u) - f(\beta) + f(u) \\
&= -(u-\beta)f'(\alpha) - f(\beta) + f(u),
\end{aligned}
$$

from which (3.50) follows immediately. Similarly, we can prove other identities. □

3.2.1. Generalizations of Sherman's inequality by Taylor's formula and Green's functions

We start this subsection with two equivalent statements that include Sherman's type inequalities for real, not necessary nonnegative values of weights $\mathbf{a}$, $\mathbf{b}$ and matrix $\mathbf{A}$.

Theorem 3.2.11: *Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha,\beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha,\beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$, $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$ and*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{A} \tag{3.54}$$

for some matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$ such that $\sum_{j=1}^l a_{ij} = 1$, $i = 1, ..., m$. Let $G_z(\cdot, s)$, $s \in [\alpha,\beta]$, $z = 1,2,3,4,5$ be defined as in (3.44)-(3.48). Then the following statement are equivalent:

(i) For every continuous convex function $f : [\alpha,\beta] \to \mathbb{R}$ we have

$$\sum_{i=1}^m b_i f(y_i) \le \sum_{j=1}^l a_j f(x_j). \tag{3.55}$$

(ii) For all $s \in [\alpha, \beta]$ we have

$$\sum_{i=1}^{m} b_i G_z(y_i, s) \leq \sum_{j=1}^{l} a_j G_z(x_j, s). \tag{3.56}$$

Furthermore, the statements (i) and (ii) are also equivalent if we change the sign of inequality in both (3.55) and (3.56).

Proof: (i)⇒(ii) Let (i) holds. Since $G_z(\cdot, s), s \in [\alpha, \beta]$, is continuous and convex on $[\alpha, \beta]$ for every $z \in \{1, ..., 5\}$, then (3.55) also holds, i.e. the inequality (3.56) holds.
(ii)⇒(i) Let (ii) holds. Let us consider Green's function G_2 defined by (3.45). Applying (3.50) on the Sherman difference $\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)$, by an easy calculation we get

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) = \int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right] f''(s) ds.$$

If the function f is also convex, then $f''(s) \geq 0$ for every $s \in [\alpha, \beta]$. Furthermore, if for every $s \in [\alpha, \beta]$ the inequality (3.56) holds, the required results follows.
Note that it is not necessary to demand the existence of the second derivative of the function f. The differentiability condition can be directly eliminated by using fact that it is possible to approximate uniformly a continuous convex function by convex polynomials (see [138, p. 172]).
The last part of our theorem can be proved analogously. The same conclusion we have for other four Green's functions. □

The next theorem gives two new identities which we use to get new generalizations.

Theorem 3.2.12: *Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$, $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$ be such that (3.54) holds for some matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$ such that $\sum_{j=1}^{l} a_{ij} = 1, i = 1, ..., m$. Let $G_z(\cdot, s), s \in [\alpha, \beta], z = 1, 2, 3, 4, 5$ be defined as in (3.44)-(3.48). Let $n \geq 3$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$. Then*

(i)

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \tag{3.57}$$
$$= \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_k(y_i, s) \right) \cdot (s-\alpha)^k ds$$
$$+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} \left[\int_{t}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds \right] f^{(n)}(t) dt,$$

(ii)

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \tag{3.58}$$
$$= \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\beta)^k ds$$
$$- \frac{1}{(n-3)!} \int_{\alpha}^{\beta} \left[\int_{\alpha}^{t} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds \right] f^{(n)}(t) dt.$$

Proof: (i) Let us consider Green's function G_2 defined by (3.45). Applying (3.50) on the difference $\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)$, by an easy calculation we get

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) = \int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right] f''(s) ds. \tag{3.59}$$

Now, applying Taylor's formula (3.1) to the function f'' at the point $c = \alpha$ we have

$$f''(s) = \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} (s-\alpha)^k + \frac{1}{(n-3)!} \int_{\alpha}^{s} f^{(n)}(t)(s-t)^{n-3} dt. \tag{3.60}$$

Similarly, applying Taylor's formula (3.1) to the function f'' at the point $c = \beta$ we have

$$f''(s) = \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} (s-\beta)^k - \frac{1}{(n-3)!} \int_{s}^{\beta} f^{(n)}(t)(s-t)^{n-3} dt. \tag{3.61}$$

Combining (3.59) and (3.60) we have

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \\
&= \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right) \cdot (s-\alpha)^k ds \qquad (3.62)\\
&+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right) \left(\int_{\alpha}^{s} f^{(n)}(t)(s-t)^{n-3} dt \right) ds.
\end{aligned}$$

Applying Fubini's theorem in the last term of (3.62) we obtain

$$\begin{aligned}
&\frac{1}{(n-3)!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right) \left(\int_{\alpha}^{s} f^{(n)}(t)(s-t)^{n-3} dt \right) ds \\
&= \frac{1}{(n-3)!} \int_{\alpha}^{\beta} \left[\int_{t}^{\beta} \left(\sum_{j=1}^{l} a_j G_2(x_j, s) - \sum_{i=1}^{m} b_i G_2(y_i, s) \right) (s-t)^{n-3} ds \right] f^{(n)}(t) dt. \qquad (3.63)
\end{aligned}$$

Now, (3.57) follows from (3.62) and (3.63).

Analogously we can prove identities for other Green's functions.

(ii) Similarly, combining (3.59) and (3.61) and applying Fubini's theorem we obtain (3.58).

□

The following generalizations of Sherman's theorem for n-convex functions hold.

Theorem 3.2.13: *Let the assumptions of Theorem 3.2.11 hold.*

(i) If f is n-convex and

$$\int_{t}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds \geq 0, \quad t \in [\alpha, \beta], \qquad (3.64)$$

then

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \tag{3.65}$$
$$\geq \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-\alpha)^k ds.$$

(ii) If f is n-convex and

$$\int_{\alpha}^{t} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds \leq 0, \quad t \in [\alpha, \beta], \tag{3.66}$$

then

$$\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \tag{3.67}$$
$$\geq \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-\beta)^k ds.$$

Proof: Since a function f is n-convex, we may assume without loss of generality that f is n-times differentiable and $f^{(n)}(t) \geq 0$, $t \in [\alpha, \beta]$ (see [138, p. 16]).
Now, applying Theorem 3.2.11 we obtain (3.65) and (3.67) respectively. □

In the previous theorem we didn't assume that weights $\mathbf{a}$, $\mathbf{b}$ and matrix $\mathbf{A}$ are with nonnegative entries. But if we take into account assumption of nonnegativity, i.e. the assumptions as in Sherman's theorem (Theorem 2.2.4), then we get the following generalizations.

Theorem 3.2.14: *Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in [0, \infty)^l$, $\mathbf{b} = (b_1, ..., b_m) \in [0, \infty)^m$ be such that (3.54) holds for some row stochastic matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$. Let $G_z(\cdot, s), s \in [\alpha, \beta], z = 1, 2, 3, 4, 5$ be defined as in (3.44)-(3.48). Let $n \geq 3$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$.*

(i) If f is n-convex then

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (3.68)\\
&\geq \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\alpha)^k ds.
\end{aligned}$$

Moreover, if $\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!}(s-\alpha)^k \geq 0$, *then*

$$\sum_{i=1}^{m} b_i f(y_i) \leq \sum_{j=1}^{l} a_j f(x_j). \qquad (3.69)$$

(ii) If n is even and f is n-convex, then

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (3.70)\\
&\geq \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\beta)^k ds.
\end{aligned}$$

Moreover, if $\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!}(s-\beta)^k \geq 0$, *then (3.69) holds.*

(iii) If n is odd and f is n-convex, then

$$\begin{aligned}
&\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \qquad (3.71)\\
&\leq \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\beta)^k ds.
\end{aligned}$$

Moreover, if $\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!}(s-\beta)^k \leq 0$, *then the reverse inequality in (3.69) holds.*

Proof: (i) Since $G_z(x_i, s), s \in [\alpha, \beta]$, is continuous and convex on $[\alpha, \beta]$ for every $z \in \{1, ..., 5\}$, then by Sherman's theorem it holds

$$\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \geq 0.$$

Also, we have that $(s-t)^{n-3} \geq 0$ for $s \in [t, \beta]$. Applying Theorem 3.2.13 we get (3.68).

Further, applying Fubini's theorem we have

$$\begin{aligned} &\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\alpha)^k ds \qquad (3.72) \\ = &\int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot \left(\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} (s-\alpha)^k \right) ds. \end{aligned}$$

If in addition $\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!}(s-\alpha)^k \geq 0$, then the identity in (3.72) is nonnegative and (3.69) immediately follows.

(ii) and (iii) The other parts can be proved similarly. $\square$

3.2.2. Related inequalities of the Ostrowski and Grüss type

We define the functions $\mathcal{R}_z, \mathcal{B}_z : [\alpha, \beta] \to \mathbb{R}$, $z = 1, 2, 3, 4, 5$, by

$$\begin{aligned} \mathcal{R}_z(t) &= \int_{t}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds, \\ \mathcal{B}_z(t) &= \int_{\alpha}^{t} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} ds, \qquad (3.73) \end{aligned}$$

where $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$ *and* $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$ are such that (3.54) holds for some matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$ with $\sum_{j=1}^{l} a_{ij} = 1$, $i = 1, ..., m$, and $G_z(\cdot, s), s \in [\alpha, \beta], z = 1, 2, 3, 4, 5$ are defined as in (3.44)-(3.48), respectively. For these functions the Čebyšev functionals are defined by

$$\begin{aligned} T(\mathcal{R}_z, \mathcal{R}_z) &= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} \mathcal{R}_z^2(t) dt - \left(\frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} \mathcal{R}_z(t) dt \right)^2, \\ T(\mathcal{B}_z, \mathcal{B}_z) &= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} \mathcal{B}_z^2(t) dt - \left(\frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} \mathcal{B}_z(t) dt \right)^2. \end{aligned}$$

Using results for the Čebyšev functionals (Theorem 1.4.12) in a similar way as in the proof of Theorem 3.1.5, we can prove the following theorem which gives upper bounds for identities related to the generalized Sherman inequality.

Theorem 3.2.15: *Let $n \geq 3$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)}$ is absolutely continuous with $(\cdot - \alpha)(\beta - \cdot)(f^{(n+1)})^2 \in L[\alpha, \beta]$. Let the functions $\mathcal{R}_z$ and $\mathcal{B}_z$, $z = 1, 2, 3, 4, 5$, be defined as in (3.73).*

(i) Then the remainder $\kappa_z^1(f; \alpha, \beta)$, defined by

$$\begin{aligned}\kappa_z^1(f;\alpha,\beta) &= \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \\ &- \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\alpha)^k ds \\ &- \frac{f^{(n-1)}(\beta) - f^{(n-1)}(\alpha)}{(\beta-\alpha)(n-3)!} \int_\alpha^\beta \mathcal{R}_z(t) dt,\end{aligned} \tag{3.74}$$

satisfies the estimation

$$\left|\kappa_z^1(f;\alpha,\beta)\right| \leq \frac{\sqrt{\beta-\alpha}}{\sqrt{2}(n-3)!} [T(\mathcal{R}_z, \mathcal{R}_z)]^{\frac{1}{2}} \left| \int_\alpha^\beta (t-\alpha)(\beta-t)[f^{(n+1)}(t)]^2 dt \right|^{\frac{1}{2}}. \tag{3.75}$$

(ii) Then the remainder $\kappa_z^2(f; \alpha, \beta)$, defined by

$$\begin{aligned}\kappa_z^2(f;\alpha,\beta) &= \sum_{i=1}^{m} b_i f(y_i) - \sum_{j=1}^{l} a_j f(x_j) \\ &+ \sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) \cdot (s-\beta)^k ds \\ &+ \frac{f^{(n-1)}(\beta) - f^{(n-1)}(\alpha)}{(\alpha-\beta)(n-3)!} \int_\alpha^\beta \mathcal{B}_z(t) dt,\end{aligned} \tag{3.76}$$

satisfies the estimation

$$\left|\kappa_z^2(f;\alpha,\beta)\right| \leq \frac{\sqrt{\beta-\alpha}}{\sqrt{2}(n-3)!} [T(\mathcal{B}_z, \mathcal{B}_z)]^{\frac{1}{2}} \left| \int_\alpha^\beta (t-\alpha)(\beta-t)[f^{(n+1)}(t)]^2 dt \right|^{\frac{1}{2}}. \tag{3.77}$$

Proof: The idea of the proof is the same as the proof of Theorem 3.1.5. □

Using other result for the Čebyšev functionals (Theorem 1.4.13) we obtain the Grüss type inequalities as follows.

Theorem 3.2.16: *Let $n \geq 3$ and $f : [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n)}$ is absolutely continuous* *and* $f^{(n+1)} \geq 0$ *on* $[\alpha, \beta]$. *Let the functions* $\mathcal{R}_z$ *and* $\mathcal{B}_z$, $z = 1, 2, 3, 4, 5$, *be defined as in (3.73).*

(i) Then the remainder $\kappa_z^1(f; \alpha, \beta)$ *defined by (3.74) satisfies the estimation*

$$\left|\kappa_z^1(f; \alpha, \beta)\right| \leq \frac{1}{(n-3)!}\left\|\mathcal{R}_z'\right\|_\infty \left[\frac{f^{(n-1)}(\beta) + f^{(n-1)}(\alpha)}{2} - \frac{f^{(n-2)}(\beta) - f^{(n-2)}(\alpha)}{\beta - \alpha}\right]. \tag{3.78}$$

(ii) Then the remainder $\kappa_z^2(f; \alpha, \beta)$ *defined by (3.76) satisfies the estimation*

$$\left|\kappa_z^2(f; \alpha, \beta)\right| \leq \frac{1}{(n-3)!}\left\|\mathcal{B}_z'\right\|_\infty \left[\frac{f^{(n-1)}(\beta) + f^{(n-1)}(\alpha)}{2} - \frac{f^{(n-2)}(\beta) - f^{(n-2)}(\alpha)}{\beta - \alpha}\right]. \tag{3.79}$$

Proof: The idea of the proof is the same as in the proof of Theorem 3.1.6. □

The next we present the Ostrowski type inequalities related to generalized Sherman's inequalities (3.65) and (3.67).

Theorem 3.2.17: *Let the assumptions of Theorem 3.2.12 hold. Assume that* (p, q) *is a pair of conjugate exponents, that is* $1 \leq p, q \leq \infty$, $1/p + 1/q = 1$. *Let* $f^{(n)} \in L_p[\alpha, \beta]$. *Then*

(i)

$$\begin{aligned}
&\left|\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)\right. \\
&\left.-\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s)\right) \cdot (s-\alpha)^k ds\right| \\
\leq\ & \frac{1}{(n-3)!}\left\|f^{(n)}\right\|_p \left[\int_\alpha^\beta \left|\int_t^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s)\right) \cdot (s-t)^{n-3} ds\right|^q dt\right]^{\frac{1}{q}}.
\end{aligned} \tag{3.80}$$

The constant on the right-hand side of (3.80) is sharp for $1 < p \leq \infty$ *and the best possible for* $p = 1$.

(ii)

$$\left|\sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \right. \tag{3.81}$$

$$\left. -\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s) \right) \cdot (s-\beta)^k ds \right|$$

$$\leq \frac{1}{(n-3)!} \left\| f^{(n)} \right\|_p \left[\int_{\alpha}^{\beta} \left| \int_{\alpha}^{t} \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s) \right) \cdot (s-t)^{n-3} ds \right|^q dt \right]^{\frac{1}{q}}.$$

The constant on the right-hand side of (3.81) is sharp for $1 < p \leq \infty$ and the best possible for $p = 1$.

Proof: The idea of the proof is the same as that of the proof of Theorem 3.1.7. □

3.2.3. Related mean value theorems and n-exponentially convex functions

In this subsection we present mean value theorems related to the results from the previous subsection.

To simplify the following results, we introduce the following definition motivated by inequalities (3.65) and (3.67).

Definition 3.2.3.: Under the assumptions of Theorem 3.2.13, equipped with conditions (3.64) and (3.66), respectively, we define

$$A_{1,z}(f) = \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i) \tag{3.82}$$

$$-\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\alpha)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s) \right) (s-\alpha)^k ds,$$

$$A_{2,z}(f) = \sum_{j=1}^{l} a_j f(x_j) - \sum_{i=1}^{m} b_i f(y_i)$$

$$-\sum_{k=0}^{n-3} \frac{f^{(k+2)}(\beta)}{k!} \int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s) \right) (s-\beta)^k ds, \quad (z = 1,2,3,4,5).$$

Remark 3.2.4.: If $f : [\alpha, \beta] \to \mathbb{R}$ is n-convex, then by Theorem 3.2.13, for the linear functionals $A_{i,z}, i = 1, 2$, it holds:

$$A_{i,z}(f) \geq 0.$$

Theorem 3.2.18: *Let $A_{i,z}$, $(i = 1, 2,\ z = 1, 2, 3, 4, 5)$, be the linear functionals defined as in (3.82). Let $f \in C^n([\alpha, \beta])$. Then there exists $\xi_i \in [\alpha, \beta]$, $(i = 1, 2)$, such that*

$$A_{i,z}(f) = f^{(n)}(\xi_i) A_{i,z}(\varphi_0), \quad (i = 1, 2,\ z = 1, 2, 3, 4, 5),$$

where $\varphi_0(x) = x^n/n!$.

Proof: The idea of the proof is the same as that of the proof of Theorem 3.1.8. □

Theorem 3.2.19: *Let $A_{i,z}$, $(i = 1, 2,\ z = 1, 2, 3, 4, 5)$, be the linear functionals defined as in (3.82). Let $f, g \in C^n([\alpha, \beta])$. Then there exists $\xi_i \in [\alpha, \beta]$, $(i = 1, 2)$ such that*

$$\frac{f^{(n)}(\xi_i)}{g^{(n)}(\xi_i)} = \frac{A_{i,z}(f)}{A_{i,z}(g)}, \quad (i = 1, 2,\ z = 1, 2, 3, 4, 5), \tag{3.83}$$

assuming neither of the denominators is equal to zero.

Proof: The idea of the proof is the same as that of the proof of Theorem 3.1.9. □

Remark 3.2.5.: If $\frac{f^{(n)}}{g^{(n)}}$ is an invertible function, then from (3.83) it follows

$$\xi_i = \left(\frac{f^{(n)}}{g^{(n)}}\right)^{-1} \left(\frac{A_{i,z}(f)}{A_{i,z}(g)}\right), \quad (i = 1, 2,\ z = 1, 2, 3, 4, 5),$$

which presents a new mean of Cauchy's type of the interval $[\alpha, \beta]$.

Using the families of exponentially convex functions given by Definition 3.1.2. in Subsection 3.1.3., we interpret our results in the form of exponentially convex functions or in the special case logarithmically convex functions as it follows.

Theorem 3.2.20: *Let $A_{i,z}$, $(i = 1, 2,\ z = 1, 2, 3, 4, 5)$, be the linear functionals defined as in (3.82) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$) . Then the following statements hold:*

(i) The function $v \mapsto A_{i,z}(f_v)$, $(i=1,2,\ z=1,2,3,4,5)$, is n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) If the function $v \mapsto A_{i,z}(f_v)$, $(i=1,2,\ z=1,2,3,4,5)$, is continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I.

Proof: *The idea of the proof is the same as that of the proof of Theorem 3.1.10.* □

Corollary 3.2.5.: *Let $A_{i,z}$, $(i=1,2,\ z=1,2,3,4,5)$, be the linear functional defined as in (3.82) associated with the family $\mathcal{F}_3$. Then the following statements hold:*

(i) If the function $v \mapsto A_{i,z}(f_v)$, $(i=1,2,\ z=1,2,3,4,5)$, is continuous on I, then it is 2-exponentially convex on I. If $v \mapsto A_{i,z}(f_v)$, $(i=1,2,\ z=1,2,3,4,5)$, is additionally positive, then it is also log-convex on I.

(ii) If the function $v \mapsto A_{i,z}(f_v)$, $(i=1,2,\ z=1,2,3,4,5)$, is positive and differentiable on I, then for all $r,s,u,w \in I$ such that $r \le u$, $s \le w$, we have

$$M_{r,s}\left(A_{i,z},\mathcal{F}_3\right) \le M_{u,w}\left(A_{i,z},\mathcal{F}_3\right), \tag{3.84}$$

where

$$M_{r,s}\left(A_{i,z},\mathcal{F}_3\right)=\begin{cases}\left(\frac{A_{i,z}(f_r)}{A_{i,z}(f_s)}\right)^{\frac{1}{r-s}}, & r\neq s\\ \exp\left(\frac{\frac{d}{dr}(A_{i,z}(f_r))}{A_{i,z}(f_r)}\right), & r=s\end{cases}, \quad (i=1,2,\ \ z=1,2,3,4,5), \tag{3.85}$$

Proof: The idea of the proof is the same as that of the proof of Corollary 3.1.4.. □

Remark 3.2.6.: Remark 3.1.3. is also valid for these functionals. In an analogous way as in Subsection 3.1.3., we could derive new class of two-parameter Cauchy-type means.

3.3. Sherman's inequality and Fink's identity

In this section, using Fink's identity we obtain new extensions and improvements of Sherman's inequality (2.10) for n-convex functions. Also, we get related results which include the Ostrowski and Grüss type inequalities. At the end of the section, we apply obtained extensions and improvements to estimate as inequalities between geometric, logarithmic and arithmetic means. The results from this section are quoted from [68].

3.3.1. Generalizations of Sherman's inequality by Fink's identity

Through this subsection, we consider simultaneously two aspect and represent two types of results. We obtain the first type of results by using only Fink's identity (2.230). To obtain another case we use Fink's identity (2.230) in combination with Green's functions (3.44)-(3.48).

First we prove two identities.

Theorem 3.3.21: *Let $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{A} \tag{3.86}$$

holds for some matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$ whose entries satisfy the condition $\sum_{j=1}^{l} a_{ij} = 1$ for $i = 1, ..., m$. Let $P(t, \cdot)$ and $G_z(\cdot, s)$, $s, t \in [\alpha, \beta]$, $z = 1, 2, 3, 4, 5$ be defined as in (2.231) and (3.44)-(3.48), respectively. Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi^{(n-1)} \in AC([\alpha, \beta])$.

(i) For $n \geq 1$, the following identity holds:

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \\
&= \frac{1}{\beta - \alpha} \sum_{w=2}^{n-1} \frac{n-w}{w!} \cdot \phi^{(w-1)}(\beta) \left(\sum_{j=1}^{l} a_j (x_j - \beta)^w - \sum_{i=1}^{m} b_i (y_i - \beta)^w \right) \\
&- \frac{1}{\beta - \alpha} \sum_{w=2}^{n-1} \frac{n-w}{w!} \cdot \phi^{(w-1)}(\alpha) \left(\sum_{j=1}^{l} a_j (x_j - \alpha)^w - \sum_{i=1}^{m} b_i (y_i - \alpha)^w \right) \\
&+ \frac{1}{(n-1)!(\beta - \alpha)} \times \\
&\int_{\alpha}^{\beta} \left[\sum_{j=1}^{l} a_j (x_j - t)^{n-1} P(t, x_j) - \sum_{i=1}^{m} b_i (y_i - t)^{n-1} P(t, y_i) \right] \phi^{(n)}(t) dt \ ;
\end{aligned} \tag{3.87}$$

(ii) For $n \geq 3$, the following identity holds:

$$\begin{aligned}
&\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \qquad (3.88)\\
&= \sum_{w=0}^{n-3} \frac{n-w-2}{(\beta-\alpha)w!}\times\\
&\int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s)\right)\left(\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w\right) ds\\
&+ \frac{1}{(n-3)!(\beta-\alpha)}\times\\
&\int_\alpha^\beta \phi^{(n)}(t)\left(\int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s)\right)(s-t)^{n-3}P(t,s)ds\right) dt.
\end{aligned}$$

Proof: (i) Using (2.230) in the difference $\sum\limits_{j=1}^{l} a_j\phi(x_j) - \sum\limits_{i=1}^{m} b_i\phi(y_i)$, we get

$$\begin{aligned}
&\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i)\\
&= \frac{1}{\beta-\alpha}\sum_{w=1}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j(x_j-\beta)^w - \sum_{i=1}^{m} b_i(y_i-\beta)^w\right)\\
&- \frac{1}{\beta-\alpha}\sum_{w=1}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^w - \sum_{i=1}^{m} b_i(y_i-\alpha)^w\right)\\
&+ \frac{1}{(n-1)!(\beta-\alpha)}\times\\
&\int_\alpha^\beta \left[\sum_{j=1}^{l} a_j(x_j-t)^{n-1}P(t,x_j) - \sum_{i=1}^{m} b_i(y_i-t)^{n-1}P(t,y_i)\right]\phi^{(n)}(t)dt.
\end{aligned}$$

Since under assumption (3.86) we have

$$\sum_{j=1}^{l} a_j(x_j-\alpha) - \sum_{i=1}^{m} b_i(y_i-\alpha) = \sum_{j=1}^{l} a_j(x_j-\beta) - \sum_{i=1}^{m} b_i(y_i-\beta) = 0, \qquad (3.89)$$

the identity (3.87) immediately follows.

(ii) Let us consider Green's function G_2 defined by (3.45). Using (3.86) and (3.50) in the difference $\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i)$, we get

$$\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) = \int_{\alpha}^{\beta}\left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s)\right)\phi''(s)ds. \tag{3.90}$$

Applying Fink's identity (2.230) to ϕ'' we have

$$\begin{aligned}\phi''(s) &= \sum_{w=0}^{n-3}\frac{n-w-2}{w!}\cdot\frac{\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w}{\beta-\alpha}\\ &+\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(s-t)^{n-3}P(t,s)\phi^{(n)}(t)dt.\end{aligned} \tag{3.91}$$

By an easy calculation, using (3.90) and (3.91) we get

$$\begin{aligned}\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) = \int_{\alpha}^{\beta}\Bigg[&\left(\sum_{j=1}^{l} a_j G_z(x_j,s) - \sum_{i=1}^{m} b_i G_z(y_i,s)\right)\times\\ &\left(\sum_{w=0}^{n-3}\frac{n-w-2}{w!}\cdot\frac{\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w}{\beta-\alpha}\right.\\ &\left.+\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(s-t)^{n-3}P(t,s)\phi^{(n)}(t)dt\right)\Bigg]ds.\end{aligned}$$

After interchanging the order of summation and integration and applying Fubini's theorem we get (3.88). □

The following generalizations of Sherman's theorem for n-convex functions hold.

Theorem 3.3.22: *Let all the assumptions of Theorem 3.3.21 be satisfied. Additionally, let ϕ be n-convex on $[\alpha,\beta]$.*

(i) If $n \geq 1$ and

$$\sum_{j=1}^{l} a_j(x_j-t)^{n-1}P(t,x_j) - \sum_{i=1}^{m} b_i(y_i-t)^{n-1}P(t,y_i) \geq 0, \quad \alpha \leq t \leq \beta, \tag{3.92}$$

then

$$\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \tag{3.93}$$
$$\geq \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j(x_j-\beta)^w - \sum_{i=1}^{m} b_i(y_i-\beta)^w\right)$$
$$-\frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^w - \sum_{i=1}^{m} b_i(y_i-\alpha)^w\right).$$

If the reverse inequality in (3.92) holds, then the reverse inequality in (3.93) holds.

(ii) If $n \geq 3$ *and*

$$\int_{\alpha}^{\beta}\left(\sum_{j=1}^{l} a_jG_z(x_j,s) - \sum_{i=1}^{m} b_iG_z(y_i,s)\right)(s-t)^{n-3}P(t,s)ds \geq 0, \quad \alpha \leq s,t \leq \beta, \tag{3.94}$$

then

$$\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \tag{3.95}$$
$$\geq \sum_{w=0}^{n-3}\frac{n-w-2}{(\beta-\alpha)w!}\times$$
$$\int_{\alpha}^{\beta}\left(\sum_{j=1}^{l} a_jG_z(x_j,s) - \sum_{i=1}^{m} b_iG_z(y_i,s)\right)\left(\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w\right)ds.$$

If the reverse inequality in (3.94) holds, then the reverse inequality in (3.95) holds.

Proof: (i) Since ϕ is n-convex on $[\alpha,\beta]$, we may assume without loss of generality that ϕ is n-times differentiable and $\phi^{(n)}(t) \geq 0$, $t \in [\alpha,\beta]$ (see [138, p. 16]).
Using this fact and the assumption (3.92), applying Theorem 3.3.21 we obtain (3.93).
(ii) Analogous to the part (i). □

The following generalizations under Sherman's assumption of nonnegativity are also valid.

Theorem 3.3.23: *Let all the assumptions of Theorem 3.3.21 be satisfied. Additionally, let* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{A}$ *be nonnegative and* ϕ *be* n*-convex on* $[\alpha,\beta]$*. Let* n *be even.*

(i) For $n \geq 1$, the inequality (3.93) holds. Moreover, if the function

$$\bar{F}(\cdot) = \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\left[\phi^{(w-1)}(\beta)(\cdot-\beta)^w - \phi^{(w-1)}(\alpha)(\cdot-\alpha)^w\right] \tag{3.96}$$

is convex on $[\alpha, \beta]$, then

$$\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \geq 0. \tag{3.97}$$

(ii) For $n \geq 3$, the inequality (3.95) holds. Moreover, if the function

$$\tilde{F}(\cdot) = \sum_{w=0}^{n-3}\frac{n-w-2}{(\beta-\alpha)w!}\int_{\alpha}^{\beta} G_z(\cdot, s)\left[\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w\right] ds, \tag{3.98}$$

is convex on $[\alpha, \beta]$, then (3.97) holds.

Proof: (i) Consider the function $\lambda : [\alpha, \beta] \to \mathbb{R}$ defined by

$$\lambda(s) = (s-t)^{n-1}P(t,s) = \begin{cases} (s-t)^{n-1}(t-\alpha), & \alpha \leq t \leq s \leq \beta \\ (s-t)^{n-1}(t-\beta), & \alpha \leq s < t \leq \beta \end{cases}. \tag{3.99}$$

Since

$$\lambda''(s) = \begin{cases} (n-1)(n-2)(s-t)^{n-3}(t-\alpha), & \alpha \leq t \leq s \leq \beta \\ (n-1)(n-2)(s-t)^{n-3}(t-\beta), & \alpha \leq s < t \leq \beta \end{cases},$$

it follows that for even $n \geq 2$, λ is convex on $[\alpha, \beta]$. Then by Sherman's theorem, the inequality (3.92) holds. Therefore, by Theorem 3.3.22, the inequality (3.93) holds.

Changing the order of summation, the right hand side of (3.93) can be written in the form

$$\sum_{j=1}^{l} a_j\bar{F}(x_j) - \sum_{i=1}^{m} b_i\bar{F}(y_i),$$

where $\bar{F}$ is defined as in (3.96). If $\bar{F}$ is convex, then by Sherman's theorem we have

$$\sum_{j=1}^{l} a_j\bar{F}(x_j) - \sum_{i=1}^{m} b_i\bar{F}(y_i) \geq 0,$$

i.e. the right hand side of (3.93) is nonnegative, so the inequality (3.97) immediately follows.

(ii) Further, the function $G_z(\cdot, s)$, $s \in [\alpha, \beta]$, is convex on $[\alpha, \beta]$ for every $z = 1, ..., 5$, and by Sherman's theorem we have

$$\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \geq 0.$$

It is easy to see that for even $n > 3$, the inequality

$$\int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} P(t, s) ds \geq 0, \quad \alpha \leq s \leq t, \tag{3.100}$$

holds, while for every $n \geq 3$, we have

$$\int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} P(t, s) ds \geq 0, \quad t \leq s \leq \beta.$$

Now applying Theorem 3.3.23, we get the inequality (3.95).

The rest of proof is analogous to the part (ii), whereby instead of $\bar{F}$ we consider the function $\tilde{F}$ defined by (3.98). □

Remark 3.3.7.: Specially, if we consider case (ii) from the previous theorem when n is odd, since the function λ, defined by (3.99), for $t \leq s$ is convex and for $s < t$ is concave, then we have conclusion: for $t \leq s$, the assumption (3.94) is satisfied and the inequality (3.95) holds, while for $s \leq t$, the reverse of (3.94) holds so the reverse of (3.95) holds.

Remark 3.3.8.: Let all the assumptions of the previous theorem be satisfied.

(i) For even $n \geq 2$, the inequality (3.93) holds.

Further, the function $s \mapsto (s-\alpha)^w$ is convex on $[\alpha, \beta]$ for every w, while $s \mapsto (s-\beta)^w$ is convex on $[\alpha, \beta]$ for even w and concave for odd w.

If for even w, $\phi^{(w-1)}(\alpha) \leq 0$ and $\phi^{(w-1)}(\beta) \geq 0$ and for odd w, $\phi^{(w-1)}(\alpha) \leq 0$ and $\phi^{(w-1)}(\beta) \leq 0$, then the right hand side of (3.93) is nonnegative. Therefore, (3.97) immediately follows.

(ii) For even $n \geq 4$, the inequality (3.95) holds.

Further, when $\alpha \leq s \leq \beta$, we have $(s-\alpha)^w \geq 0$ for every w while $(s-\beta)^w \geq 0$ for even w

and $(s-\beta)^w \le 0$ for odd w.

If for even w, $\phi^{(w+1)}(\alpha) \le 0$ and $\phi^{(w+1)}(\beta) \ge 0$ and for odd w, $\phi^{(w+1)}(\alpha) \le 0$ and $\phi^{(w+1)}(\beta) \le 0$, then the right hand side of (3.95) is nonnegative and the inequality (3.97) immediately follows.

3.3.2. New Ostrowski and Grüss type inequalities

To avoid many notations, we define the functions $\mathcal{B}, \tilde{\mathcal{B}}_z : [\alpha,\beta] \to \mathbb{R}$, $z = 1, ..., 5$, by

$$\mathcal{B}(t) = \sum_{j=1}^{l} a_j (x_j - t)^{n-1} P(t, x_j) - \sum_{i=1}^{m} b_i (y_i - t)^{n-1} P(t, y_i), \quad n \ge 1, \tag{3.101}$$

$$\tilde{\mathcal{B}}_z(t) = \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s) \right) (s-t)^{n-3} P(t, s) ds, \quad n \ge 3,$$

where $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ *and* $\mathbf{b} \in \mathbb{R}^m$ are such that (3.86) holds for some matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$ whose entries satisfy the condition $\sum_{j=1}^{l} a_{ij} = 1$ for $i = 1, ..., m$, and P, G_z, $z = 1, 2, 3, 4, 5$, are defined as in (2.231) and (3.44)-(3.48), respectively.

For such functions we consider the Čebyšev functionals defined by

$$T(\mathcal{B}, \mathcal{B}) = \frac{1}{\beta - \alpha} \int_\alpha^\beta \mathcal{B}^2(t) dt - \left(\frac{1}{\beta - \alpha} \int_\alpha^\beta \mathcal{B}(t) dt \right)^2,$$

$$T(\tilde{\mathcal{B}}_z, \tilde{\mathcal{B}}_z) = \frac{1}{\beta - \alpha} \int_\alpha^\beta \tilde{\mathcal{B}}_z^2(t) dt - \left(\frac{1}{\beta - \alpha} \int_\alpha^\beta \tilde{\mathcal{B}}_z(t) dt \right)^2, \quad z = 1, ..., 5.$$

Theorem 3.3.24: *Let $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^m$, $\mathbf{a} \in \mathbb{R}^l$ and $\mathbf{b} \in \mathbb{R}^m$ be such that (3.86) holds for some matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$ whose entries satisfy the condition $\sum_{j=1}^{l} a_{ij} = 1$ for $i = 1, ..., m$. Let $\mathcal{B}$, $\tilde{\mathcal{B}}_z$, $z = 1, ..., 5$, be defined as in (3.101). Let $\phi : [\alpha,\beta] \to \mathbb{R}$ be such that $\phi^{(n)} \in AC([\alpha,\beta])$ with $(\cdot - \alpha)(\beta - \cdot)(\phi^{(n+1)})^2 \in L_1[\alpha,\beta]$.*

(i) For $n \geq 1$, the remainder $R_n(\phi;\alpha,\beta)$, defined by

$$\begin{aligned} R_n(\phi;\alpha,\beta) &= \sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\ &- \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j(x_j-\beta)^w - \sum_{i=1}^{m} b_i(y_i-\beta)^w\right) \\ &+ \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^w - \sum_{i=1}^{m} b_i(y_i-\alpha)^w\right) \\ &- \frac{\phi^{(n-1)}(\beta)-\phi^{(n-1)}(\alpha)}{(\beta-\alpha)^2(n-1)!}\int_\alpha^\beta \mathcal{B}(t)dt, \end{aligned} \tag{3.102}$$

satisfies the estimation

$$|R_n(\phi;\alpha,\beta)| \leq \frac{1}{\sqrt{2(\beta-\alpha)}(n-1)!}[T(\mathcal{B},\mathcal{B})]^{\frac{1}{2}}\left(\int_\alpha^\beta (t-\alpha)(\beta-t)[\phi^{(n+1)}(t)]^2dt\right)^{\frac{1}{2}}. \tag{3.103}$$

(ii) For $n \geq 3$, the remainder $\tilde{R}_{nz}(\phi;\alpha,\beta)$, $z \in \{1,...,5\}$, defined by

$$\begin{aligned} \tilde{R}_{nz}(\phi;\alpha,\beta) &= \sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\ &- \sum_{w=0}^{n-3}\frac{n-w-2}{(\beta-\alpha)w!}\int_\alpha^\beta \mathcal{G}_z(s)\left(\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w\right)ds \\ &- \frac{\phi^{(n-1)}(\beta)-\phi^{(n-1)}(\alpha)}{(n-3)!(\beta-\alpha)^2}\int_\alpha^\beta \tilde{\mathcal{B}}_z(t)dt, \end{aligned} \tag{3.104}$$

where $\mathcal{G}_z(s) = \sum_{j=1}^{l} a_jG_z(x_j,s) - \sum_{i=1}^{m} b_iG_z(y_i,s)$ for $G_z(\cdot,\cdot)$, $z = 1,...,5$, defined by (3.44)-(3.48), satisfies the estimation

$$\left|\tilde{R}_{nz}(\phi;\alpha,\beta)\right| \leq \frac{1}{\sqrt{2(\beta-\alpha)}(n-3)!}[T(\tilde{\mathcal{B}}_z,\tilde{\mathcal{B}}_z)]^{\frac{1}{2}}\left(\int_\alpha^\beta (t-\alpha)(\beta-t)[\phi^{(n+1)}(t)]^2dt\right)^{\frac{1}{2}}. \tag{3.105}$$

Proof: We prove the case (ii): Comparing (3.104) and (3.88) we have

$$
\begin{aligned}
\tilde{R}_{nz}(\phi;\alpha,\beta) &= \frac{1}{(n-3)!(\beta-\alpha)}\int_\alpha^\beta \tilde{\mathcal{B}}_z(t)\phi^{(n)}(t)dt - \frac{\phi^{(n-1)}(\beta)-\phi^{(n-1)}(\alpha)}{(n-3)!(\beta-\alpha)^2}\int_\alpha^\beta \tilde{\mathcal{B}}_z(t)dt \\
&= \frac{1}{(n-3)!(\beta-\alpha)}\int_\alpha^\beta \tilde{\mathcal{B}}_z(t)\phi^{(n)}(t)dt \\
&- \frac{1}{(n-3)!(\beta-\alpha)}\int_\alpha^\beta \tilde{\mathcal{B}}_z(t)dt\cdot\frac{1}{\beta-\alpha}\int_\alpha^\beta \phi^{(n)}(t)dt \\
&= \frac{1}{(n-3)!}T(\tilde{\mathcal{B}}_z,\phi^{(n)}).
\end{aligned}
$$

Applying Theorem 1.4.12 to $\mathcal{B}_k$ and $\phi^{(n)}$, we obtain (3.105).

Case (i) we can prove on analog way. □

The following theorem gives the Grüss type inequalities related to the identities (3.87) and (3.88).

Theorem 3.3.25: *Let $\phi:[\alpha,\beta]\to\mathbb{R}$ be such that $\phi^{(n)}\in AC([\alpha,\beta])$ and $\phi^{(n+1)}\geq 0$ on $[\alpha,\beta]$. Let $\mathcal{B}$, $\tilde{\mathcal{B}}_z$, $z=1,...,5$, be defined as in (3.101).*

(i) For $n\geq 1$, the remainder $R_n(\phi;\alpha,\beta)$, defined by (3.102), satisfies the estimation

$$
|R_n(\phi;\alpha,\beta)| \leq \frac{1}{(n-1)!}\left\|\mathcal{B}'\right\|_\infty\left[\frac{\phi^{(n-1)}(\beta)+\phi^{(n-1)}(\alpha)}{2}-\frac{\phi^{(n-2)}(\beta)-\phi^{(n-2)}(\alpha)}{\beta-\alpha}\right]. \tag{3.106}
$$

(ii) For $n\geq 3$, the remainder $\tilde{R}_{nz}(\phi;\alpha,\beta)$, $z\in\{1,...,5\}$, defined by the representation (3.104), satisfies the estimation

$$
\left|\tilde{R}_{nz}(\phi;\alpha,\beta)\right| \leq \frac{1}{(n-3)!}\left\|\tilde{\mathcal{B}}_z'\right\|_\infty\left[\frac{\phi^{(n-1)}(\beta)+\phi^{(n-1)}(\alpha)}{2}-\frac{\phi^{(n-2)}(\beta)-\phi^{(n-2)}(\alpha)}{\beta-\alpha}\right]. \tag{3.107}
$$

Proof: We prove the case (ii): We have $\tilde{R}_{nz}(\phi;\alpha,\beta)=\frac{1}{(n-3)!}T(\tilde{\mathcal{B}}_z,\phi^{(n)})$.

Applying Theorem 1.4.13 on the functions $\tilde{\mathcal{B}}_z$ and $\phi^{(n)}$ and using the identity

$$
\begin{aligned}
&\int_\alpha^\beta (t-\alpha)(\beta-t)\phi^{(n+1)}(t)dt = \int_\alpha^\beta \left[2t-(\alpha+\beta)\right]\phi^{(n)}(t)dt \\
&= (\beta-\alpha)\left[\phi^{(n-1)}(\beta)+\phi^{(n-1)}(\alpha)\right] - 2\left[\phi^{(n-2)}(\beta)-\phi^{(n-2)}(\alpha)\right],
\end{aligned}
$$

we obtain (3.107).

Case (i) we can prove on analog way. □

In the next theorem we give the Ostrowski type inequalities related to the identities (3.87) and (3.88). We omit the proof since the technique of proving is same as in the proof of Theorem 3.1.7.

Theorem 3.3.26: *Let the assumptions of Theorem 3.3.21 hold. Furthermore, let $\mathcal{B}$, $\tilde{\mathcal{B}}_z$, $z = 1, ..., 5$, be defined as in (3.101). Let (p, q) be a pair of conjugate exponents, that is $1 \le p, q \le \infty$, $1/p + 1/q = 1$ and $\phi^{(n)} \in L_p[\alpha, \beta]$.*

(i) For $n \ge 1$, we have

$$\begin{aligned}&\left|\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i)\right. \\ &- \frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\beta)\left(\sum_{j=1}^{l} a_j(x_j-\beta)^w - \sum_{i=1}^{m} b_i(y_i-\beta)^w\right) \\ &\left.+\frac{1}{\beta-\alpha}\sum_{w=2}^{n-1}\frac{n-w}{w!}\cdot\phi^{(w-1)}(\alpha)\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^w - \sum_{i=1}^{m} b_i(y_i-\alpha)^w\right)\right| \\ &\le \frac{1}{(n-1)!(\beta-\alpha)}\left(\int_\alpha^\beta |\mathcal{B}(t)|^q\,dt\right)^{\frac{1}{q}}\left\|\phi^{(n)}\right\|_p .\end{aligned} \tag{3.108}$$

The constant on the right hand side of (3.108) is sharp for $1 < p \le \infty$ and the best possible for $p = 1$.

(ii) For $n \ge 3$, we have

$$\begin{aligned}&\left|\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i)\right. \\ &\left.-\sum_{w=0}^{n-3}\frac{n-w-2}{(\beta-\alpha)w!}\int_\alpha^\beta \mathcal{G}_z(s)\left(\phi^{(w+1)}(\beta)(s-\beta)^w - \phi^{(w+1)}(\alpha)(s-\alpha)^w\right)ds\right| \\ &\le \frac{1}{(n-3)!(\beta-\alpha)}\left(\int_\alpha^\beta \left|\tilde{\mathcal{B}}_z(t)\right|^q dt\right)^{\frac{1}{q}}\left\|\phi^{(n)}\right\|_p ,\end{aligned} \tag{3.109}$$

where $\mathcal{G}_z(s) = \sum_{j=1}^{l} a_j G_z(x_j, s) - \sum_{i=1}^{m} b_i G_z(y_i, s)$ for $G_z(\cdot,\cdot)$, $z = 1, ..., 5$, defined by (3.44)-(3.48).

The constant on the right hand side of (3.109) is sharp for $1 < p \leq \infty$ *and the best possible for* $p = 1$.

3.3.3. Applications to means

Considering the particular cases of results from the previous subsection, we present some consequences. As applications we obtain extensions and improvements of majorization inequalities, discrete Jensen's inequality as well as inequalities between geometric, logarithmic and arithmetic means.

As a direct consequence of Theorem 3.3.22 (ii) we get the following corollary.

Corollary 3.3.6.: *Let* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{A}$, $P(t,s)$ *and* $G_k(\cdot, s)$, $s,t \in [\alpha,\beta]$, $k = 1,2,3,4,5$, *be as in Theorem 3.3.22. Let a function* $\phi : [\alpha,\beta] \to \mathbb{R}$ *be* 3*-convex such that* $\phi'' \in AC([\alpha,\beta])$.

(i) If $t \leq s$, *then*

$$\int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_k(x_j, s) - \sum_{i=1}^{m} b_i G_k(y_i, s) \right) P(t,s) ds \geq 0 \tag{3.110}$$

and

$$\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \geq \frac{\phi'(\beta) - \phi'(\alpha)}{\beta - \alpha} \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_k(x_j, s) - \sum_{i=1}^{m} b_i G_k(y_i, s) \right) ds. \tag{3.111}$$

Moreover, if in addition $\phi'(\beta) \geq \phi'(\alpha)$, *then the left hand side of (3.111) is nonnegative.*

(ii) If $s < t$, *then the reverse inequalities in (3.110) and (3.111) hold. Moreover, if in addition* $\phi'(\beta) \leq \phi'(\alpha)$, *then the left hand side of the reversed (3.111) is nonpositive.*

Remark 3.3.9.: Note that (3.111) gives a nonnegative lower bound for Sherman's difference $\sum\limits_{j=1}^{l} a_j \phi(x_j) - \sum\limits_{i=1}^{m} b_i \phi(y_i)$ when $\phi'(\beta) \geq \phi'(\alpha)$.

Specially, from (3.111), for $m = 1$, $b_1 = 1$, $y_1 = \sum_{i=1}^{l} a_i x_i = \bar{x}$, we get Jensen's type inequality

$$\sum_{j=1}^{l} a_j \phi(x_j) - \phi(\bar{x}) \geq \frac{\phi'(\beta) - \phi'(\alpha)}{\beta - \alpha} \int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_k(x_j, s) - G_k(\bar{x}, s) \right) ds.$$

If in addition $\phi'(\beta) \geq \phi'(\alpha)$, then we get the following double inequality

$$\begin{aligned}\phi(\bar{x}) &\leq \frac{\phi'(\beta)-\phi'(\alpha)}{\beta-\alpha}\int_\alpha^\beta \left(\sum_{j=1}^{l} a_j G_k(x_j,s) - G_k(\bar{x},s)\right) ds + \phi(\bar{x}) \\ &\leq \sum_{j=1}^{l} a_j \phi(x_j)\end{aligned} \tag{3.112}$$

which improves Jensen's inequality.

Further, from (3.111), choosing Green's function G_2, defined by (3.45), and setting $l=2$, $a_1=a_2=\frac{1}{2}$, $x_1=\alpha$, $x_2=\beta$, $\bar{x}=\frac{\alpha+\beta}{2}$, we have

$$\int_\alpha^\beta \left(\sum_{j=1}^{2} a_j G_2(x_j,s) - G_2(\bar{x},s)\right) ds = \frac{(\beta-\alpha)^2}{4},$$

and (3.112) reduces to

$$\phi\left(\frac{\alpha+\beta}{2}\right) \leq \frac{\phi'(\beta)-\phi'(\alpha)}{4}(\beta-\alpha) + \phi\left(\frac{\alpha+\beta}{2}\right) \leq \frac{\phi(\alpha)+\phi(\beta)}{2}. \tag{3.113}$$

In the sequel, appying the double inequality (3.113) to some concrete functions, we derive some new inequalities for means.

Example 3.: *Consider the function $\phi : [\alpha,\beta] \to \mathbb{R}$ defined by $\phi(x) = e^x$. Since ϕ is 3-convex, $\phi'' \in AC([\alpha,\beta])$ and $\phi'(\beta) \geq \phi'(\alpha)$ is satisfied, then by (3.113) we have*

$$e^{\frac{\alpha+\beta}{2}} \leq \frac{e^\beta - e^\alpha}{4}(\beta-\alpha) + e^{\frac{\alpha+\beta}{2}} \leq \frac{e^\alpha + e^\beta}{2}.$$

By substituting $x = e^\alpha$, $y = e^\beta$, we get

$$\sqrt{xy} \leq \frac{y-x}{\ln y - \ln x}\frac{(\ln y - \ln x)^2}{4} + \sqrt{xy} \leq \frac{x+y}{2},$$

i.e. we get the inequalities between the geometric mean $G = \sqrt{xy}$, the logarithmic mean $L = \frac{y-x}{\ln y - \ln x}$ and the arithmetic mean $A = \frac{x+y}{2}$, in the form

$$G \leq L\frac{(\ln y - \ln x)^2}{4} + G \leq A.$$

Example 4.: *Let* $0 < \alpha < \beta$ *and the function* $\phi : [\alpha, \beta] \to \mathbb{R}$ *be defined by* $\phi(x) = x^{r+1}$, $r \geq 1$. *Since* ϕ *is* 3*-convex,* $\phi'' \in AC([\alpha, \beta])$ *and* $\phi'(\beta) \geq \phi'(\alpha)$ *is satisfied, then by (3.113) we have*

$$\left(\frac{\alpha+\beta}{2}\right)^{r+1} \leq \frac{r+1}{4}(\beta-\alpha)(\beta^r-\alpha^r) + \left(\frac{\alpha+\beta}{2}\right)^{r+1} \leq \frac{\alpha^{r+1}+\beta^{r+1}}{2}.$$

The following corollary is direct consequence of Theorem 3.3.26 (ii).

Corollary 3.3.7.: *Assume that* (p, q) *is a pair of conjugate exponents, i.e.* $1 \leq p, q \leq \infty$, $1/p + 1/q = 1$. *Let* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{A}$, $P(t, s)$ *and* $G_k(\cdot, s)$, $s, t \in [\alpha, \beta]$, $k = 1, 2, 3, 4, 5$, *be as in Theorem 3.3.26. Let a function* $\phi : [\alpha, \beta] \to \mathbb{R}$ *be such that* ϕ'' *is absolutely continuous and* $\phi''' \in L_p[\alpha, \beta]$. *Then*

$$\left|\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) - \frac{\phi'(\beta)-\phi'(\alpha)}{\beta-\alpha}\int_\alpha^\beta \left(\sum_{j=1}^{l} a_jG_k(x_j, s) - \sum_{i=1}^{m} b_iG_k(y_i, s)\right) ds\right| \tag{3.114}$$

$$\leq \frac{1}{\beta-\alpha}\left(\int_\alpha^\beta \left(\int_\alpha^\beta \left(\sum_{j=1}^{l} a_jG_k(x_j, s) - \sum_{i=1}^{m} b_iG_k(y_i, s)\right) P(t, s)ds\right)^q dt\right)^{\frac{1}{q}} \left\|\phi'''\right\|_p.$$

The constant on the right hand side is sharp for $1 < p \leq \infty$ *and the best possible for* $p = 1$.

Remark 3.3.10.: Choosing Green's function G_2, defined by (3.45), and setting $l = 2$, $a_1 = a_2 = \frac{1}{2}$, $x_1 = \alpha$, $x_2 = \beta$, and $m = 1$, $b_1 = 1, y_1 = \bar{x} = \frac{\alpha+\beta}{2}$ with $t \leq s$, from (3.114) we get

$$\left|\phi(\alpha) + \phi(\beta) - 2\phi\left(\frac{\alpha+\beta}{2}\right) - \frac{\phi\prime(\beta)-\phi\prime(\alpha)}{2}(\beta-\alpha)\right|$$
$$\leq \frac{\beta-\alpha}{2}\left(\int_\alpha^\beta (t-\alpha)^q dt\right)^{\frac{1}{q}} \left\|\phi'''\right\|_p.$$

If we choose $q = 1$, $p = \infty$, we obtain

$$\left|\phi(\alpha) + \phi(\beta) - 2\phi\left(\frac{\alpha+\beta}{2}\right) - \frac{\phi\prime(\beta) - \phi\prime(\alpha)}{2}(\beta-\alpha)\right|$$
$$\leq \frac{(\beta-\alpha)^3}{4}\left\|\phi'''\right\|_{\infty}.$$

Choosing $q = \infty$, $p = 1$, we have

$$\left|\phi(\alpha) + \phi(\beta) - 2\phi\left(\frac{\alpha+\beta}{2}\right) - \frac{\phi\prime(\beta) - \phi\prime(\alpha)}{2}(\beta-\alpha)\right|$$
$$\leq \frac{(\beta-\alpha)^2}{2}(\phi''(\beta) - \phi''(\alpha)).$$

If in addition ϕ is 3-convex, then

$$0 \leq \phi(\alpha) + \phi(\beta) - 2\phi\left(\frac{\alpha+\beta}{2}\right) - \frac{\phi\prime(\beta) - \phi\prime(\alpha)}{2}(\beta-\alpha) \leq \frac{(\beta-\alpha)^2}{2}(\phi''(\beta) - \phi''(\alpha)),$$

i.e. we get the double inequality

$$\begin{aligned}\phi\left(\frac{\alpha+\beta}{2}\right) &\leq \frac{\phi(\alpha)+\phi(\beta)}{2} - \frac{\beta-\alpha}{4}(\phi\prime(\beta) - \phi\prime(\alpha)) \\ &\leq \phi\left(\frac{\alpha+\beta}{2}\right) + \frac{(\beta-\alpha)^2}{4}(\phi''(\beta) - \phi''(\alpha)).\end{aligned} \tag{3.115}$$

In the following examples, applying the double inequality (3.115) to some concrete functions, we derive some new inequalities for means.

Example 5.: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be defined by $\phi(x) = e^x$. Then*

$$e^{\frac{\alpha+\beta}{2}} \leq \frac{e^{\alpha}+e^{\beta}}{2} - \frac{e^{\beta}-e^{\alpha}}{4}(\beta-\alpha) \leq e^{\frac{\alpha+\beta}{2}} + \frac{e^{\beta}-e^{\alpha}}{4}(\beta-\alpha)^2.$$

By substituting $x = e^{\alpha}$, $y = e^{\beta}$, we get the inequalities between means $G = \sqrt{xy}$, $L = \frac{y-x}{\ln y - \ln x}$ and $A = \frac{x+y}{2}$:

$$L \leq A + L\frac{(\ln y - \ln x)^2}{4} \leq G + L\frac{(\ln y - \ln x)^3}{4}.$$

Example 6.: *Let* $0 < \alpha < \beta$, $\phi : [\alpha, \beta] \to \mathbb{R}$ *be defined by* $\phi(x) = x^{r+1}$, $r \geq 1$. *Then*

$$\begin{aligned}\left(\frac{\alpha+\beta}{2}\right)^{r+1} &\leq \frac{\alpha^{r+1}+\beta^{r+1}}{2} - \frac{r+1}{4}(\beta-\alpha)(\beta^r - \alpha^r) \\ &\leq \left(\frac{\alpha+\beta}{2}\right)^{r+1} + \frac{r(r+1)}{4}(\beta-\alpha)^2(\beta^{r-1} - \alpha^{r-1}).\end{aligned}$$

Remark 3.3.11.: Choosing other Green's functions (3.44)-(3.48), we can estimate analogous related results.

3.4. Sherman's inequality and Lidstone's interpolation polynomials with Green's functions

In 1929, Lidstone [103] introduced a generalization of Taylor's series that approximates a given function in the neighborhood of two points instead of one. This series includes the polynomials, later called Lidstone's polynomials, which have been studied in the works of Boas [24], Poritsky [142], Widder [157] and others.

In this section, using Lidstone's polynomials in combination with Green's function we obtain new generalizations of Sherman's inequality for $2n$-convex functions (see [10]).

The following definition and lemma can be found in [11].

Definition 3.4.4.: Let $f \in C^{\infty}([0,1])$, then the Lidstone series has the form

$$\sum_{k=0}^{\infty} \left(f^{(2k)}(0)\Lambda_k(1-x) + f^{(2k)}(1)\Lambda_k(x) \right),$$

where Λ_k is polynomial of degree $2n+1$ defined by the relations

$$\begin{aligned}&\Lambda_0(t) = t, \\ &\Lambda_n''(t) = \Lambda_{n-1}(t), \quad n \geq 1, \\ &\Lambda_n(0) = \Lambda_n(1) = 0, \quad n \geq 1.\end{aligned} \tag{3.116}$$

In [157], Widder proved the following fundamental lemma.

Lemma 3.4.2.: *If* $f \in C^{2n}([0,1])$, *then*

$$f(t) = \sum_{k=0}^{n-1} \left(f^{(2k)}(0)\Lambda_k(1-t) + f^{(2k)}(1)\Lambda_k(t) \right) + \int_0^1 G_n(t,s) f^{(2n)}(s)ds,$$

where

$$G_1(t,s) = G(t,s) = \begin{cases} (t-1)s, & s \leq t \\ (s-1)t, & t \leq s \end{cases}, \tag{3.117}$$

is homogeneous Green's function of the differential operator $\frac{d^2}{ds^2}$ *on* $[0,1]$ *and with the successive iterates of* $G(t,s)$

$$G_n(t,s) = \int_0^1 G_1(t,p) G_{n-1}(p,s)dp, \quad n \geq 2. \tag{3.118}$$

Remark 3.4.12.: Green's function G, defined by (3.117), is convex and continuous with respect to both variables s and t.

The Lidstone polynomial can be expressed in terms of $G_n(t,s)$ as

$$\Lambda_n(t) = \int_0^1 G_n(t,s)sds. \tag{3.119}$$

3.4.1. Generalizations of Sherman's inequality by Lidstone's interpolation polynomials and Green's functions

By using interpolation with Lidstone's interpolating polynomials we prove the following identity.

Theorem 3.4.27: *Let* $n \in \mathbb{N}$, $\phi \in C^{2n}([\alpha,\beta])$ *and* $\mathbf{x} = (x_1, ..., x_l) \in [\alpha,\beta]^l$, $\mathbf{y} =$

$(y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$ *and* $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$. *Then*

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \qquad (3.120)\\
&= \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\alpha) \left[\sum_{j=1}^{l} a_j \Lambda_k \left(\frac{\beta - x_j}{\beta - \alpha} \right) - \sum_{i=1}^{m} b_i \Lambda_k \left(\frac{\beta - y_i}{\beta - \alpha} \right) \right] \\
&+ \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\beta) \left[\sum_{j=1}^{l} a_j \Lambda_k \left(\frac{x_j - \alpha}{\beta - \alpha} \right) - \sum_{i=1}^{m} b_i \Lambda_k \left(\frac{y_i - \alpha}{\beta - \alpha} \right) \right] \\
&+ (\beta-\alpha)^{2n-1} \int_\alpha^\beta \left[\sum_{j=1}^{l} a_j G_n \left(\tfrac{x_j-\alpha}{\beta-\alpha}, \tfrac{t-\alpha}{\beta-\alpha} \right) - \sum_{i=1}^{m} b_i G_n \left(\tfrac{y_i-\alpha}{\beta-\alpha}, \tfrac{t-\alpha}{\beta-\alpha} \right) \right] \phi^{(2n)}(t) dt,
\end{aligned}$$

where Λ_k *and* G_n *are defined as in (3.116) and (3.117)-(3.118), respectively.*

Proof: By Widder's lemma, we can represent every function $\phi \in C^{2n}([\alpha, \beta])$ in the form:

$$\begin{aligned}
\phi(x) = &\sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \left[\phi^{(2k)}(\alpha) \Lambda_k \left(\frac{\beta - x}{\beta - \alpha} \right) + \phi^{(2k)}(\beta) \Lambda_k \left(\frac{x - \alpha}{\beta - \alpha} \right) \right] \\
&+ (\beta-\alpha)^{2n-1} \int_\alpha^\beta G_n \left(\frac{x-\alpha}{\beta-\alpha}, \frac{s-\alpha}{\beta-\alpha} \right) \phi^{(2n)}(s) ds. \qquad (3.121)
\end{aligned}$$

Using (3.121), we have

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \qquad (3.122)\\
&= \sum_{j=1}^{l} a_j \left\{ \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \left[\phi^{(2k)}(\alpha) \Lambda_k \left(\frac{\beta - x_j}{\beta - \alpha} \right) + \phi^{(2k)}(\beta) \Lambda_k \left(\frac{x_j - \alpha}{\beta - \alpha} \right) \right] \right\} \\
&+ \sum_{j=1}^{l} a_j \left[(\beta-\alpha)^{2n-1} \int_\alpha^\beta G_n \left(\frac{x_j-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha} \right) \phi^{(2n)}(t) dt \right] \\
&- \sum_{i=1}^{m} b_i \left\{ \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \left[\phi^{(2k)}(\alpha) \Lambda_k \left(\frac{\beta - y_i}{\beta - \alpha} \right) + \phi^{(2k)}(\beta) \Lambda_k \left(\frac{y_i - \alpha}{\beta - \alpha} \right) \right] \right\} \\
&- \sum_{i=1}^{m} b_i \left[(\beta-\alpha)^{2n-1} \int_\alpha^\beta G_n \left(\frac{y_i-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha} \right) \phi^{(2n)}(t) dt \right].
\end{aligned}$$

By an easy calculation, from (3.122) we get (3.120). □

Using the previous identity, we get the following generalizations of Sherman's theorem for $(2n)$-convex functions.

Theorem 3.4.28: *Let $n \in \mathbb{N}$, $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$ and $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$. Let Λ_k and G_n be defined as in (3.116) and (3.117)-(3.118), respectively. If for all $t \in [\alpha, \beta]$, the following assumption*

$$\sum_{i=1}^{m} b_i G_n\left(\frac{y_i-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right) \le \sum_{j=1}^{l} a_j G_n\left(\frac{x_j-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right), \tag{3.123}$$

holds, then for every $(2n)$-convex function $\phi : [\alpha, \beta] \to \mathbb{R}$, we have

$$\begin{aligned} &\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \qquad (3.124)\\ &\ge \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\alpha) \left[\sum_{j=1}^{l} a_j \Lambda_k\left(\frac{\beta-x_j}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i \Lambda_k\left(\frac{\beta-y_i}{\beta-\alpha}\right)\right] \\ &+ \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\beta) \left[\sum_{j=1}^{l} a_j \Lambda_k\left(\frac{x_j-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i \Lambda_k\left(\frac{y_i-\alpha}{\beta-\alpha}\right)\right]. \end{aligned}$$

If the reverse inequality in (3.123) holds, then the reverse inequality in (3.124) holds.

Proof: Since a function ϕ is $(2n)$-convex, we may assume without loss of generality that ϕ is $2n$-times differentiable and $\phi^{(2n)}(t) \ge 0$, $t \in [\alpha, \beta]$ (see [138, p. 16]).
Using this fact and the assumption (3.123), applying Theorem 3.4.27 we obtain (3.124).
□

Specially, if we use assumptions as in Sherman's theorem (Theorem 2.2.4), then we get the following generalizations.

Theorem 3.4.29: *Let $n \in \mathbb{N}$, $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in [0, \infty)^l$ and $\mathbf{b} = (b_1, ..., b_m) \in [0, \infty)^m$ be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{A} \tag{3.125}$$

holds for some row stochastic matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$. Let Λ_k and G_n be defined as in (3.116) and (3.117)-(3.118), respectively.

(i) If n is odd, then for every $2n$-convex function $\phi : [\alpha, \beta] \to \mathbb{R}$, it holds

$$\begin{aligned} &\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\ &\geq \sum_{k=1}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\alpha)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{\beta-x_j}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{\beta-y_i}{\beta-\alpha}\right)\right] \\ &+ \sum_{k=1}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\beta)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{x_j-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{y_i-\alpha}{\beta-\alpha}\right)\right]. \end{aligned} \tag{3.126}$$

Moreover, if $\phi^{(2k)}(\alpha) \geq 0$, $\phi^{(2k)}(\beta) \geq 0$ for $k = 1, 3, ..., n-2$ and $\phi^{(2k)}(\alpha) \leq 0$, $\phi^{(2k)}(\beta) \leq 0$ for $k = 2, 4, ..., n-1$, then

$$\sum_{i=1}^{m} b_i\phi(y_i) \leq \sum_{j=1}^{l} a_j\phi(x_j). \tag{3.127}$$

(ii) If n is even, then for every $2n$-convex function $\phi : [\alpha, \beta] \to \mathbb{R}$, the reverse inequality in (3.126) holds. Moreover, if $\phi^{(2k)}(\alpha) \leq 0$, $\phi^{(2k)}(\beta) \leq 0$ for $k = 1, 3, ..., n-1$ and $\phi^{(2k)}(\alpha) \geq 0$, $\phi^{(2k)}(\beta) \geq 0$ for $k = 2, 4, ..., n-2$, then the reverse inequality in (3.127) holds.

Proof: (i) From (3.117), it follows that $G_1(t, s) \leq 0$ for $0 \leq t, s \leq 1$.

From (3.118), it follows that $G_n(t, s) \leq 0$ for odd n and $G_n(t, s) \geq 0$ for even n.

Now, since G_1 is convex and G_{n-1} is nonnegative for odd n, using (3.118), we conclude that G_n is convex in first variable if n is odd. Similarly, we have that G_n is concave in first variable if n is even.

Hence, if n is odd, then by Sherman's theorem we have

$$\sum_{i=1}^{m} b_i G_n\left(\frac{y_i-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right) \leq \sum_{j=1}^{l} a_j G_n\left(\frac{x_j-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right). \tag{3.128}$$

Therefore, in this case, by Theorem 3.4.28, the inequality (3.126) holds.

Using representation (3.119), we also conclude that Λ_n is convex if n is odd and concave if n is even. Then by Sherman's theorem, it holds

$$\sum_{i=1}^{m} b_i \Lambda_k \left(\frac{y_i - \alpha}{\beta - \alpha}\right) \leq \sum_{j=1}^{l} a_j \Lambda_k \left(\frac{x_j - \alpha}{\beta - \alpha}\right) \tag{3.129}$$

if k is odd and the reverse inequality in (3.129) holds if k is even.

Moreover, if $\phi^{(2k)}(\alpha) \geq 0$, $\phi^{(2k)}(\beta) \geq 0$ for $k = 1, 3, ..., n-2$ and $\phi^{(2k)}(\alpha) \leq 0$, $\phi^{(2k)}(\beta) \leq 0$ for $k = 2, 4, ..., n-1$, then the right hand side in (3.124) is nonnegative and (3.127) immediately follows .

(ii) Similar to the part (i). □

3.4.2. Bounds for identities related to generalizations of Sherman's inequality

For the sake of simplicity and to avoid many notations, we define two functions as follows.

Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{a} = (a_1, ..., a_l) \in \mathbb{R}^l$, $\mathbf{b} = (b_1, ..., b_m) \in \mathbb{R}^m$ and the function G_n be defined as in (3.117)-(3.118). We define the function $\mathcal{R} : [\alpha, \beta] \to \mathbb{R}$ by

$$\mathcal{R}(t) = \sum_{j=1}^{l} a_j G_n \left(\frac{x_j - \alpha}{\beta - \alpha}, \frac{t - \alpha}{\beta - \alpha}\right) - \sum_{i=1}^{m} b_i G_n \left(\frac{y_i - \alpha}{\beta - \alpha}, \frac{t - \alpha}{\beta - \alpha}\right). \tag{3.130}$$

Let $T(\mathcal{R}, \mathcal{R})$ denotes the Čebyšev functional of $\mathcal{R}$.

Using theorem for the Čebyšev functional (see Theorem 1.4.12), we obtain an upper bound for the identity (3.120).

Theorem 3.4.30: *Let $n \in \mathbb{N}$, $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi \in C^{2n}([\alpha, \beta])$ with $(\cdot - \alpha)(\beta - \cdot)(\phi^{(2n+1)})^2 \in L[\alpha, \beta]$. Let Λ_k be defined as in (3.116 and $\mathcal{R}$ as in (3.130). Then* *the remainder* $\kappa_n(\phi;\alpha,\beta)$*, defined by*

$$\begin{aligned}\kappa_n(\phi;\alpha,\beta) &= \sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\ &- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\alpha)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{\beta-x_j}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{\beta-y_i}{\beta-\alpha}\right)\right] \\ &- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\beta)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{x_j-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{y_i-\alpha}{\beta-\alpha}\right)\right] \\ &- (\beta-\alpha)^{2n-2}\left[\phi^{(2n-1)}(\beta) - \phi^{(2n-1)}(\alpha)\right]\int_\alpha^\beta \mathcal{R}(t)dt,\end{aligned} \tag{3.131}$$

satisfies the estimation

$$|\kappa_n(\phi;\alpha,\beta)| \le \frac{(\beta-\alpha)^{2n-\frac{1}{2}}}{\sqrt{2}}[T(\mathcal{R},\mathcal{R})]^{\frac{1}{2}}\left(\int_\alpha^\beta (t-\alpha)(\beta-t)[\phi^{(2n+1)}(t)]^2dt\right)^{\frac{1}{2}}. \tag{3.132}$$

Proof: Our proof proceeds similarly to the proof of Theorem 3.1.5. □

The following Grüss type inequality holds.

Theorem 3.4.31: *Let* $\phi:[\alpha,\beta]\to\mathbb{R}$ *be such that* $\phi\in C^{2n}([\alpha,\beta])$ *for* $n\in\mathbb{N}$ *and* $\phi^{(2n+1)}\ge 0$ *on* $[\alpha,\beta]$*. Let the function* $\mathcal{R}$ *be defined as in (3.130). Then the remainder* $\kappa_n^1(\phi;\alpha,\beta)$*, defined by (3.131), satisfies the bound*

$$\left|\kappa_n^1(\phi;\alpha,\beta)\right| \le (\beta-\alpha)^{2n-1}\left\|\mathcal{R}'\right\|_\infty\left[\frac{\phi^{(2n-1)}(\beta)+\phi^{(2n-1)}(\alpha)}{2} - \frac{\phi^{(2n-2)}(\beta)-\phi^{(2n-2)}(\alpha)}{\beta-\alpha}\right]. \tag{3.133}$$

Proof: In the proof, we use a similar technique as the proof of Theorem 3.1.6. □

Next we present related the Ostrowski type inequality.

Theorem 3.4.32: *Suppose that all assumptions of Theorem 3.4.27 hold. Assume that* (p,q) *is a pair of conjugate exponents, that is* $1\le p,q\le\infty$, $1/p+1/q=1$*. Let* $\phi^{(2n)}\in$

$L_p[\alpha, \beta]$. *Then*

$$\begin{aligned}
&\Bigg|\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\
&- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\alpha)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{\beta-x_j}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{\beta-y_i}{\beta-\alpha}\right)\right] \\
&- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\beta)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{x_j-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{y_i-\alpha}{\beta-\alpha}\right)\right]\Bigg| \\
&\leq (\beta-\alpha)^{2n-1}\left\|\phi^{(2n)}\right\|_p \left[\int_\alpha^\beta \left|\sum_{j=1}^{l} a_j G_n\left(\tfrac{x_j-\alpha}{\beta-\alpha}, \tfrac{t-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i G_n\left(\tfrac{y_i-\alpha}{\beta-\alpha}, \tfrac{t-\alpha}{\beta-\alpha}\right)\right|^q dt\right]^{\frac{1}{q}},
\end{aligned} \tag{3.134}$$

The constant on the right-hand side of (3.134) is sharp for $1 < p \leq \infty$ *and the best possible for* $p = 1$.

Proof: Let us denote

$$\mathcal{S}(t) = (\beta-\alpha)^{2n-1}\left[\sum_{j=1}^{l} a_j G_n\left(\frac{x_j-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i G_n\left(\frac{y_i-\alpha}{\beta-\alpha}, \frac{t-\alpha}{\beta-\alpha}\right)\right].$$

Using the identity (3.120) and applying well known Hölder's inequality, we obtain

$$\begin{aligned}
&\Bigg|\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{m} b_i\phi(y_i) \\
&- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\alpha)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{\beta-x_j}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{\beta-y_i}{\beta-\alpha}\right)\right] \\
&- \sum_{k=0}^{n-1}(\beta-\alpha)^{2k}\phi^{(2k)}(\beta)\left[\sum_{j=1}^{l} a_j\Lambda_k\left(\frac{x_j-\alpha}{\beta-\alpha}\right) - \sum_{i=1}^{m} b_i\Lambda_k\left(\frac{y_i-\alpha}{\beta-\alpha}\right)\right]\Bigg| \\
&= \left|\int_\alpha^\beta \mathcal{S}(t)\phi^{(2n)}(t)dt\right| \leq \left\|\phi^{(2n)}\right\|_p \left(\int_\alpha^\beta |\mathcal{S}(t)|^q\, dt\right)^{\frac{1}{q}}.
\end{aligned}$$

The proof of the sharpness is analogous to one in proof of Theorem 3.1.7. □

3.4.3. Related mean value theorems with the application of the exp-convex method

Motivated by the inequality (3.124), we define under the assumptions of Theorem 3.4.28, equipped with condition (3.123), the linear functional as follows:

$$\begin{aligned} A(\phi) &= \sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{m} b_i \phi(y_i) \\ &- \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\alpha) \left[\sum_{j=1}^{l} a_j \Lambda_k \left(\frac{\beta - x_j}{\beta - \alpha} \right) - \sum_{i=1}^{m} b_i \Lambda_k \left(\frac{\beta - y_i}{\beta - \alpha} \right) \right] \\ &- \sum_{k=0}^{n-1} (\beta-\alpha)^{2k} \phi^{(2k)}(\beta) \left[\sum_{j=1}^{l} a_j \Lambda_k \left(\frac{x_j - \alpha}{\beta - \alpha} \right) - \sum_{i=1}^{m} b_i \Lambda_k \left(\frac{y_i - \alpha}{\beta - \alpha} \right) \right]. \end{aligned} \tag{3.135}$$

Remark 3.4.13.: It should be noticed that if $\phi : [\alpha, \beta] \to \mathbb{R}$ is $2n$-convex, by Theorem 3.4.28, it holds

$$A(\phi) \geq 0.$$

Theorem 3.4.33: *Let A be the linear functional defined as in (3.135). Let $\phi \in C^{2n}([\alpha, \beta])$. Then there exists $\xi \in [\alpha, \beta]$ such that*

$$A(\phi) = \phi^{(2n)}(\xi) A(\varphi_0),$$

where $\varphi_0(x) = x^{2n}/(2n)!$.

Proof: The proof technique is similar to the proof of Theorem 3.1.8. □

Theorem 3.4.34: *Let A be the linear functional defined as in (3.135). Let $\phi, \psi \in C^{2n}([\alpha, \beta])$. Then there exists $\xi \in [\alpha, \beta]$ such that*

$$\frac{\phi^{(2n)}(\xi)}{\psi^{(2n)}(\xi)} = \frac{A(\phi)}{A(\psi)}, \tag{3.136}$$

assuming neither of the denominators is equal to zero.

Proof: We prove the theorem similarly to the proof of Theorem 3.1.9. □

Remark 3.4.14.: If $\frac{\phi^{(2n)}}{\psi^{(2n)}}$ is an invertible function, then from (3.136) we get new mean of Cauchy's type of the interval $[\alpha, \beta]$,

$$\xi = \left(\frac{\phi^{(2n)}}{\psi^{(2n)}}\right)^{-1} \left(\frac{A(\phi)}{A(\psi)}\right).$$

Through the rest of this subsection, I denotes an open interval in $\mathbb{R}$.

In order to obtain results regarding the exponential convexity, we define the families of functions as follows.

Definition 3.4.5.: Let I be an open interval in $\mathbb{R}$. For every choice of $2l+1$ mutually different points $z_0, z_1, ..., z_{2l} \in [\alpha, \beta]$ we define:

- $\mathcal{F}_1 = \{f_t : [\alpha, \beta] \to \mathbb{R} : t \in I$ and $t \mapsto [z_0, z_1, ..., z_{2l}; f_t]$ is n-exponentially convex in the Jensen sense on $I\}$;
- $\mathcal{F}_2 = \{f_t : [\alpha, \beta] \to \mathbb{R} : t \in I$ and $t \mapsto [z_0, z_1, ..., z_{2l}; f_t]$ is exponentially convex in the Jensen sense on $I\}$;
- $\mathcal{F}_3 = \{f_t : [\alpha, \beta] \to \mathbb{R} : t \in I$ and $t \mapsto [z_0, z_1, ..., z_{2l}; f_t]$ is 2-exponentially convex in the Jensen sense on $I\}$.

Theorem 3.4.35: *Let A be the linear functional defined as in (3.135) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$). Then the following statements hold:*

(i) The function $t \mapsto A(f_t)$ is n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) If the function $t \mapsto A(f_t)$ is continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I.

Proof: *Our proof proceeds similarly to the proof of Theorem 3.1.10.* □

Corollary 3.4.8.: *Let A be the linear functional defined as in (3.135) associated with the family $\mathcal{F}_3$. Then the following statements hold:*

(i) If the function $t \mapsto A(f_t)$ is continuous on I, then it is 2-exponentially convex on I. If $t \mapsto A(f_t)$ is additionally positive, then it is also log-convex on I. Furthermore, for every choice $r, s, t \in I$, such that $r < s < t$, it holds

$$[A(f_s)]^{t-r} \leq [A(f_r)]^{t-s} [A(f_r)]^{s-r} .$$

(ii) If the function $t \mapsto A(f_t)$ is positive and differentiable on I, then for all $r, s, u, v \in I$ such that $r \leq u$, $s \leq v$, we have

$$M_{r,s}(A, \mathcal{F}_3) \leq M_{u,v}(A, \mathcal{F}_3),$$

where

$$M_{r,s}(A, \mathcal{F}_3) = \begin{cases} \left(\frac{A(f_r)}{A(f_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(\frac{\frac{d}{dr}(A(f_r))}{A(f_r)}\right), & r = s. \end{cases} \tag{3.137}$$

Proof: (i) The first part of statement is an easy consequence of Theorem 3.4.35 and the second one of Remark 1.2.10..

Since the function $t \mapsto A(f_t)$ is log-convex on I, i.e. the function $t \mapsto \log A(f_t)$ is convex on I, then applying Lemma 1.1.1. a) we have

$$(r-t)\log A(f_t) + (t-s)\log A(f_r) + (s-r)\log A(f_r) \geq 0$$

for every choice $r, s, t \in I$, such that $r < s < t$. Therefore, we have

$$[A(f_s)]^{t-r} \leq [A(f_r)]^{t-s} [A(f_r)]^{s-r} .$$

(ii) Applying Lemma 1.1.1. b) to the convex function $t \mapsto \log A(f_t)$, we get

$$\frac{\log A(f_r) - \log A(f_s)}{r-s} \leq \frac{\log A(f_u) - \log A(f_v)}{u-v} \tag{3.138}$$

for $r \leq u$, $s \leq v$, $r \neq u$, $s \neq v$. Therefore, we have

$$M_{r,s}(A, \mathcal{F}_3) \leq M_{u,v}(A, \mathcal{F}_3) .$$

Case $r = s$, $u = v$ follows from (3.138) as limiting case. $\square$

Remark 3.4.15.: Note that, with assumption that the functions from $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ are differentiable, the results from Theorem 3.4.35 and Corollary 3.4.8. still hold when two of the points $z_0, z_1, ..., z_{2l} \in [\alpha, \beta]$ coincide. Further, if the functions from $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ are $2l$-times differentiable, the results still hold when all points coincide. These can be easily proved using (3.42) and some fact of the exponential convexity.

3.4.4. New families of exponentially convex functions with applications to means

Using few families of convex functions which are given below, we construct different examples of exponentially convex functions. As consequences, applying mean-value theorem of Cauchy type from the previous section to these special families of functions, we establish new classes of two-parameter Cauchy-type means that are symmetric and have monotone properties over both parameters.

Throughout this subsection id denotes the identity function, i.e. $id(x) = x$ for each $x \in \mathbb{R}$.

Example 7.: *Consider the family of functions*

$$\Omega_1 = \{\varphi_t : (0, \infty) \to \mathbb{R} : t \in \mathbb{R}\}$$

defined by

$$\varphi_t(x) = \begin{cases} \frac{x^t}{t(t-1)\cdots(t-2n+1)}, & t \notin \{0, 1, ..., 2n-1\}, \\ \frac{x^j \log x}{(-1)^{2n-1-j} j!(2n-1-j)!}, & t = j \in \{0, 1, ..., 2n-1\}. \end{cases}$$

Since $\frac{d^{2n}\varphi_t}{dx^{2n}}(x) = x^{t-2n} = e^{(t-2n)\log x} > 0$, $t \in \mathbb{R}$, it follows that φ_t is $2n$-convex on $(0, \infty)$ for every $t \in \mathbb{R}$ and $t \mapsto \frac{d^{2n}\varphi_t}{dx^{2n}}(x)$ is exponentially convex (for more explanation see [77]). Therefore, using the same arguments as in proof of Theorem 3.4.35 we conclude that the function $t \mapsto [z_0, z_1, ..., z_{2l}; \varphi_t]$ is exponentially convex (and so exponentially convex in the

Jensen sense). Also, it follows that $t \mapsto A(\varphi_t)$ *is exponentially convex in the Jensen sense. It is easy to verify that the function* $t \mapsto A(\varphi_t)$ *is continuous so then also exponentially convex. In this case we assume that* $[\alpha, \beta] \subset (0, \infty)$.

For this family of functions, with assumption that $t \mapsto A(\varphi_t)$ *is positive, (3.137) becomes*

$$M_{r,s}(A, \Omega_1) = \begin{cases} \left(\frac{A(\varphi_r)}{A(\varphi_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(-(2n-1)!\frac{A(\varphi_r \cdot \varphi_0)}{A(\varphi_r)} + \sum\limits_{i=0}^{2n-1} \frac{1}{i-r}\right), & r = s \notin \{0, 1, ..., 2n-1\}, \\ \exp\left(-(2n-1)!\frac{A(\varphi_r \cdot \varphi_0)}{2A(\varphi_r)} + \sum\limits_{i=0, i\neq r}^{2n-1} \frac{1}{i-r}\right), & r = s \in \{0, 1, ..., 2n-1\}. \end{cases} \tag{3.139}$$

Extensions are obtained by continuity for a different choice of parameters $r, s \in \mathbb{R}$.

We establish a new class of two-parameter Cauch-type means which is monotonic over both parameters.

Example 8.: *Consider the family of functions*

$$\Omega_2 = \{\psi_t : \mathbb{R} \to [0, \infty) : t \in \mathbb{R}\}$$

defined by

$$\psi_t(x) = \begin{cases} \frac{e^{tx}}{t^{2n}}, & t \neq 0, \\ \frac{x^{2n}}{(2n)!}, & t = 0. \end{cases}$$

Since $\frac{d^{2n}\psi_t}{dx^{2n}}(x) = e^{tx} > 0$, $t \in \mathbb{R}$, *it follows that* ψ_t *is* $2n$*-convex on* $\mathbb{R}$ *for every* $t \in \mathbb{R}$ *and* $t \mapsto \frac{d^{2n}\psi_t}{dx^{2n}}(x) = e^{tx}$ *is exponentially convex by definition. Therefore, we have that the function* $t \mapsto A(\psi_t)$ *is exponentially convex.*

For this family of functions, with assumption that $t \mapsto A(\psi_t)$ *is positive, (3.137) becomes*

$$M_{r,s}(A, \Omega_2) = \begin{cases} \left(\frac{A(\psi_r)}{A(\psi_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(\frac{A(id \cdot \psi_r)}{A(\psi_r)} - \frac{2n}{r}\right), & r = s \neq 0, \\ \exp\left(\frac{1}{2n+1}\frac{A(id \cdot \psi_0)}{A(\psi_0)}\right), & r = s = 0. \end{cases}$$

Applying Theorem 3.4.34 to this family of functions, it follows that

$$\overline{M}_{r,s}\left(A,\Omega_2\right) := \log M_{r,s}\left(A,\Omega_2\right),$$

which presents a new class of two-parameter Cauchy-type means of interval $[\alpha,\beta]$. *Monotonicity over both parameters follows from (ii)-part of Corollary 3.4.8..*

Example 9.: *Consider the family of functions*

$$\Omega_3 = \{\lambda_t : (0,\infty) \to (0,\infty) : t \in (0,\infty)\}$$

defined by

$$\lambda_t(x) = \begin{cases} \frac{t^{-x}}{(\log t)^{2n}}, & t \neq 1, \\ \frac{x^{2n}}{(2n)!}, & t = 1. \end{cases}$$

Since $t \mapsto \frac{d^{2n}\lambda_t}{dx^{2n}}(x) = t^{-x} > 0$, $t \in (0,\infty)$, *it follows that* λ_t *is* $2n$*-convex on* $(0,\infty)$ *for every* $t \in (0,\infty)$ *and* $t \mapsto \frac{d^{2n}\lambda_t}{dx^{2n}}(x) = t^{-x}$ *is exponentially convex since it is the Laplace transform of a non-negative function (for more explanation see [77], [146]). Therefore, we have that the function* $t \mapsto A(\lambda_t)$ *is exponentially convex. In this case we assume that* $[\alpha,\beta] \subset (0,\infty)$.

For this family of functions, with assumption that $t \mapsto A(\varphi_t)$ *is positive, (3.137) becomes*

$$M_{r,s}\left(A,\Omega_3\right) = \begin{cases} \left(\frac{A(\lambda_r)}{A(\lambda_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(-\frac{A(id\cdot\lambda_r)}{rA(\lambda_r)} - \frac{2n}{r\log r}\right), & r = s \neq 1, \\ \exp\left(-\frac{1}{2n+1}\frac{A(id\cdot\lambda_1)}{A(\lambda_1)}\right), & r = s = 1. \end{cases}$$

For this family of functions, applying Theorem 3.4.34, we define a new class of two-parameter Cauchy-type means as follows

$$\overline{M}_{r,s}\left(A,\Omega_3\right) := -L(r,s)\log M_{r,s}\left(A,\Omega_3\right).$$

Here $L(r,s)$ presents logarithmic mean which is defined as follows

$$L(r,s) = \begin{cases} \frac{r-s}{\log r - \log s}, & r \neq s, \\ s, & r = s. \end{cases}$$

Monotonicity with respect to the both parameters follows from (ii)-part of Corollary 3.4.8..

Example 10.: *Consider the family of functions*

$$\Omega_4 = \{\kappa_t : (0,\infty) \to (0,\infty) : t \in (0,\infty)\}$$

defined by

$$\kappa_t(x) = \frac{e^{-x\sqrt{t}}}{t^n}.$$

Since $t \mapsto \frac{d^{2n}\kappa_t}{dx^{2n}}(x) = e^{-x\sqrt{t}}$, $t \in (0,\infty)$, it follows that κ_t is $2n$-convex on $(0,\infty)$ for every $t \in (0,\infty)$ and $t \mapsto \frac{d^{2n}\kappa_t}{dx^{2n}}(x) = e^{-x\sqrt{t}}$ is exponentially convex since it is the Laplace transform of a non-negative function (for more details see [77], [146]). Therefore, using the same arguments as in the previous example, we have that the function $t \mapsto A(\kappa_t)$ is exponentially convex. In this case we assume that $[\alpha,\beta] \subset (0,\infty)$.

For this family of function, with assumption that $t \mapsto A(\varphi_t)$ is positive, (3.137) becomes

$$M_{r,s}(A,\Omega_4) = \begin{cases} \left(\frac{A(\kappa_r)}{A(\kappa_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(-\frac{A(id\cdot\kappa_r)}{2\sqrt{r}A(\kappa_r)} - \frac{n}{r}\right), & r = s. \end{cases}$$

For this family of functions, applying Theorem 3.4.34, we define a new class of two-parameter Cauchy-type means as follows

$$\overline{M}_{r,s}(A,\Omega_4) := -(\sqrt{r}+\sqrt{s})\log M_{r,s}(A,\Omega_3).$$

Monotonicity with respect to the both parameters follows from (ii)-part of Corollary 3.4.8..

3.5. Sherman's inequality and Hermite's interpolation polynomials

In this section, we give generalizations of Sherman's inequality obtained by applying interpolation with Hermite interpolating polynomials. We discuss the results for particular cases. Namely, we consider interpolation for particular cases known as Lagrange interpolating polynomials, $(m, n-m)$ interpolating polynomials and two-point Taylor interpolating polynomials. We give bounds for the identities related to new generalizations in the form of the Čebyšev functionals. We also give generalized Grüss and Ostrowski type inequalities related to obtained generalizations. At the end of the section, applying exp-convex method we construct new families of exponentially and logarithically convex functions and related Stolarsky type means. The results from this section are quoted from [6].

We start this section with the following Hermite interpolating polynomial representations by using notations and terminology from [11].

Let $\alpha \leq a_1 < a_2 < \cdots < a_r \leq \beta$, $(r \geq 2)$ be the given points. For $\phi \in C^n([\alpha, \beta])$ $(n \geq r)$, a unique polynomial $\rho_H(s)$ of degree $(n-1)$ exists, fulfilling one of the following **Hermite conditions**:

$$\rho_H^{(i)}(a_j) = \phi^{(i)}(a_j);\ 0 \leq i \leq k_j,\ 1 \leq j \leq r,\ \sum_{j=1}^{r} k_j + r = n. \tag{H}$$

In particular:

a) **Lagrange conditions** (for $r = n$, $k_j = 0$ for all j):

$$\rho_L(a_j) = \phi(a_j), \quad 1 \leq j \leq n.$$

b) **Type** $(m, n-m)$ **conditions** ($r = 2, 1 \leq m \leq n-1,\ k_1 = m-1,\ k_2 = n-m-1$):

$$\rho_{(m,n)}^{(i)}(\alpha) = \phi^{(i)}(\alpha), \quad 0 \leq i \leq m-1,$$

$$\rho_{(m,n)}^{(i)}(\beta) = \phi^{(i)}(\beta), \quad 0 \leq i \leq n-m-1.$$

c) **Two-point Taylor conditions** ($n = 2m$, $r = 2$, $k_1 = k_2 = m - 1$):

$$\rho_{2T}^{(i)}(\alpha) = \phi^{(i)}(\alpha),\ \rho_{2T}^{(i)}(\beta) = \phi^{(i)}(\beta), \quad 0 \le i \le m-1.$$

Theorem 3.5.36: *Let $-\infty < \alpha < \beta < \infty$ and $\alpha \le a_1 < a_2 < \cdots < a_r \le \beta$ be r ($r \ge 2$) distinct points and $\phi \in C^n([\alpha,\beta])$. Then*

$$\phi(t) = \rho_H(t) + R_{H,n}(\phi, t), \tag{3.140}$$

where $\rho_H(t)$ is the Hermite inrepolating polynomial, i.e.

$$\rho_H(t) = \sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i)}(a_j),$$

H_{ij} are fundamental polynomials of the Hermite basis defined by

$$H_{ij}(t) = \frac{1}{i!}\frac{\omega(t)}{(t-a_j)^{k_j+1-i}}\sum_{k=0}^{k_j-i}\frac{1}{k!}\frac{d^k}{dt^k}\left(\frac{(t-a_j)^{k_j+1}}{\omega(t)}\right)\Bigg|_{t=a_j}(t-a_j)^k, \tag{3.141}$$

with

$$\omega(t) = \prod_{j=1}^{r}(t-a_j)^{k_j+1}, \tag{3.142}$$

and the remainder is given by

$$R_{H,n}(\phi,t) = \int_\alpha^\beta G_{H,n}(t,s)\phi^{(n)}(s)ds,$$

where $G_{H,n}(t,s)$ is defined by

$$G_{H,n}(t,s) = \begin{cases} \sum\limits_{j=1}^{l}\sum\limits_{i=0}^{k_j}\frac{(a_j-s)^{n-i-1}}{(n-i-1)!}H_{ij}(t); & s \le t, \\ -\sum\limits_{j=l+1}^{r}\sum\limits_{i=0}^{k_j}\frac{(a_j-s)^{n-i-1}}{(n-i-1)!}H_{ij}(t); & s \ge t, \end{cases} \tag{3.143}$$

for all $a_l \le s \le a_{l+1}$, $l = 0, \ldots, r$ with $a_0 = \alpha$ and $a_{r+1} = \beta$.

The associated error $R_{H,n}(\phi,t)$ is represented in the terms of Green's function $G_{H,n}(t,s)$

(3.143) for multipoint boundary value problem

$$z^{(n)}(t) = 0,\ z^{(i)}(a_j) = 0,\ 0 \le i \le k_j,\ 1 \le j \le r.$$

Remark 3.5.16.: a) For Lagrange conditions, from Theorem 3.5.36 we have

$$\phi(t) = \rho_L(t) + R_L(\phi, t)$$

where $\rho_L(t)$ is the Lagrange interpolating polynomial i.e.

$$\rho_L(t) = \sum_{j=1}^{n} \prod_{\substack{k=1\\k\neq j}}^{n} \left(\frac{t-a_k}{a_j-a_k}\right)\phi(a_j)$$

and the remainder $R_L(\phi, t)$ is given by

$$R_L(\phi, t) = \int_\alpha^\beta G_L(t,s)\phi^{(n)}(s)ds$$

with

$$G_L(t,s) = \frac{1}{(n-1)!}\begin{cases} \sum\limits_{j=1}^{l}(a_j - s)^{n-1}\prod\limits_{\substack{k=1\\k\neq j}}^{n}\left(\frac{t-a_k}{a_j-a_k}\right), & s \le t \\ -\sum\limits_{j=l+1}^{n}(a_j - s)^{n-1}\prod\limits_{\substack{k=1\\k\neq j}}^{n}\left(\frac{t-a_k}{a_j-a_k}\right), & s \ge t \end{cases}, \tag{3.144}$$

$a_l \le s \le a_{l+1}$, $l = 1, 2, ..., n-1$ with $a_1 = \alpha$ and $a_n = \beta$.

b) For type $(m, n-m)$ conditions, from Theorem 3.5.36 we have

$$\phi(t) = \rho_{(m,n)}(t) + R_{(m,n)}(\phi, t),$$

where $\rho_{(m,n)}(t)$ is $(m, n-m)$ interpolating polynomial, i.e.

$$\rho_{(m,n)}(t) = \sum_{i=0}^{m-1}\tau_i(t)\phi^{(i)}(\alpha) + \sum_{i=0}^{n-m-1}\eta_i(t)\phi^{(i)}(\beta),$$

with

$$\tau_i(t) = \frac{1}{i!}(t-\alpha)^i\Big(\frac{t-\beta}{\alpha-\beta}\Big)^{n-m}\sum_{p=0}^{m-1-i}\binom{n-m+p-1}{p}\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^p, \tag{3.145}$$

$$\eta_i(t) = \frac{1}{i!}(t-\beta)^i\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^{m}\sum_{p=0}^{n-m-1-i}\binom{m+p-1}{p}\Big(\frac{t-\beta}{\alpha-\beta}\Big)^p, \tag{3.146}$$

and the remainder is given by

$$R_{(m,n)}(\phi,t) = \int_\alpha^\beta G_{(m,n)}(t,s)\phi^{(n)}(s)ds$$

with

$$G_{(m,n)}(t,s) = \begin{cases} \sum\limits_{j=0}^{m-1}\Big[\sum\limits_{p=0}^{m-1-j}\binom{n-m+p-1}{p}\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^p\Big]\times \\ \frac{(t-\alpha)^j(\alpha-s)^{n-j-1}}{j!(n-j-1)!}\Big(\frac{\beta-t}{\beta-\alpha}\Big)^{n-m}, & \alpha\le s\le t\le\beta \\ -\sum\limits_{i=0}^{n-m-1}\Big[\sum\limits_{q=0}^{n-m-i-1}\binom{m+q-1}{q}\Big(\frac{\beta-t}{\beta-\alpha}\Big)^q\Big]\times \\ \frac{(t-\beta)^i(\beta-s)^{n-i-1}}{i!(n-i-1)!}\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^{m}, & \alpha\le t\le s\le\beta. \end{cases} \tag{3.147}$$

c) For Two-point Taylor conditions, from Theorem 3.5.36 we have

$$\phi(t) = \rho_{2T}(t) + R_{2T}(\phi,t),$$

where $\rho_{2T}(t)$ is the two-point Taylor interpolating polynomial i.e,

$$\begin{aligned}\rho_{2T}(t) = \sum_{i=0}^{m-1}\sum_{p=0}^{m-1-i}\binom{m+p-1}{p}\Big[&\frac{(t-\alpha)^i}{i!}\Big(\frac{t-\beta}{\alpha-\beta}\Big)^m\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^p\phi^{(i)}(\alpha) \\ &+\frac{(t-\beta)^i}{i!}\Big(\frac{t-\alpha}{\beta-\alpha}\Big)^m\Big(\frac{t-\beta}{\alpha-\beta}\Big)^p\phi^{(i)}(\beta)\Big]\end{aligned} \tag{3.148}$$

and the remainder is given by

$$R_{2T}(\phi,t) = \int_\alpha^\beta G_{2T}(t,s)\phi^{(n)}(s)ds$$

with

$$G_{2T}(t,s)=\begin{cases}\frac{(-1)^m}{(2m-1)!}p^m(t,s)\sum\limits_{j=0}^{m-1}\binom{m-1+j}{j}(t-s)^{m-1-j}q^j(t,s), & s\le t\\ \frac{(-1)^m}{(2m-1)!}q^m(t,s)\sum\limits_{j=0}^{m-1}\binom{m-1+j}{j}(s-t)^{m-1-j}p^j(t,s), & s\ge t\end{cases}, \tag{3.149}$$

where $p(t,s)=\frac{(s-\alpha)(\beta-t)}{\beta-\alpha}$, $q(t,s)=p(s,t)$, $\forall\, t,s\in[\alpha,\beta]$.

The following lemma describes the positivity of $G_{H,n}(t,s)$ (see [20], [161]).

Lemma 3.5.3.: *Let ω be given by (3.142). The function $G_{H,n}(t,s)$, defined by (3.143), has the following properties:*

(i) $\frac{G_{H,n}(t,s)}{\omega(t)}>0,\ a_1\le t\le a_r,\ a_1<s<a_r;$

(ii) $G_{H,n}(t,s)\le\frac{1}{(n-1)!(\beta-\alpha)}|\omega(t)|;$

(iii) $\int_\alpha^\beta G_{H,n}(t,s)ds=\frac{\omega(t)}{n!}.$

3.5.1. Generalized Sherman's inequality by Hermite interpolation polynomials

First we prove new identity by using interpolation with Hermite interpolating polynomials.

Theorem 3.5.37: *Let $\phi\in C^n([\alpha,\beta])$ and let $\infty<\alpha\le a_1<a_2<\cdots<a_r\le\beta<\infty$, $(r\ge 2)$ be the given points, $k_j\in\mathbb{N}$, $1\le j\le r$, $\sum\limits_{j=1}^{r}k_j+r=n$. Let $\mathbf{x}=(x_1,...,x_l)\in[\alpha,\beta]^l$, $\mathbf{y}=(y_1,...,y_m)\in[\alpha,\beta]^m$, $\mathbf{u}=(u_1,...,u_l)\in\mathbb{R}^l$ and $\mathbf{v}=(v_1,...,v_m)\in\mathbb{R}^m$.*

Then

$$\sum_{p=1}^{l}u_p\phi(x_p)-\sum_{q=1}^{m}v_q\phi(y_q)=\sum_{j=1}^{r}\sum_{i=0}^{k_j}\phi^{(i)}(a_i)\left[\sum_{p=1}^{l}u_pH_{ij}(x_p)-\sum_{q=1}^{m}v_qH_{ij}(y_q)\right] \tag{3.150}$$
$$+\int_\alpha^\beta\left[\sum_{p=1}^{l}u_pG_{H,n}(x_p,s)-\sum_{q=1}^{m}v_qG_{H,n}(y_q,s)\right]\phi^{(n)}(s)ds,$$

where H_{ij} and $G_{H,n}$ are defined as in (3.141) and (3.143), respectively.

Proof: By an easy calculation, using (3.140) in $\sum\limits_{p=1}^{l}u_p\phi(x_p)-\sum\limits_{q=1}^{m}v_q\phi(y_q)$ we obtain

(3.150). □

In the following theorem, under Hermite conditions, we obtain new generalization of Sherman's inequality.

Theorem 3.5.38: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be n-convex and let $\infty < \alpha \leq a_1 < a_2 < \cdots < a_r \leq \beta < \infty$, $(r \geq 2)$ be the given points, $k_j \in \mathbb{N},\ 1 \leq j \leq r$, $\sum\limits_{j=1}^{r} k_j + r = n$. Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{u} = (u_1, ..., u_l) \in \mathbb{R}^l$ and $\mathbf{v} = (v_1, ..., v_m) \in \mathbb{R}^m$.*
If

$$\sum_{p=1}^{l} u_p G_{H,n}(x_p, s) - \sum_{q=1}^{m} v_q G_{H,n}(y_q, s) \geq 0, \quad s \in [\alpha, \beta], \tag{3.151}$$

then

$$\begin{aligned} &\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) \\ &\geq \sum_{p=1}^{l} u_p \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) H_{ij}(x_p) - \sum_{q=1}^{m} v_q \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) H_{ij}(y_q), \end{aligned} \tag{3.152}$$

where H_{ij} and $G_{H,n}$ are defined as in (3.141) and (3.143), respectively.

Proof: Since the function ϕ is n-convex, then without loss of generality we can assume that ϕ is n-times differentiable and $\phi^{(n)} \geq 0$ (see [138, p. 16 and p. 293]). Hence, we can apply Theorem 3.5.37 to obtain (3.152). □

Specially, if we assume that weights $\mathbf{u}, \mathbf{v}$ and matrix $\mathbf{A}$ are such that satified Sherman's conditions, i.e. with nonnegative entries and (3.153) holds, then we get the next generalization.

Theorem 3.5.39: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be n-convex and let $-\infty < \alpha \leq a_1 < a_2 < \cdots < a_r \leq \beta < \infty$, $(r \geq 2)$ be the given points, $k_j \in \mathbb{N}$, $1 \leq j \leq r$, $\sum\limits_{j=1}^{r} k_j + r = n$. Let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{u} = (u_1, ..., u_l) \in [0, \infty)^l$ and $\mathbf{v} = (v_1, ..., v_m) \in [0, \infty)^m$ be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \text{and} \quad \mathbf{u} = \mathbf{v}\mathbf{A} \tag{3.153}$$

holds for some row stochastic matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$.

If (3.152) holds and the function

$$\bar{F}(\cdot) = \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) H_{ij}(\cdot) \tag{3.154}$$

is convex on $[\alpha, \beta]$, *then*

$$\sum_{q=1}^{m} v_q \phi(y_q) \leq \sum_{p=1}^{l} u_p \phi(x_p). \tag{3.155}$$

Proof: Since the inequality (3.152) holds, the right hand side of (3.152) can be rewriting in the form

$$\sum_{p=1}^{l} u_p \bar{F}(x_p) - \sum_{q=1}^{m} v_q \bar{F}(y_q),$$

where $\bar{F}$ is defined by (3.154). If $\bar{F}$ is convex, then by Sherman's theorem we have

$$\sum_{p=1}^{l} u_p \bar{F}(x_p) - \sum_{q=1}^{m} v_q \bar{F}(y_q) \geq 0,$$

i.e. the right hand side of (3.152) is nonnegative so the inequality (3.155) immediately follows. □

As a direct consequence of the previous two theorems, by using Lagrange conditions we get the following generalization of Sherman's theorem.

Corollary 3.5.9.: *Let* $\phi : [\alpha, \beta] \to \mathbb{R}$ *be* n*-convex and let* $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{u} = (u_1, ..., u_l) \in [0, \infty)^l$ *and* $\mathbf{v} = (v_1, ..., v_m) \in [0, \infty)^m$ *be such that (3.153) holds for some row stochastic matrix* $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$.

(i) If

$$\sum_{p=1}^{l} u_p G_L(x_p, s) - \sum_{q=1}^{m} v_q G_L(y_q, s) \geq 0, \quad s \in [\alpha, \beta],$$

then

$$\begin{aligned}
&\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{m} v_q\phi(y_q) \qquad (3.156)\\
&\geq \sum_{p=1}^{l} u_p \sum_{j=1}^{n} \phi(a_j) \prod_{\substack{k=1\\k\neq j}}^{n} \left(\frac{x_p - a_k}{a_j - a_k}\right) - \sum_{q=1}^{m} v_q \sum_{j=1}^{n} \phi(a_j) \prod_{\substack{k=1\\k\neq j}}^{n} \left(\frac{y_q - a_k}{a_j - a_k}\right),
\end{aligned}$$

where G_L is defined as in (3.144).

(ii) If (3.156) holds and the function

$$\tilde{F}(\cdot) = \sum_{j=1}^{n} \phi(a_j) \prod_{\substack{k=1\\k\neq j}}^{n} \left(\frac{\cdot - a_k}{a_j - a_k}\right)$$

is convex on $[\alpha, \beta]$, then

$$\sum_{q=1}^{m} v_q\phi(y_q) \leq \sum_{p=1}^{l} u_p\phi(x_p).$$

By using type $(m, n-m)$ conditions we can give the following result.

Corollary 3.5.10.: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be n-convex and let $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_u) \in [\alpha, \beta]^u$, $\mathbf{u} = (u_1, ..., u_l) \in [0, \infty)^l$ and $\mathbf{v} = (v_1, ..., v_u) \in [0, \infty)^u$ be such that (3.153) holds for some row stochastic matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ul}(\mathbb{R})$.*

(i) If

$$\sum_{p=1}^{l} u_p G_{(m,n)}(x_p, s) - \sum_{q=1}^{u} v_q G_{(m,n)}(y_q, s) \geq 0, \quad s \in [\alpha, \beta],$$

then

$$\begin{aligned}
&\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{u} v_q\phi(y_q) \qquad (3.157)\\
&\geq \sum_{p=1}^{l} u_p \left(\sum_{i=0}^{m-1} \tau_i(x_p)\phi^i(\alpha) + \sum_{i=0}^{n-m-1} \eta_i(x_p)\phi^i(\beta) \right)\\
&- \sum_{q=1}^{u} v_q \left(\sum_{i=0}^{m-1} \tau_i(y_q)\phi^i(\alpha) + \sum_{i=0}^{n-m-1} \eta_i(y_q)\phi^i(\beta) \right),
\end{aligned}$$

where τ_i, η_i *and* $G_{(m,n)}$ *are defined as in (3.145), (3.146) and (3.147), respectively.*

(ii) If (3.157) holds and the function

$$\hat{F}(\cdot) = \sum_{i=0}^{m-1} \tau_i(\cdot)\phi^i(\alpha) + \sum_{i=0}^{n-m-1} \eta_i(\cdot)\phi^i(\beta)$$

is convex on $[\alpha, \beta]$, *then*

$$\sum_{p=1}^{l} u_p\phi(x_p) \le \sum_{q=1}^{u} v_q\phi(y_q).$$

By using Two-point Taylor conditions we can give the following result.

Corollary 3.5.11.: *Let* $\phi : [\alpha, \beta] \to \mathbb{R}$ *be* $2m$*-convex and let* $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_u) \in [\alpha, \beta]^u$, $\mathbf{u} = (u_1, ..., u_l) \in [0, \infty)^l$ *and* $\mathbf{v} = (v_1, ..., v_u) \in [0, \infty)^u$ *be such that (3.153) holds for some row stochastic matrix* $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{ul}(\mathbb{R})$.

(i) If

$$\sum_{p=1}^{l} u_p G_{2T}(x_p, s) - \sum_{q=1}^{u} v_q G_{2T}(y_q, s) \ge 0, \quad s \in [\alpha, \beta],$$

then

$$\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{u} v_q\phi(y_q) \ge \sum_{p=1}^{l} u_p\rho_{2T}(x_p) - \sum_{q=1}^{u} v_q\rho_{2T}(y_q), \tag{3.158}$$

where ρ_{2T} *and* G_{2T} *are defined as in (3.148) and (3.149), respectively.*

(ii) If (3.158) holds and the function ρ_{2T} *is convex on* $[\alpha, \beta]$, *then*

$$\sum_{q=1}^{u} v_q\phi(y_q) \le \sum_{p=1}^{l} u_p\phi(x_p).$$

3.5.2. The Ostrowski and Grüss type inequalities

For the sake of simplicity and to avoid many notations, we define the function $\mathfrak{R} : [\alpha, \beta] \to \mathbb{R}$ by

$$\mathfrak{R}(s) = \sum_{p=1}^{l} u_p G_{H,n}(x_p, s) - \sum_{q=1}^{m} v_q G_{H,n}(y_q, s), \tag{3.159}$$

where $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{u} = (u_1, ..., u_l) \in \mathbb{R}^l$, $\mathbf{v} = (v_1, ..., v_m) \in \mathbb{R}^m$ and the function $G_{H,n}$ is defined by (3.143). We also consider the the Čebyšev functional $T(\mathfrak{R}, \mathfrak{R})$ for the function $\mathfrak{R}$.

The following bounds for identities related to generalization of Sherman's inequality holds.

Theorem 3.5.40: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi \in C^n([\alpha, \beta])$ with $(\cdot - \alpha)(\beta - \cdot)(\phi^{(n+1)})^2 \in L[\alpha, \beta]$. Let the function $\mathfrak{R}$ be defined as in (3.159). Then the remainder $\kappa(\phi; \alpha, \beta)$, defined by*

$$\begin{aligned}\kappa(\phi; \alpha, \beta) = &\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) - \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) \left[\sum_{p=1}^{l} u_p H_{ij}(x_p) - \sum_{q=1}^{m} v_q H_{ij}(y_q) \right] \\ &- \frac{\phi^{(n-1)}(\beta) - \phi^{(n-1)}(\alpha)}{\beta - \alpha} \int_{\alpha}^{\beta} \mathfrak{R}(s) ds + \kappa(\phi; \alpha, \beta), \end{aligned} \tag{3.160}$$

satisfies the estimation

$$|\kappa(\phi; \alpha, \beta)| \leq \frac{\sqrt{\beta - \alpha}}{\sqrt{2}} \left[T(\mathfrak{R}, \mathfrak{R}) \right]. \tag{3.161}$$

Proof: The idea of the proof is the same as the proof of Theorem 3.1.5. □

Using Theorem 1.4.13 we obtain the Grüss type inequalities.

Theorem 3.5.41: *Let $\phi \in C^n([\alpha, \beta])$ such that $\phi^{(n)}$ is monotonic nondecreasing on $[\alpha, \beta]$ and let $\mathfrak{R}$ be defined by (3.159). Then remainder $\kappa(\phi; \alpha, \beta)$, defined by (3.160), satisfies the estimation*

$$|\kappa(\phi; \alpha, \beta)| \leq \|\mathfrak{R}'\|_{\infty} \left[\frac{\phi^{(n-1)}(\beta) + \phi^{(n-1)}(\alpha)}{2} - \frac{\phi^{(n-2)}(\beta) - \phi^{(n-2)}(\alpha)}{\beta - \alpha} \right]. \tag{3.162}$$

Proof: The proof is analogous to the proof of Theorem 3.1.6. □

We present the Ostrowski type inequalities related to generalizations of Sherman's inequality.

Theorem 3.5.42: *Let the assumptions of Theorem 3.5.37 hold. Assume (p,q) is a pair of conjugate exponents, that is $1 \leq p, q \leq \infty$, $1/p + 1/q = 1$. Let $\phi^{(n)} \in L_p[\alpha, \beta]$. Then we have:*

$$\left| \sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) - \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) \left[\sum_{p=1}^{l} u_p H_{ij}(x_p) - \sum_{q=1}^{m} v_q H_{ij}(y_q) \right] \right|$$
$$\leq \left\| \phi^{(n)} \right\|_p \|\mathfrak{R}\|_q , \tag{3.163}$$

where $\mathfrak{R}$ is defined in (3.159).

The constant on the right-hand side of (3.163) is sharp for $1 < p \leq \infty$ and the best possible for $p = 1$.

Proof: In the proof we proceed similarly as in the proof of Theorem 3.1.7. □

3.5.3. Mean-value theorems and n-exponential convexity

Now, we present the Lagrange and the Cauchy type of mean-value theorems by using results from the previous subsection. Also, we use exp-convex method in order to interpret our results in the form of exponentially convex functions or in the special case logarithmically convex functions.

Motivated by the inequality (3.152), under the assumptions of Theorems 3.5.38 we define

$$F(\phi) = \sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) \tag{3.164}$$
$$- \sum_{j=1}^{r} \sum_{i=0}^{k_j} \phi^{(i)}(a_i) \left[\sum_{p=1}^{l} u_p H_{ij}(x_p) - \sum_{q=1}^{m} v_q H_{ij}(y_q) \right].$$

Remark 3.5.17.: Under the assumptions of Theorems 3.5.38, the linear functional (3.164) satisfies $F(\phi) \geq 0$ for all n-convex functions ϕ.

Theorem 3.5.43: *Let $\phi \in C^n[\alpha,\beta]$. If the inequality in (3.151) holds, then there exists $\xi \in [\alpha,\beta]$ such that*

$$F(\phi) = \phi^{(n)}(\xi)F(\varphi),$$

where $\varphi(x) = \frac{x^n}{n!}$ and F is defined by (3.164).

Proof: We proceed analogously to the proof of Theorem 3.1.8. □

Theorem 3.5.44: *Let $\phi, \psi \in C^n[\alpha,\beta]$. If the inequality in (3.151) holds, then there exists $\xi \in [\alpha,\beta]$ such that*

$$\frac{F(\phi)}{F(\psi)} = \frac{\phi^{(n)}(\xi)}{\psi^{(n)}(\xi)}, \tag{3.165}$$

provided that the denominators are non-zero and F is defined by (3.164).

Proof: We use the technique of proving Theorem 3.1.9. □

Remark 3.5.18.: If $\frac{\phi^{(n)}}{\psi^{(n)}}$ is an invertible function, then from (3.165) we obtain new mean of Cauchy's type of the interval $[\alpha,\beta]$,

$$\xi = \left(\frac{\phi^{(n)}}{\psi^{(n)}}\right)^{-1}\left(\frac{F(\phi)}{F(\psi)}\right).$$

In order to obtain results regarding to the exponential convexity, we consider the families of exponentially convex functions $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ as given in Subsection 3.1.3. in Definition 3.1.2..

Theorem 3.5.45: *Let F be the linear functional defined as in (3.164) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$). Then the following statements hold:*

(i) The function $t \mapsto F(\phi_t)$ is n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) If the function $t \mapsto F(\phi_t)$ is continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I.

Proof: *The idea of the proof is the same as that of the proof of Theorem 3.1.10.* □

Corollary 3.5.12.: *Let F be the linear functional defined as in (3.164) associated with the family $\mathcal{F}_3$. Then the following statements hold:*

(i) If the function $t \mapsto F(\phi_t)$ is continuous on I, then it is 2-exponentially convex on I. If $t \mapsto F(\phi_t)$ is additionally positive, then it is also log-convex on I. Furthermore, for every choice $r, s, t \in I$, such that $r < s < t$, it holds

$$[F(\phi_s)]^{t-r} \leq [F(\phi_r)]^{t-s} [F(\phi_r)]^{s-r}.$$

(ii) If the function $t \mapsto F(\phi_t)$ is positive and differentiable on I, then for all $r, s, u, v \in I$ such that $r \leq u$, $s \leq v$, we have

$$M_{r,s}(F, \mathcal{F}_3) \leq M_{u,v}(F, \mathcal{F}_3), \tag{3.166}$$

where

$$M_{r,s}(F, \mathcal{F}_3) = \begin{cases} \left(\frac{F(\phi_r)}{F(\phi_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(\frac{\frac{d}{dr}(F(\phi_r))}{F(\phi_r)}\right), & r = s. \end{cases} \tag{3.167}$$

Proof: The idea of the proof is the same as that of the proof of Corollary 3.1.4.. □

Remark 3.5.19.: Remark 3.1.3. is also valid for these functionals. Analogous results can be discussed as given in Section 3.1.3..

3.6. Sherman's inequality and Hermite interpolation polynomials with Green's functions

The main aim of this section is to give generalizations of Sherman's result for n-convex functions by using Hermite interpolating polynomials in combination with Green's functions (3.44)-(3.48). We also consider the results for particular cases, namely, $(m, n-m)$ interpolating polynomial, two-point Taylor interpolating polynomial. We give upper bounds for identity related to generalized Sherman's result and discusse about some applications. The results from this section are quoted from [7].

3.6.1. Generalizations of Sherman's inequality by Hermite interpolation polynomials and Green's functions

Theorem 3.6.46: *Let $\alpha \leq a_1 < a_2 < \cdots < a_r \leq \beta$ $(r \geq 2)$ be distinct points and $\phi \in C^n([\alpha,\beta])$ $(n \geq r)$. Let $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^k$, $\mathbf{u} \in \mathbb{R}^l$ and $\mathbf{v} \in \mathbb{R}^k$ be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \text{and} \quad \mathbf{u} = \mathbf{v}\mathbf{A} \tag{3.168}$$

holds for some matrix $\mathbf{A} \in \mathcal{M}_{kl}(\mathbb{R})$ whose entries satisfy the condition $\sum\limits_{j=1}^{l} a_{ij} = 1$, $i = 1,\dots,k$. Then the following identity holds:

$$\begin{aligned}
&\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) \\
&= \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G(x_p,t) - \sum_{q=1}^{k} v_q G(y_q,t)\right] \sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i+2)}(a_j)dt \\
&+ \int_\alpha^\beta \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G(x_p,t) - \sum_{q=1}^{k} v_q G(y_q,t)\right] G_{H,n-2}(t,s)\phi^{(n)}(s)dsdt,
\end{aligned} \tag{3.169}$$

where G_z, $z = 1,2,3,4,5$, H_{ij} and $G_{H,n}$ are defined as in (3.44)-(3.48), (3.141) and (3.143), respectively.

Proof: Let us consider Green's function G_2, defined by (3.45). Applying (3.50) on the difference $\sum\limits_{p=1}^{l} u_p\phi(x_p) - \sum\limits_{q=1}^{k} v_q\phi(y_q)$ and using (3.168), we have

$$\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) = \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G_2(x_p,t) - \sum_{q=1}^{k} v_q G_2(y_q,t)\right] \phi''(t)dt. \tag{3.170}$$

By Theorem 3.5.36, the function $\phi''(t)$ can be expressed as

$$\phi''(t) = \sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i+2)}(a_j) + \int_\alpha^\beta G_{H,n-2}(t,s)\phi^{(n)}(s)ds. \tag{3.171}$$

Now, combining (3.170) and (3.171), we get (3.169). □

Using the previous identity we get the following generalizations of Sherman's theorem.

Theorem 3.6.47: *Suppose that all the assumptions of Theorem 3.6.46 hold. Additionally, let ϕ be n-convex on $[\alpha, \beta]$ and*

$$\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \geq 0, \quad t \in [\alpha, \beta]. \tag{3.172}$$

(T_1) If k_j is odd for each $j = 2, .., r$, then

$$\begin{aligned} &\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{k} v_q \phi(y_q) \\ &\geq \int_{\alpha}^{\beta} \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \right] \sum_{j=1}^{r} \sum_{i=0}^{k_j} H_{ij}(t) \phi^{(i+2)}(a_j) dt. \end{aligned} \tag{3.173}$$

(T_2) If k_j is odd for each $j = 2, .., r - 1$ and k_r is even, then the reverse inequality in (3.173) holds.

Proof: (T_1) Since the function ϕ is n-convex, then ϕ is n-times differentiable and $\phi^{(n)} \geq 0$. If k_j is odd for each $j = 2, .., r$, then ω, given by (3.142), satisfies $\omega(t) \geq 0$. Therefore, by Lemma 3.5.3. (i) it follows $G_{H,n-2}(t, s) \geq 0$. Hence, we can apply Theorem 3.6.46 to obtain (3.173).

(T_2) This part we can prove similarly. □

Theorem 3.6.48: *Suppose that all the assumptions of Theorem 3.6.46 hold. Additionally, let vectors* **u**, **v** *and matrix* **A** *be nonnegative and ϕ be n-convex on $[\alpha, \beta]$. Then the statements (T_1) and (T_2) hold. Moreover, if (3.173) holds and the function*

$$\bar{F}_z(\cdot) = \sum_{j=1}^{r} \sum_{i=0}^{k_j} \int_{\alpha}^{\beta} G_z(\cdot, t) H_{ij}(t) \phi^{(i+2)}(a_j) dt \tag{3.174}$$

is convex on $[\alpha, \beta]$, then

$$\sum_{p=1}^{l} u_p \phi(x_p) \geq \sum_{q=1}^{k} v_q \phi(y_q). \tag{3.175}$$

Proof: Since the function $G_z(., t)$, $t \in [\alpha, \beta]$, is convex for every $z = 1, ..., 5$, then by

Sherman's theorem it holds that

$$\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \geq 0, \quad t \in [\alpha, \beta].$$

Applying Theorem 3.6.47 and Lemma 3.5.3. (i) we get (3.173). Similarly we can prove the second statement (T_2).

Since (3.173) holds, the right hand side of (3.173) can be rewriting in the form

$$\sum_{p=1}^{l} u_p \bar{F}_z(x_p) - \sum_{q=1}^{k} v_q \bar{F}_z(y_q),$$

where $\bar{F}_z$ is defined by (3.174). If $\bar{F}_z$ is convex, then by Sherman's theorem we have

$$\sum_{p=1}^{l} u_p \bar{F}_z(x_p) - \sum_{q=1}^{k} v_q \bar{F}_z(y_q) \geq 0,$$

i.e. the right hand side of (3.173) is nonnegative, so the inequality (3.175) immediately follows. □

As a direct consequence of the previous results, considering particular case of Hermite interpolating polynomial with type $(m, n-m)$ conditions, we get the following corollary.

Corollary 3.6.13.: *Suppose that all the assumptions of Theorem 3.5.36 hold with type $(m, n-m)$ conditions. Additionally, let ϕ be n-convex on $[\alpha, \beta]$ and $\mathbf{x} \in [\alpha, \beta]^l$, $\mathbf{y} \in [\alpha, \beta]^m$, $\mathbf{u} \in [0, \infty)^l$ and $\mathbf{v} \in [0, \infty)^m$ be such that (3.153) holds for some row stochastic matrix $\mathbf{A} \in \mathcal{M}_{ml}(\mathbb{R})$.*

(i) If $n-m$ is even, then

$$\begin{aligned} &\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{k} v_q \phi(y_q) \qquad (3.176)\\ &\geq \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t)\right] \left(\sum_{i=0}^{m-1} \tau_i(t) \phi^{(i+2)}(\alpha) + \sum_{i=0}^{n-m-1} \eta_i(t) \phi^{(i+2)}(\beta)\right) dt, \end{aligned}$$

where τ_i and η_i are defined as in (3.145) and (3.146), respectively.

(ii) If $n-m$ is odd, then the reverse inequality in (3.176) holds.

(iii) If (3.176) holds and the function

$$\tilde{F}_z(\cdot)=\int_\alpha^\beta G_z(\cdot,t)\left(\sum_{i=0}^{m-1}\tau_i(t)\phi^i(\alpha)+\sum_{i=0}^{n-m-1}\eta_i(t)\phi^i(\beta)\right)dt \tag{3.177}$$

is convex on $[\alpha,\beta]$, then (3.175) holds.

Considering a particular case of Hermite interpolating polynomial with two-point Taylor conditions we get the next generalizations.

Corollary 3.6.14.: *Suppose that all the assumptions of Theorem 3.5.36 hold with two-point Taylor conditions. Additionally, let ϕ be n-convex on $[\alpha,\beta]$ and $\mathbf{x}\in[\alpha,\beta]^l$, $\mathbf{y}\in[\alpha,\beta]^m$, $\mathbf{u}\in[0,\infty)^l$ and $\mathbf{v}\in[0,\infty)^m$ be such that (3.168) holds for some row stochastic matrix $\mathbf{A}\in\mathcal{M}_{ml}(\mathbb{R})$.*

(i) If m is even, then

$$\sum_{p=1}^{l}u_p\phi(x_p)-\sum_{q=1}^{k}v_q\phi(y_q)\geq\int_\alpha^\beta\left[\sum_{p=1}^{l}u_pG_z(x_p,t)-\sum_{q=1}^{k}v_qG_z(y_q,t)\right]F(t)dt, \tag{3.178}$$

where

$$\begin{aligned}F(t)=\sum_{i=0}^{m-1}\sum_{k=0}^{m-1-i}\binom{m+k-1}{k}&\frac{(t-\alpha)^i}{i!}\left(\frac{t-\beta}{\alpha-\beta}\right)^m\left(\frac{t-\alpha}{\beta-\alpha}\right)^k\phi^{(i+2)}(\alpha)\\&+\frac{(t-\beta)^i}{i!}\left(\frac{t-\alpha}{\beta-\alpha}\right)^m\left(\frac{t-\beta}{\alpha-\beta}\right)^k\phi^{(i+2)}(\beta).\end{aligned}$$

(ii) If m is odd, then the reverse inequality in (3.178) holds.

(iii) If (3.178) holds and the function

$$\hat{F}_z(\cdot)=\int_\alpha^\beta G_z(\cdot,t)F(t)dt$$

is convex on $[\alpha,\beta]$, then (3.175) holds.

3.6.2. The Ostrowski and Grüss type inequalities

To avoid many notations, under the assumptions of Theorem 3.6.46, we define the function $\mathcal{R}_z : [\alpha, \beta] \to \mathbb{R}$, $z = 1, ..., 5$, by

$$\mathcal{R}_z(s) = \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{m} v_q G_z(y_q, t) \right] G_{H,n-2}(t, s) dt. \tag{3.179}$$

Then $T(\mathcal{R}_z, \mathcal{R}_z)$ denotes the Čebyšev functional

$$T(\mathcal{R}_z, \mathcal{R}_z) = \frac{1}{\beta - \alpha} \int_\alpha^\beta \mathcal{R}_z^2(t) dt - \left(\frac{1}{\beta - \alpha} \int_\alpha^\beta \mathcal{R}_z(t) dt \right)^2.$$

Theorem 3.6.49: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi^{(n)} \in AC([\alpha, \beta])$ and $(\cdot - \alpha)(\beta - \cdot)(\phi^{(n+1)})^2 \in L[\alpha, \beta]$. Let $\mathcal{R}_z, z = 1, ..., 5$ be defined as in (3.179). Then*

$$\begin{aligned} \kappa_z(\phi; \alpha, \beta) &= \sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{k} v_q \phi(y_q) \\ &- \int_\alpha^\beta \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \right] \sum_{j=1}^{r} \sum_{i=0}^{k_j} H_{ij}(t) \phi^{(i+2)}(a_j) dt \\ &- \frac{\phi^{(n-1)}(\beta) - \phi^{(n-1)}(\alpha)}{\beta - \alpha} \int_\alpha^\beta \mathcal{R}_z(s) ds \end{aligned} \tag{3.180}$$

and $\kappa_z(\phi; \alpha, \beta)$ satisfies the estimation

$$|\kappa_z(\phi; \alpha, \beta)| \leq \frac{\sqrt{\beta - \alpha}}{\sqrt{2}} [T(\mathcal{R}_z, \mathcal{R}_z)]^{\frac{1}{2}} \left| \int_\alpha^\beta (s - \alpha)(\beta - s)[\phi^{(n+1)}(s)]^2 ds \right|^{\frac{1}{2}} \tag{3.181}$$

Proof: The idea of the proof is the same as the proof of Theorem 3.1.5. □

Theorem 3.6.50: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi^{(n)}$ is monotonic nondecreasing on $[\alpha, \beta]$ and $\mathcal{R}_z, z = 1, ..., 5$ be defined as in (3.179). Then (3.180) holds and $\kappa_z(\phi; \alpha, \beta)$ satisfies the estimation*

$$|\kappa_z(\phi; \alpha, \beta)| \leq \|\mathcal{R}_z'\|_\infty \left\{ \frac{\phi^{(n-1)}(\beta) + \phi^{(n-1)}(\alpha)}{2} - \frac{\phi^{(n-2)}(\beta) - \phi^{(n-2)}(\alpha)}{\beta - \alpha} \right\}. \tag{3.182}$$

Proof: The idea of the proof is the same as in the proof of Theorem 3.1.6. □

Theorem 3.6.47 gives the lower bound for the expression

$$\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) \\ -\int_{\alpha}^{\beta}\left[\sum_{p=1}^{l} u_p G_z(x_p,t) - \sum_{q=1}^{k} v_q G_z(y_q,t)\right]\sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i+2)}(a_j)dt. \tag{3.183}$$

Next result gives an upper bound for (3.183).

Theorem 3.6.51: *Let (p, q) be a pair of conjugate exponents, i.e. $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\left|\phi^{(n)}\right|^p \in L[\alpha, \beta]$ and $\mathcal{R}_z, z = 1, ..., 5$ be defined as in (3.179). Then*

$$\left|\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) \right. \\ \left. -\int_{\alpha}^{\beta}\left[\sum_{p=1}^{l} u_p G_z(x_p,t) - \sum_{q=1}^{k} v_q G_z(y_q,t)\right]\sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i+2)}(a_j)dt\right| \\ \leq \left\|\phi^{(n)}\right\|_p \left\|\mathcal{R}_z\right\|_q. \tag{3.184}$$

The constant $\|\mathcal{R}_z\|_q$ is sharp for $1 < p \leq \infty$ and the best possible for $p = 1$.

Proof: The idea of the proof is the same as that of the proof of Theorem 3.1.7. □

3.6.3. Mean-value theorems and n-exponential convexity

Motivated by the inequality (3.173), under the assumptions of Theorems 3.6.47, we define the linear functional $\Lambda : C^n([\alpha, \beta]) \to \mathbb{R}$ by

$$\Lambda(\phi) = \sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) \\ -\int_{\alpha}^{\beta}\left[\sum_{p=1}^{l} u_p G(x_p,t) - \sum_{q=1}^{k} v_q G(y_q,t)\right]\sum_{j=1}^{r}\sum_{i=0}^{k_j} H_{ij}(t)\phi^{(i+2)}(a_j)dt. \tag{3.185}$$

Remark 3.6.20.: By Theorem 3.6.47, for every n-convex functions $\phi : [\alpha, \beta] \to \mathbb{R}$ we have

$$\Lambda(\phi) \geq 0.$$

Using the linearity of this functional we derive mean value theorems of Lagrange and Cauchy type.

Theorem 3.6.52: *Let $\phi \in C^n([\alpha, \beta])$ and $\Lambda : C^n([\alpha, \beta]) \to \mathbb{R}$ be the linear functional defined by (3.185). Then there exists $\xi \in [\alpha, \beta]$ such that*

$$\Lambda(\phi) = \phi^{(n)}(\xi)\Lambda(\varphi),$$

where $\varphi(x) = \frac{x^n}{n!}$.

Proof: Similar to the proof of Theorem 3.1.8. □

Theorem 3.6.53: *Let $\phi, \psi \in C^n([\alpha, \beta])$ and $\Lambda : C^n([\alpha, \beta]) \to \mathbb{R}$ be the linear functional defined by (3.185). Then there exists $\xi \in [\alpha, \beta]$ such that*

$$\frac{\Lambda(\phi)}{\Lambda(\psi)} = \frac{\phi^{(n)}(\xi)}{\psi^{(n)}(\xi)}, \tag{3.186}$$

provided that the denominators are non-zero.

Proof: Similar to the proof of Theorem 3.1.9. □

Remark 3.6.21.: With assumption that $\frac{\phi^{(n)}}{\psi^{(n)}}$ is an invertible function, as a consequence of the previous theorem we get the single number

$$\xi = \left(\frac{\phi^{(n)}}{\psi^{(n)}}\right)^{-1} \left(\frac{\Lambda(\phi)}{\Lambda(\psi)}\right)$$

which is exactly mean of Chauchy type of the segment $[\alpha, \beta]$.

Through the rest of this section, I denotes an open interval in $\mathbb{R}$.

Let us consider the families of exponentially convex functions $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ as given in Subsection 3.1.3. in Definition 3.1.2..

Theorem 3.6.54: *Let Λ be the linear functional defined as in (3.185) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$). Then the following statements hold:*

(i) The function $t \mapsto \Lambda(\phi_t)$ is n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) If the function $t \mapsto \Lambda(\phi_t)$ is continuous on I, then it is n-exponentially convex (resp. exponentially convex) on I.

Proof: *The idea of the proof is the same as that of the proof of Theorem 3.1.10.* □

Corollary 3.6.15.: *Let Λ be the linear functional defined as in (3.185) associated with the family $\mathcal{F}_3$. Then the following statements hold:*

(i) If the function $t \mapsto \Lambda(\phi_t)$ is continuous on I, then it is 2-exponentially convex on I. If $t \mapsto \Lambda(\phi_t)$ is additionally positive, then it is also log-convex on I. Furthermore, for every choice $r, s, t \in I$, such that $r < s < t$, it holds

$$[\Lambda(\phi_s)]^{t-r} \leq [\Lambda(\phi_r)]^{t-s} [\Lambda(\phi_r)]^{s-r}.$$

(ii) If the function $t \mapsto \Lambda(\phi_t)$ is positive and differentiable on I, then for all $r, s, u, v \in I$ such that $r \leq u$, $s \leq v$, we have

$$M_{r,s}(\Lambda, \mathcal{F}_3) \leq M_{u,v}(\Lambda, \mathcal{F}_3), \tag{3.187}$$

where

$$M_{r,s}(\Lambda, \mathcal{F}_3) = \begin{cases} \left(\frac{\Lambda(\phi_r)}{\Lambda(\phi_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(\frac{\frac{d}{dr}(\Lambda(\phi_r))}{\Lambda(\phi_r)}\right), & r = s. \end{cases} \tag{3.188}$$

Proof: The idea of the proof is the same as that of the proof of Corollary 3.1.4.. □

Example 11.: *As an example of application of the previous results consider the family of functions*

$$\Omega_4 = \{\varphi_t : (0, \infty) \to (0, \infty) : t \in (0, \infty)\}$$

defined by

$$\varphi_t(x) = \frac{e^{-x\sqrt{t}}}{\left(-\sqrt{t}\right)^n}.$$

Since $\frac{d^n\varphi_t}{dx^n}(x) = e^{-x\sqrt{t}} > 0$, *the function* φ_t *is* n*-convex function for every* $t > 0$. *Moreover, the function* $t \mapsto \frac{d^n\varphi_t}{dx^n}(x)$ *is exponentially convex. Therefore, the function* $t \mapsto [x_0, x_1, ..., x_l; \varphi_t]$ *is exponentially convex (and so exponentially convex in the Jensen sense). Also, the function* $t \mapsto \Lambda(\varphi_t)$ *is exponentially convex in the Jensen sense. It is easy to verify that the function* $t \mapsto \Lambda(\varphi_t)$ *is continuous, so it is exponentially convex.*

For this family of functions, with assumption that $[\alpha, \beta] \subset (0, \infty)$ *and* $t \mapsto \Lambda(\varphi_t)$ *is positive, (3.188) becomes*

$$\mu_{\eta,\zeta} = \left(\frac{\zeta^n}{\eta^n} \cdot \frac{\sum\limits_{p=1}^{l} u_p e^{-x_p\sqrt{\eta}} - \sum\limits_{q=1}^{k} v_q e^{-y_q\sqrt{\eta}} - A_1}{\sum\limits_{p=1}^{l} u_p e^{-x_p\sqrt{\zeta}} - \sum\limits_{q=1}^{k} v_q e^{-y_q\sqrt{\zeta}} - B_1}\right)^{\frac{1}{\eta-\zeta}}, \quad \eta \neq \zeta,$$

$$\mu_{\eta,\eta} = \exp\left(\frac{\sum\limits_{q=1}^{k} v_q y_q e^{-y_q\sqrt{\eta}} - \sum\limits_{p=1}^{l} u_p x_p e^{-x_p\sqrt{\eta}} + A_2}{2\sqrt{\eta}\left(\sum\limits_{p=1}^{l} u_p e^{-x_p\sqrt{\eta}} - \sum\limits_{q=1}^{k} v_q e^{-y_q\sqrt{\eta}} - A_1\right)} - \frac{n}{\eta}\right),$$

where

$$A_1 = \int_\alpha^\beta (\sum_{p=1}^{l} u_p G(x_p, t) - \sum_{q=1}^{k} v_q G(y_q, t)) \sum_{j=1}^{r} \sum_{i=0}^{k_j} H_{ij}(t)(-1)^{i+2}\eta^{1+\frac{i}{2}} e^{-a_j\sqrt{\eta}} dt,$$

$$A_2 = \int_\alpha^\beta (\sum_{p=1}^{l} u_p G(x_p, t) - \sum_{q=1}^{k} v_q G(y_q, t)) \sum_{j=1}^{r} \sum_{i=0}^{k_j} H_{ij}(t) \frac{d^{i+2}}{dx^{i+2}}(xe^{-x\sqrt{\eta}})|_{x=a_j} dt,$$

$$B_1 = \int_\alpha^\beta (\sum_{p=1}^{l} u_p G(x_p, t) - \sum_{q=1}^{k} v_q G(y_q, t))) \sum_{j=1}^{r} \sum_{i=0}^{k_j} H_{ij}(t)(-1)^{i+2}\zeta^{1+\frac{i}{2}} e^{-a_j\sqrt{\zeta}} dt.$$

Using Theorem 3.6.53 it follows that

$$M_{\eta,\zeta}(\Lambda, \Omega_4) = -\left(\sqrt{\eta} + \sqrt{\zeta}\right) \log \mu_{\eta,\zeta}(\Lambda, \Omega_4)$$

satisfies

$$\alpha \leq M_{\eta,\zeta}(\Lambda, \Omega_4) \leq \beta.$$

It is obviously symmetric. By Corollary 3.6.15., using (3.187), it follows that this mean is monotonic.

Remark 3.6.22.: Analogous results can be discussed as given in Section 3.1.3..

3.7. Sherman's inequality and Montgomery's identity with Green's functions

In this section, we obtain generalized Sherman's type inequalities for n-convex functions by using Montgomery identity with Green's functions. We also give bounds for identities related to obtained generalizations in the form of the Grüss and Ostrowski type inequalities. Applying exp-convex method we establish new examples of exponentially convex functions and related generalized Cauchy type means. The results from this section are quoted from [8, 9].

In our results we use the following generalized Montgomery identities via Taylor's formula as given in [12], [13].

Proposition 3.7.55: Let $n \in \mathbb{N}$, $\phi : I \to \mathbb{R}$ be such that $\phi^{(n-1)} \in AC(I)$, where I is an open interval in $\mathbb{R}$, $\alpha, b \in I$, $\alpha < b$. Then the following identity holds

$$\phi(x) = \frac{1}{b-\alpha}\int_\alpha^b \phi(t)\,dt + \sum_{k=0}^{n-2}\frac{\phi^{(k+1)}(\alpha)}{k!\,(k+2)}\frac{(x-\alpha)^{k+2}}{b-\alpha} - \sum_{k=0}^{n-2}\frac{\phi^{(k+1)}(b)}{k!\,(k+2)}\frac{(x-b)^{k+2}}{b-\alpha} + \frac{1}{(n-1)!}\int_\alpha^b T_n(x,s)\,\phi^{(n)}(s)\,ds, \tag{3.189}$$

where

$$T_n(x,s) = \begin{cases} -\dfrac{(x-s)^n}{n(b-\alpha)} + \dfrac{x-\alpha}{b-\alpha}(x-s)^{n-1}, & \alpha \le s \le x, \\ -\dfrac{(x-s)^n}{n(b-\alpha)} + \dfrac{x-b}{b-\alpha}(x-s)^{n-1}, & x < s \le b. \end{cases} \tag{3.190}$$

Remark 3.7.23.: For $n = 1$, the sum on the right side of (3.189) is equal to zero, so the identity (3.189) reduces to the well known **Montgomery identity**

$$\phi(x) = \frac{1}{b-\alpha}\int_\alpha^b \phi(t)\,dt + \int_\alpha^b P(x,s)\,\phi'(s)\,ds,$$

where $P(x,s)$ is the Peano kernel, defined by

$$P(x,s)=\begin{cases}\dfrac{s-\alpha}{b-\alpha}, & \alpha\leq s\leq x\\ \dfrac{s-b}{b-\alpha}, & x<s\leq b\end{cases},$$

(see for instance [112]).

3.7.1. Generalizations of Sherman's inequality by Montgomery's identity

First, applying Montgomery identity we establish new identity for Sherman's difference.

Theorem 3.7.56: *Let $n\in\mathbb{N}$, $\phi: I\to\mathbb{R}$ be such that $\phi^{(n-1)}\in AC(I)$, where I is an open interval in $\mathbb{R}$, $\alpha,\beta\in I$, $\alpha<\beta$. Suppose that $\mathbf{x}=(x_1,...,x_l)\in[\alpha,\beta]^l$, $\mathbf{y}=(y_1,...,y_m)\in[\alpha,\beta]^m$, $\mathbf{u}=(u_1,...,u_l)\in\mathbb{R}^l$, $\mathbf{v}=(v_1,...,v_m)\in\mathbb{R}^m$ be such that*

$$\mathbf{y}=\mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{u}=\mathbf{v}\mathbf{A} \tag{3.191}$$

holds for some matrix A$=(a_{ij})\in\mathcal{M}_{lm}(\mathbb{R})$ satisfying the condition $\sum\limits_{j=1}^{m}a_{ij}=1$, $i=1,2,...,l$. Then

$$\begin{aligned}&\sum_{p=1}^{l}u_p\phi(x_p)-\sum_{q=1}^{m}v_q\phi(y_q) \\ =&\ \frac{1}{\beta-\alpha}\sum_{k=0}^{n-2}\frac{1}{k!(k+2)}\left[\phi^{(k+1)}(\alpha)\left(\sum_{p=1}^{l}u_p(x_p-\alpha)^{k+2}-\sum_{q=1}^{m}v_q(y_q-\alpha)^{k+2}\right)\right. \\ -&\ \left.\phi^{(k+1)}(\beta)\left(\sum_{p=1}^{l}u_p(x_p-\beta)^{k+2}-\sum_{q=1}^{m}v_q(y_q-\beta)^{k+2}\right)\right] \\ +&\ \frac{1}{(n-1)!}\int_{\alpha}^{\beta}\left(\sum_{p=1}^{l}u_pT_n(x_p,s)-\sum_{q=1}^{m}v_qT_n(y_q,s)\right)\phi^{(n)}(s)ds.\end{aligned} \tag{3.192}$$

Proof: Applying (3.189) to the differenece $\sum\limits_{p=1}^{l}u_p\phi(x_p)-\sum\limits_{q=1}^{m}v_q\phi(y_q)$ we obtain (3.192). □

Now we state generalized Sherman's type inequality by using the above identity.

Theorem 3.7.57: *Let all the assumptions of Theorem 3.7.56 be satisfied. For $T_n(\cdot,s)$, $s\in$*

$[\alpha, \beta]$, *defined by (3.190), let*

$$\sum_{q=1}^{m} v_q T_n(y_p, s) \leq \sum_{p=1}^{l} u_p T_n(x_p, s). \tag{3.193}$$

If ϕ is n-convex on I, then the following inequality holds:

$$\begin{aligned}
&\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) \qquad (3.194)\\
\geq\ & \frac{1}{\beta - \alpha} \sum_{k=0}^{n-2} \frac{1}{k!(k+2)} \left[\phi^{(k+1)}(\alpha) \left(\sum_{p=1}^{l} u_p (x_p - \alpha)^{k+2} - \sum_{q=1}^{m} v_q (y_q - \alpha)^{k+2} \right) \right.\\
-\ & \left. \phi^{(k+1)}(\beta) \left(\sum_{p=1}^{l} u_p (x_p - \beta)^{k+2} - \sum_{q=1}^{m} v_q (y_q - \beta)^{k+2} \right) \right].
\end{aligned}$$

Proof: Under the assumptions of theorem the identity (3.192) holds. Since ϕ is n-convex, then $\phi^{(n)} \geq 0$ on I. Using this fact in combination with (3.193) we get the required result. □

Theorem 3.7.58: *Let all the assumptions of Theorem 3.7.56 be satisfied. Additionally, let* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{A}$ *are with nonnegative entries. If ϕ is $2n$-convex function, then*

$$\begin{aligned}
&\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{m} v_q \phi(y_q) \qquad (3.195)\\
\geq\ & \frac{1}{\beta - \alpha} \sum_{k=0}^{2n-2} \frac{1}{k!(k+2)} \left[\phi^{(k+1)}(\alpha) \left(\sum_{p=1}^{l} u_p (x_p - \alpha)^{k+2} - \sum_{q=1}^{m} v_q (y_q - \alpha)^{k+2} \right) \right.\\
-\ & \left. \phi^{(k+1)}(\beta) \left(\sum_{p=1}^{l} u_p (x_p - \beta)^{k+2} - \sum_{q=1}^{m} v_q (y_q - \beta)^{k+2} \right) \right].
\end{aligned}$$

Moreover, if $\phi^{(k)}(\alpha) \geq 0$ and $(-1)^k \phi^{(k)}(\beta) \geq 0$ for $k = 1, ..., 2n-1$, then

$$\sum_{p=1}^{l} u_p \phi(x_p) \geq \sum_{q=1}^{m} v_q \phi(y_q).$$

Proof: For $T_n(\cdot, s), s \in [\alpha, \beta]$, defined by (3.190), we have

$$\frac{d^2T_n(x,s)}{dx^2} = \begin{cases} \frac{n-1}{\beta-\alpha}[(x-s)^{n-2} + (n-2)(x-\alpha)(x-s)^{n-3}], & \alpha \le s \le x \le \beta \\ \frac{n-1}{\beta-\alpha}[(x-s)^{n-2} + (n-2)(x-\beta)(x-s)^{n-3}], & \alpha \le x \le s \le \beta \end{cases},$$

so $T_n(\cdot, s), s \in [\alpha, \beta]$ is continuous for every $n \ge 2$ and convex for even n. Hence, by Sherman's Theorem the inequality (3.193) holds for even n and by using Theorem 3.7.60 the inequality (3.195) hold if n is replaced by $2n$.

To prove the second part, let us consider the function $f(x) = (x-\alpha)^{k+2}$, which is continuous convex function on $[\alpha, \beta]$. Again, applying Sherman's Theorem to f, we get

$$\sum_{p=1}^{l} u_p(x_p-\alpha)^{k+2} \ge \sum_{q=1}^{m} v_q(y_q-\alpha)^{k+2}. \tag{3.196}$$

Similarly, the function $g(x) = (x-\beta)^{k+2}$ is convex on $[\alpha, \beta]$ if k is even and concave on $[\alpha, \beta]$ if k is odd. So we have

$$\sum_{p=1}^{l} u_p(x_p-\beta)^{k+2} \ \ge \ \sum_{q=1}^{m} v_q(y_q-\beta)^{k+2}, \quad \text{if } k \text{ is even }, \tag{3.197}$$

$$\sum_{p=1}^{l} u_p(x_p-\beta)^{k+2} \ \le \ \sum_{q=1}^{m} v_q(y_q-\beta)^{k+2}, \quad \text{if } k \text{ is odd }. \tag{3.198}$$

Now using (3.196), (3.197), (3.198) and the the assumptions $\phi^{(k)}(\alpha) \ge 0$ and $(-1)^k\phi^{(k)}(\beta) \ge 0$, $k = 1, 2, ..., 2n-1$, we have

$$\frac{1}{\beta-\alpha}\sum_{k=0}^{2n-2}\left[\left\{\sum_{p=1}^{l} u_p(x_p-\alpha)^{k+2} - \sum_{q=1}^{m} v_q(y_q-\alpha)^{k+2}\right\}\frac{\phi^{(k+1)}(\alpha)}{k!(k+2)}\right]$$
$$-\frac{1}{\beta-\alpha}\sum_{k=0}^{2n-2}\left[\left\{\sum_{p=1}^{l} u_p(x_p-\beta)^{k+2} - \sum_{q=1}^{m} v_q(y_q-\beta)^{k+2}\right\}\frac{\phi^{(k+1)}(\beta)}{k!(k+2)}\right] \ge 0.$$

Hence the result follows. □

3.7.2. Generalized Sherman's inequality by Montgomery's identity and Green's functions

In this subsection, using Montgomery identity with Green's functions we obtain new identity.

Theorem 3.7.59: *Let $n \in \mathbb{N}$, $n \geq 4$, $\phi : I \to \mathbb{R}$ be such that $\phi^{(n-1)} \in AC(I)$, where $I \subset \mathbb{R}$ is an open interval, $\alpha, \beta \in I$, $\alpha < \beta$. Suppose that $\mathbf{x} = (x_1, ..., x_l) \in [\alpha, \beta]^l$, $\mathbf{y} = (y_1, ..., y_m) \in [\alpha, \beta]^m$, $\mathbf{u} = (u_1, ..., u_l) \in \mathbb{R}^l$, $\mathbf{v} = (v_1, ..., v_m) \in \mathbb{R}^m$ be such that (3.191) holds for some matrix $\mathbf{A} = (a_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$ satisfying the condition $\sum\limits_{j=1}^{m} a_{ij} = 1$, $i = 1, 2, ..., l$. Let T_n and G_z be defined as in (3.190) and (3.44)-(3.48), respectively. Then*

$$
\begin{aligned}
&\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{k} v_q \phi(y_q) \qquad (3.199)\\
&= \int_{\alpha}^{\beta} \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \right] \\
&\cdot \left[\frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} \phi(t)\, dt + \sum_{k=0}^{n-4} \frac{\phi^{(k+1)}(\alpha)}{k!\,(k+2)} \frac{(t-\alpha)^{k+2}}{\beta - \alpha} - \sum_{k=0}^{n-4} \frac{\phi^{(k+1)}(\beta)}{k!\,(k+2)} \frac{(t-\beta)^{k+2}}{\beta - \alpha} \right] dt \\
&+ \frac{1}{(n-1)!} \int_{\alpha}^{\beta} \int_{\alpha}^{\beta} \left[\sum_{p=1}^{l} u_p G_z(x_p, t) - \sum_{q=1}^{k} v_q G_z(y_q, t) \right] T_{n-2}(t, s)\, \phi^{(n)}(s)\, ds dt.
\end{aligned}
$$

Proof: Let us consider Green's function G_2 defined by (3.45). Applying (3.50) on the difference $\sum\limits_{p=1}^{l} u_p \phi(x_p) - \sum\limits_{q=1}^{k} v_q \phi(y_q)$ and applying (3.191), we have

$$
\sum_{p=1}^{l} u_p \phi(x_p) - \sum_{q=1}^{k} v_q \phi(y_q) = \int_{\alpha}^{\beta} \left[\sum_{p=1}^{l} u_p G_2(x_p, t) - \sum_{q=1}^{k} v_q G_2(y_q, t) \right] \phi''(t) dt. \qquad (3.200)
$$

By Proposition 3.7.55, the function $\phi''(t)$ can be expressed as

$$\phi''(t) = \frac{1}{\beta-\alpha}\int_{\alpha}^{\beta}\phi(t)\,dt + \sum_{k=0}^{n-4}\frac{\phi^{(k+1)}(\alpha)}{k!\,(k+2)}\frac{(t-\alpha)^{k+2}}{\beta-\alpha} - \sum_{k=0}^{n-4}\frac{\phi^{(k+1)}(\beta)}{k!\,(k+2)}\frac{(t-\beta)^{k+2}}{\beta-\alpha} \tag{3.201}$$

$$+\frac{1}{(n-1)!}\int_{\alpha}^{\beta}T_{n-2}(t,s)\,\phi^{(n)}(s)\,ds.$$

Now, combining (3.200) and (3.201), we get (3.199) for Green's function G_2. Analogously we can prove identities for other Green's functions. □

Now we state generalization of Sherman's results by using the previous identity.

Theorem 3.7.60: *Suppose that all the assumptions of Theorem 3.7.59 hold. Additionally, if for any even n the function $\phi : I \to \mathbb{R}$ is n-convex and*

$$\sum_{p=1}^{l} u_p G_z(x_p,t) - \sum_{q=1}^{k} v_q G_z(y_q,t) \geq 0, \quad t \in [\alpha,\beta], \tag{3.202}$$

then the following inequality holds:

$$\begin{aligned} &\sum_{p=1}^{l} u_p\phi(x_p) - \sum_{q=1}^{k} v_q\phi(y_q) \qquad (3.203)\\ &\geq \int_{\alpha}^{\beta}\left[\sum_{p=1}^{l} u_p G_z(x_p,t) - \sum_{q=1}^{k} v_q G_z(y_q,t)\right]\\ &\cdot\left[\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta}\phi(t)\,dt + \sum_{k=0}^{n-4}\frac{\phi^{(k+1)}(\alpha)}{k!\,(k+2)}\frac{(t-\alpha)^{k+2}}{\beta-\alpha} - \sum_{k=0}^{n-4}\frac{\phi^{(k+1)}(\beta)}{k!\,(k+2)}\frac{(t-\beta)^{k+2}}{\beta-\alpha}\right]dt. \end{aligned}$$

Proof: Since ϕ is n-convex, then $\phi^{(n)} \geq 0$. Also it obvious that if n is even then $T_n \geq 0$ because

Case I: If $\alpha \leq s \leq x$, then $x - a \geq x - s$ and hence $(x-\alpha)(x-s)^{n-1} \geq (x-s)^n$ that is

$$\frac{(x-\alpha)(x-s)^{n-1}}{\beta-\alpha} \geq \frac{(x-s)^n}{\beta-\alpha} \geq \frac{(x-s)^n}{n(\beta-\alpha)}.$$

So in this case from (3.190) we have $T_n \geq 0$.

Case II: If $x < s \leq \beta$, then $x - s < 0$ and $s - \beta \leq 0$. As n is even so we have $(x-s)^{n-1} < 0$, therefore we have

$$\begin{aligned}(x-s)^{n-1}(s-\beta) &\geq 0\\ &\Rightarrow (x-s)^{n-1}(x-\beta+s-x)\geq 0\\ &\Rightarrow (x-s)^{n-1}(x-\beta)-(x-s)^{n}\geq 0\end{aligned}$$

that is

$$\frac{(x-s)^{n-1}(x-\beta)}{\beta-\alpha}\geq\frac{(x-s)^{n}}{\beta-\alpha}\geq\frac{(x-s)^{n}}{n(\beta-\alpha)}.$$

So in this case from (3.190) we have $T_n\geq 0$.

Now using (3.7.60) and the positivity of T_n and $\phi^{(n)}$ in (3.199) we get (3.203). □

The next theorem gives generalization of Sherman's theorem under Sherman's condition of nonnegative entries weights $\mathbf{u}$, $\mathbf{v}$ and matrix $\mathbf{A}$.

Theorem 3.7.61: *Suppose that all the assumptions of Theorem 3.7.59 hold. Additionally, let $\mathbf{u},\mathbf{v}$ and $\mathbf{A}$ be with nonnegative entries.*

(i) If n is even and ϕ is n-convex function, then (3.203) holds.

(ii) If (3.203) holds and the functions $L_z:[\alpha,\beta]\to\mathbb{R}$, $z=1,...,5$, defined by

$$L_z(\cdot)=G_z(\cdot,t)\left(\tfrac{1}{\beta-\alpha}\int_{\alpha}^{\beta}\phi\left(t\right)dt+\sum_{k=0}^{n-4}\tfrac{\phi^{(k+1)}(\alpha)}{k!(k+2)}\tfrac{(t-\alpha)^{k+2}}{\beta-\alpha}-\sum_{k=0}^{n-4}\tfrac{\phi^{(k+1)}(\beta)}{k!(k+2)}\tfrac{(t-\beta)^{k+2}}{\beta-\alpha}\right),\tag{3.204}$$

where $G_z, z=1,...,5$ is defined as in (3.44)-(3.48), are convex on $[\alpha,\beta]$, then

$$\sum_{p=1}^{l}u_p\phi(x_p)-\sum_{q=1}^{k}v_q\phi(y_q)\geq 0.\tag{3.205}$$

Proof: Since the function $G_z(\cdot,t)$, $t\in[\alpha,\beta]$, is convex for every $z=1,...,5$, by Sherman's theorem we have

$$\sum_{p=1}^{l}u_pG_z(x_p,t)-\sum_{q=1}^{k}v_qG_z(y_q,t)\geq 0.$$

Applying Theorem 3.7.60 we get (3.203).

Since (3.203) holds, the right hand side of (3.203) can be rewritten in the form

$$\sum_{p=1}^{l} u_p L_z(x_p) - \sum_{q=1}^{k} v_q L_z(y_q),$$

where L_z is defined by (3.204). If L_z is convex, then by Sherman's theorem we have

$$\sum_{p=1}^{l} u_p L_z(x_p) - \sum_{q=1}^{k} v_q L_z(y_q) \geq 0,$$

i.e. the right hand side of (3.203) is nonnegative, so the inequality (3.205) immediately follows. □

3.7.3. The Ostrowski and Grüss type inequalities

To avoid many notations, under the assumptions of Theorem 3.7.60, we define the function $\mathcal{P}_z : [\alpha, \beta] \to \mathbb{R}$, $z = 1, ..., 5$, by

$$\mathcal{P}_z(s) = \int_{\alpha}^{\beta} \left[\sum_{p=1}^{l} u_p G(x_p, t) - \sum_{q=1}^{k} v_q G(y_q, t) \right] T_{n-2}(t, s)\, dt. \tag{3.206}$$

Then $T(\mathcal{P}_z, \mathcal{P}_z)$ denotes the Čebyšev functional

$$T(\mathcal{P}_z, \mathcal{P}_z) = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} \mathcal{P}_z^2(s) ds - \left(\frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} \mathcal{P}_z(s) ds \right)^2, \quad z = 1, ..., 5.$$

Using previous two theorems we obtain upper bounds for the identity (3.199) related to generalizations of Sherman's inequality.

Theorem 3.7.62: *Let $\phi : [\alpha, \beta] \to \mathbb{R}$ be such that $\phi^{(n)} \in AC([\alpha, \beta])$ with $(\cdot - \alpha)(\beta -$*

$\cdot)(\phi^{(n+1)})^2 \in L[\alpha,\beta]$. *Let* $\mathcal{P}_z, z=1,...,5,$ *be defined as in (3.206). Then*

$$\kappa_z(\phi;\alpha,\beta)=\sum_{p=1}^{l}u_p\phi(x_p)-\sum_{q=1}^{k}v_q\phi(y_q)-\int_\alpha^\beta\left[\sum_{p=1}^{l}u_pG_z(x_p,t)-\sum_{q=1}^{k}v_qG_z(y_q,t)\right]$$
$$\cdot\left[\frac{1}{\beta-\alpha}\int_\alpha^\beta\phi\left(t\right)dt+\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\alpha\right)}{k!\left(k+2\right)}\frac{\left(t-\alpha\right)^{k+2}}{\beta-\alpha}-\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\beta\right)}{k!\left(k+2\right)}\frac{\left(t-\beta\right)^{k+2}}{\beta-\alpha}\right]dt$$
$$-\frac{\phi^{(n-1)}(\beta)-\phi^{(n-1)}(\alpha)}{\beta-\alpha}\int_\alpha^\beta\mathcal{P}_z(s)ds \tag{3.207}$$

satisfies the estimation

$$|\kappa_z(\phi;\alpha,\beta)|\leq\frac{\sqrt{\beta-\alpha}}{\sqrt{2}(n-1)!}\left[T(\mathcal{P}_z,\mathcal{P}_z)\right]^{\frac{1}{2}}\left|\int_\alpha^\beta(s-\alpha)(\beta-s)[\phi^{(n+1)}(s)]^2ds\right|^{\frac{1}{2}}. \tag{3.208}$$

Proof: The idea of the proof is the same as the proof of Theorem 3.1.5. □

Theorem 3.7.63: *Let* $\phi:[\alpha,\beta]\to\mathbb{R}$ *be such that* $\phi^{(n)}$ *is monotonic nondecreasing on* $[\alpha,\beta]$ *and* $\mathcal{P}_z, z=1,...,5,$ *be defined as in (3.206). Then* $\kappa_z(\phi;\alpha,\beta)$, *defined by (3.207), satisfies the estimation*

$$|\kappa_z(\phi;\alpha,\beta)|\leq\frac{\|\mathcal{P}_z'\|_\infty}{(n-1)!}\left\{\frac{\phi^{(n-1)}(\beta)+\phi^{(n-1)}(\alpha)}{2}-\frac{\phi^{(n-2)}(\beta)-\phi^{(n-2)}(\alpha)}{\beta-\alpha}\right\}. \tag{3.209}$$

Proof: The steps in the proof are similar to the proof of Theorem 3.1.6. □

In the following theorem we present Ostrowski type inequality related to the previous obtained Sherman's type inequalities.

Theorem 3.7.64: *Let* (p,q) *be a pair of conjugate exponents, i.e.* $1\leq p,q\leq\infty$ *and* $1/p+1/q=1$. *Let* $\phi:[\alpha,\beta]\to\mathbb{R}$ *be such that* $\left|\phi^{(n)}\right|^p\in L[\alpha,\beta]$ *and* $\mathcal{P}_z, z=1,...,5,$ *be* *defined as in (3.206). Then*

$$
\begin{aligned}
&-\int_{\alpha}^{\beta}\left[\sum_{p=1}^{l}u_pG_z(x_p,t)-\sum_{q=1}^{k}v_qG_z(y_q,t)\right]\\
&\cdot\left[\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta}\phi\left(t\right)dt+\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\alpha\right)}{k!\left(k+2\right)}\frac{\left(t-\alpha\right)^{k+2}}{\beta-\alpha}-\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\beta\right)}{k!\left(k+2\right)}\frac{\left(t-\beta\right)^{k+2}}{\beta-\alpha}\right]dt\\
&\leq\frac{1}{(n-1)!}\left\|\phi^{(n)}\right\|_p\left\|\mathcal{P}_z\right\|_q. \qquad (3.210)
\end{aligned}
$$

The constant $\left\|\mathcal{P}_z\right\|_q$ is sharp for $1<p\leq\infty$ and the best possible for $p=1$.

Proof: See the proof of Theorem 3.1.7. □

3.7.4. Mean value theorems and exponential convexity

Motivated by the inequality (3.203), under the assumptions of Theorems 3.7.60, we define the linear functionals $\Lambda_z : C^n([\alpha,\beta])\to\mathbb{R}$, $z=1,...,5$, by

$$
\begin{aligned}
\Lambda_z(\phi)=&-\int_{\alpha}^{\beta}\left[\sum_{p=1}^{l}u_pG_z(x_p,t)-\sum_{q=1}^{k}v_qG_z(y_q,t)\right] \qquad (3.211)\\
&\cdot\left[\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta}\phi\left(t\right)dt+\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\alpha\right)}{k!\left(k+2\right)}\frac{\left(t-\alpha\right)^{k+2}}{\beta-\alpha}-\sum_{k=0}^{n-4}\frac{\phi^{(k+1)}\left(\beta\right)}{k!\left(k+2\right)}\frac{\left(t-\beta\right)^{k+2}}{\beta-\alpha}\right]dt.
\end{aligned}
$$

Remark 3.7.24.: By Theorem 3.203, for every n-convex functions $\phi:[\alpha,\beta]\to\mathbb{R}$ we have

$$\Lambda_z(\phi)\geq 0,\quad z=1,...,5.$$

Using the linearity of this functional we derive mean-value theorems of Lagrange and Cauchy type.

Theorem 3.7.65: *Let $\phi\in C^n([\alpha,\beta])$ and $\Lambda_z : C^n([\alpha,\beta])\to\mathbb{R}$, $z\in\{1,...,5\}$, be the linear functionals defined by (3.211). Then there exists $\xi_z\in[\alpha,\beta]$, $z\in\{1,...,5\}$, such that*

$$\Lambda_z(\phi)=\phi^{(n)}(\xi_z)\Lambda_z(\varphi),$$

where $\varphi(x)=\frac{x^n}{n!}$..

Proof: Similar to the proof of Theorem 3.1.8. □

Theorem 3.7.66: *Let $\phi, \psi \in C^n([\alpha, \beta])$ and $\Lambda_z : C^n([\alpha, \beta]) \to \mathbb{R}$, $z \in \{1, ..., 5\}$, be the linear functionals defined by (3.211). Then there exists $\xi_z \in [\alpha, \beta], z \in \{1, ..., 5\}$, such that*

$$\frac{\Lambda_z(\phi)}{\Lambda_z(\psi)} = \frac{\phi^{(n)}(\xi_z)}{\psi^{(n)}(\xi_z)}, \tag{3.212}$$

provided that the denominators are non-zero

Proof: The method of the proof is the same as in the proof of Theorem 3.1.9. □

Remark 3.7.25.: With assumption that $\frac{\phi^{(n)}}{\psi^{(n)}}$ is an invertible function, as a consequence of the previous theorem we get

$$\xi_z = \left(\frac{\phi^{(n)}}{\psi^{(n)}}\right)^{-1} \left(\frac{\Lambda_z(\phi)}{\Lambda_z(\psi)}\right), \quad z = 1, ..., 5,$$

new means of Chauchy type of the segment $[\alpha, \beta]$.

Let $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ be the families of exponentially convex functions as given in Subsection 3.1.3. in Definition 3.1.2..

Theorem 3.7.67: *Let Λ_z, $z = 1, ..., 5$, be the linear functional defined as in (3.211) associated with the family $\mathcal{F}_1$ (resp. $\mathcal{F}_2$). Then the following statements hold:*

(i) The functions $t \mapsto \Lambda_z(\phi_t)$, $z = 1, ..., 5$, are n-exponentially convex (resp. exponentially convex) in the Jensen sense on I.

(ii) If the functions $t \mapsto \Lambda_z(\phi_t)$, $z = 1, ..., 5$, are continuous on I, then n-exponentially convex (resp. exponentially convex) on I.

Proof: *The same proof method as for Theorem 3.1.10 is used so we omit it.* □

Corollary 3.7.16.: *Let Λ_z, $z = 1, ..., 5$, be the linear functional defined as in (3.211) associated with the family $\mathcal{F}_3$. Then the following statements hold:*

(i) If the functions $t \mapsto \Lambda_z(\phi_t)$, $z = 1, ..., 5$, are continuous on I, then they are 2-

exponentially convex on I. If $t \mapsto \Lambda_z(\phi_t)$, $z = 1, ..., 5$, are additionally positive, then they are also log-convex on I. Furthermore, for every choice $r, s, t \in I$, such that $r < s < t$, it holds

$$[\Lambda_z(\phi_s)]^{t-r} \leq [\Lambda_z(\phi_r)]^{t-s} [\Lambda_z(\phi_r)]^{s-r}.$$

(ii) If the functions $t \mapsto \Lambda_z(\phi_t)$, $z = 1, ..., 5$, are positive and differentiable on I, then for all $r, s, u, v \in I$ such that $r \leq u$, $s \leq v$, we have

$$M_{r,s}(\Lambda_z, \mathcal{F}_3) \leq M_{u,v}(\Lambda_z, \mathcal{F}_3), \tag{3.213}$$

where

$$M_{r,s}(\Lambda_z, \mathcal{F}_3) = \begin{cases} \left(\frac{\Lambda_z(\phi_r)}{\Lambda_z(\phi_s)}\right)^{\frac{1}{r-s}}, & r \neq s, \\ \exp\left(\frac{\frac{d}{dr}(\Lambda_z(\phi_r))}{\Lambda_z(\phi_r)}\right), & r = s. \end{cases} \tag{3.214}$$

Proof: The idea of the proof is the same as that of the proof of Corollary 3.1.4.. □

3.8. Sherman's inequality and the Abel-Gontscharoff interpolating polynomials

The techniques that we use in this section are based on application of interpolation by the Abel-Gontscharoff interpolation polynomials. The Abel-Gontscharoff interpolation problem in the real case was introduced in 1935 by Whittaker [156] and subsequently by Gontscharoff [57] and Davis [36]. Here we use notations and terminology from [11].

Let $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$, and let $\alpha \leq a_1 < a_2 < ... < a_n \leq \beta$ be the given points. For $\phi \in C^n([\alpha, \beta])$ **the Abel-Gontscharoff interpolating polynomial** P_{AG} of degree $(n-1)$ satisfying the Abel-Gontscharoff conditions

$$P_{AG}^{(i)}(a_{i+1}) = \phi^{(i)}(a_{i+1}), \quad 0 \leq i \leq n-1$$

exists uniquely.

This conditions in particular include two-point right focal conditions

$$\begin{aligned}
&P_{AG2}^{(i)}(a_1) = \phi^{(i)}(a_1), \ \ 0 \le i \le m, \\
&P_{AG2}^{(i)}(a_2) = \phi^{(i)}(a_2), \ \ m+1 \le i \le n-1, \ \alpha \le a_1 < a_2 \le \beta.
\end{aligned}$$

First, we give representations of the Abel-Gontscharoff interpolating polynomial. For details and proofs see [11].

Remark 3.8.26.: The Abel-Gontscharoff interpolating polynomial P_{AG} of the function $\phi \in C^n([\alpha, \beta])$ can be expressed as

$$P_{AG}(t) = \sum_{i=0}^{n-1} T_i(t)\phi^{(i)}(a_{i+1}), \tag{3.215}$$

where $T_0(t) = 1$ and T_i, $1 \le i \le n-1$ is the unique polynomial of degree i satisfying

$$\begin{aligned}
&T_i^{(k)}(a_{k+1}) = 0, \ 0 \le k \le i-1, \\
&T_i^{(i)}(a_{i+1}) = 1,
\end{aligned}$$

and it can be written as

$$\begin{aligned}
T_i(t) &= \frac{1}{1!2!\cdots i!}\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{i-1} & a_1^i \\ 0 & 1 & 2a_2 & \dots & (i-1)a_2^{i-2} & ia_2^{i-1} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & (i-1)! & i!a_i \\ 1 & t & t^2 & \dots & t^{i-1} & t^i \end{vmatrix} \\
&= \int_{a_1}^{t}\int_{a_2}^{t_1}\cdots\int_{a_i}^{t_{i-1}} dt_i dt_{i-1}\cdots dt_1, \ (t_0 = t).
\end{aligned} \tag{3.216}$$

In particular, we have

$$T_0(t) = 1$$
$$T_1(t) = t - a_1$$
$$T_2(t) = \frac{1}{2}\left[t^2 - 2a_2 t + a_1(2a_2 - a_1)\right].$$

The associated error $e_{AG}(t) = \phi(t) - P_{AG}(t)$ can be represented in terms of Green's function $G_{AG}(u,t)$ of the boundary value problem

$$z^{(n)} = 0,\ z^{(i)}(a_{i+1}) = 0,\ 0 \leq i \leq n-1$$

and appears as

$$G_{AG}(u,t) = \begin{cases} \sum\limits_{i=0}^{k-1} \frac{T_i(u)}{(n-i-1)!}(a_{i+1} - t)^{n-i-1}, & a_k \leq t \leq u \\ -\sum\limits_{i=k}^{n-1} \frac{T_i(u)}{(n-i-1)!}(a_{i+1} - t)^{n-i-1}, & u \leq t \leq a_{k+1} \\ & k = 0, 1, \ldots, n\ (a_0 = \alpha, a_{n+1} = \beta) \end{cases}. \tag{3.217}$$

Theorem 3.8.68: *Let $\phi \in C^n([\alpha,\beta])$, and let P_{AG} be its Abel-Gontscharoff interpolating polynomial given in (3.215). Then*

$$\begin{aligned} \phi(t) &= P_{AG}(t) + e_{AG}(t) \\ &= \sum_{i=0}^{n-1} T_i(t)\phi^{(i)}(a_{i+1}) + \int_{\alpha}^{\beta} G_{AG}(u,s)\phi^{(n)}(s)ds, \end{aligned} \tag{3.218}$$

where T_i is defined by (3.216) *and $G_{AG}(u,s)$ is defined by* (3.217).

Remark 3.8.27.: The two-point right focal interpolating polynomial P_{AG2} of the function $\phi \in C^n([\alpha,\beta])$ can be written as

$$\begin{aligned} P_{AG2}(t) &= \sum_{s=0}^{m} \frac{(t-a_1)^s}{s!}\phi^{(s)}(a_1) \\ &+ \sum_{r=0}^{n-m-2}\left[\sum_{s=0}^{r} \frac{(t-a_1)^{m+1+s}(a_1-a_2)^{r-s}}{(m+1+s)!\,(r-s)!}\right]\phi^{(m+1+r)}(a_2). \end{aligned} \tag{3.219}$$

The associated error $e_{AG2}(t) = \phi(t) - P_{AG2}(t)$ can be represented in terms of Green's function of the boundary value problem

$$z^{(n)} = 0,\ z^{(i)}(a_1) = 0,\ 0 \le i \le m,\ z^{(i)}(a_2) = 0, m+1 \le i \le n-1,$$

given by

$$G_{mn}(u,t) = \frac{1}{(n-1)!}\begin{cases} \sum\limits_{s=0}^{m}\binom{n-1}{s}(u-a_1)^s(a_1-t)^{n-s-1}\ , & \alpha \le t \le u; \\ -\sum\limits_{s=m+1}^{n-1}\binom{n-1}{s}(u-a_1)^s(a_1-t)^{n-s-1}, & u \le t \le \beta. \end{cases} \tag{3.220}$$

Theorem 3.8.69: *Let $n, m \in \mathbb{N}$, $n \ge 2$, $0 \le m \le n-1$ and $\phi \in C^n([\alpha,\beta])$. Let P_{AG2} be its two-point right focal Abel-Gontscharoff interpolating polynomial given by (3.219). Then*

$$\begin{aligned} \phi(t) &= P_{AG2}(t) + e_{AG2}(t) \\ &= \sum_{s=0}^{m} \frac{(t-a_1)^s}{s!}\phi^{(s)}(a_1) + \sum_{r=0}^{n-m-2}\left[\sum_{s=0}^{r} \frac{(t-a_1)^{m+1+s}(a_1-a_2)^{r-s}}{(m+1+s)!\,(r-s)!}\right]\phi^{(m+1+r)}(a_2) \\ &\quad + \int_{\alpha}^{\beta} G_{mn}(u,s)\phi^{(n)}(s)ds, \end{aligned} \tag{3.221}$$

where $G_{mn}(u,s)$ is defined by (3.220).

Remark 3.8.28.: Further, for $a_1 \le t, u \le a_2$ the following inequalities hold

$$\begin{aligned} (-1)^{n-m-1}\frac{\partial^s G_{mn}(u,t)}{\partial u^s} &\ge 0, \quad 0 \le s \le m, \\ (-1)^{n-s}\frac{\partial^s G_{mn}(u,t)}{\partial u^s} &\ge 0, \quad m+1 \le s \le n-1. \end{aligned}$$

3.8.1. Generalizations of Sherman's inequality by the Abel-Gontscharoff interpolating polynomials

Applying interpolation with the Abel-Gontscharoff polynomials we derive new identity for Sherman's difference $\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{k} b_i\phi(y_i)$ as follows.

Theorem 3.8.70: *Let* $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^k$, $\mathbf{a} \in \mathbb{R}^l$ *and* $\mathbf{b} \in \mathbb{R}^k$ *be such that*

$$\mathbf{y} = \mathbf{x}\mathbf{A}^T \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{A} \tag{3.222}$$

holds for some matrix $\mathbf{A} \in \mathcal{M}_{kl}(\mathbb{R})$ *whose entries satisfy the condition* $\sum_{j=1}^{l} a_{ij} = 1$, $i = 1, ..., k$. *Let* $n, m \in \mathbb{N}$, $n \geq 2$, $0 \leq m \leq n-1$, $\phi \in C^n([\alpha,\beta])$ *and* G_{mn} *be defined by (3.220). Then*

$$\begin{aligned}
&\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{k} b_i\phi(y_i) \qquad (3.223)\\
&= \sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^s - \sum_{i=1}^{k} b_i(y_i-\alpha)^s\right)\\
&+ \sum_{r=0}^{n-m-2}\sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s}\phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^{m+1+s} - \sum_{i=1}^{k} b_i(y_i-\alpha)^{m+1+s}\right)\\
&+ \int_{\alpha}^{\beta}\left(\sum_{j=1}^{l} a_j G_{mn}(x_j,t) - \sum_{i=1}^{k} b_i G_{mn}(y_i,t)\right)\phi^{(n)}(t)dt.
\end{aligned}$$

Proof: Using Theorem 3.8.69 we can represent every function $\phi \in C^n([\alpha,\beta])$ in the form

$$\begin{aligned}
\phi(u) &= \sum_{s=0}^{m} \frac{(u-\alpha)^s}{s!}\phi^{(s)}(\alpha) \\
&+ \sum_{r=0}^{n-m-2}\left[\sum_{s=0}^{r} \frac{(u-\alpha)^{m+1+s}(-1)^{r-s}(\beta-\alpha)^{r-s}}{(m+1+s)!(r-s)!}\right]\phi^{(m+1+r)}(\beta)\\
&+ \int_{\alpha}^{\beta} G_{mn}(u,t)\phi^{(n)}(t)dt.
\end{aligned} \tag{3.224}$$

By an easy calculation, using (3.224) in $\sum\limits_{j=1}^{l} a_j\phi(x_j) - \sum\limits_{i=1}^{k} b_i\phi(y_i)$, we get

$$\begin{aligned}
&\sum_{j=1}^{l} a_j\phi(x_j) - \sum_{i=1}^{k} b_i\phi(y_i) \\
&= \sum_{s=0}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^s - \sum_{i=1}^{k} b_i(y_i-\alpha)^s\right) \\
&+ \sum_{r=0}^{n-m-2}\sum_{s=0}^{r} \tfrac{(-1)^{r-s}(\beta-\alpha)^{r-s}\phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^{m+1+s} - \sum_{i=1}^{k} b_i(y_i-\alpha)^{m+1+s}\right) \\
&+ \int_{\alpha}^{\beta}\left(\sum_{j=1}^{l} a_j G_{mn}(x_j,t) - \sum_{i=1}^{k} b_i G_{mn}(y_i,t)\right)\phi^{(n)}(t)dt.
\end{aligned}$$

Since (3.222) holds, then

$$\sum_{s=0}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^s - \sum_{i=1}^{k} b_i(y_i-\alpha)^s\right) = 0, \quad m = 0, 1.$$

Therefore, it follows (3.223). □

Using the previous identity we obtain the following Sherman's type inequality for n-convex functions for real, not necessary nonnegative entries of weights **a**, **b** and matrix **A**.

Theorem 3.8.71: *Let the assumptions of Theorem 3.8.70 hold. Additionally, let ϕ be n-convex on $[\alpha, \beta]$ and*

$$\sum_{i=1}^{k} b_i G_{mn}(y_i, t) \leq \sum_{j=1}^{l} a_j G_{mn}(x_j, t), \quad t \in [\alpha, \beta]. \tag{3.225}$$

Then

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \phi\left(x_j\right) - \sum_{i=1}^{k} b_i \phi\left(y_i\right) \qquad (3.226)\\
&\geq \sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^s - \sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^s\right)\\
&+ \sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s}\phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^{m+1+s} - \sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^{m+1+s}\right).
\end{aligned}$$

If the reverse inequality in (3.225) holds, then the reverse inequality in (3.226) holds.

Proof: Since $\phi \in C^n([\alpha,\beta])$ is n-convex, then $\phi^{(n)} \geq 0$ on $[\alpha,\beta]$. Hence, we can apply Theorem 3.8.70 to get (3.226). □

Notice that when we take into account Sherman's condition of nonnegativity of weights $\mathbf{a}$, $\mathbf{b}$ and matrix $\mathbf{A}$, the assumption (3.225) is equivalent to the requirement that $G_{mn}(\cdot,t)$, $t \in [\alpha,\beta]$, must be convex on $[\alpha,\beta]$. So, for $n=2$ and $0 \leq m \leq 1$, the assumption (3.225) is immediately satisfied and then the inequality (3.226) holds. Moreover, in that case, the right hand side of (3.226) is equal to zero, so we have

$$\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{k} b_i \phi(y_i) \geq 0,$$

i.e. we get Sherman's inequality as direct consequence. For an arbitrary $n \geq 3$ and $0 \leq m \leq 1$, we use Remark 3.8.28., i.e. we consider the following inequality:

$$(-1)^{n-2}\frac{\partial^2 G_{mn}(u,t)}{\partial u^2} \geq 0.$$

Hence, the convexity of $G_{mn}(\cdot,t)$ depends of a parity of n. If n is even, then $\frac{\partial^2 G_{mn}(u,t)}{\partial u^2} \geq 0$, i.e. $G_{mn}(\cdot,t)$ is convex and assumption (3.225) is satisfied. Moreover, the inequality (3.226) holds. For odd n we get the reverse inequality. For all other choices, the following generalization holds.

Theorem 3.8.72: *Suppose that all the assumptions of Theorem 3.8.70 hold. Additionally, let* $\mathbf{a},\mathbf{b}$ *and* $\mathbf{A}$ *be with nonnegative entries and* $\phi \in C^n([\alpha,\beta])$ *be* n*-convex.*

(i) If $n-m$ is odd, then the inequality (3.226) holds.

(ii) If $n-m$ is even, then the reverse inequality in (3.226) holds.

Proof: (i) By Remark 3.8.28., the following inequality holds

$$(-1)^{n-m-1}\frac{\partial^2 G_{mn}(u,t)}{\partial u^2} \geq 0, \quad \alpha \leq u, t \leq \beta.$$

In case $n-m$ is odd ($n-m-1$ is even), we have

$$\frac{\partial^2 G_{mn}(u,t)}{\partial u^2} \geq 0,$$

i.e. $G_{mn}(\cdot, t)$, $t \in [\alpha, \beta]$, is convex on $[\alpha, \beta]$. Then by Sherman's theorem we have

$$\sum_{i=1}^{k} b_i G_{mn}(y_i, t) \leq \sum_{j=1}^{l} a_j G_{mn}(x_j, t),$$

i.e. the assumption (3.225) is satisfied. Hence, applying Theorem 3.8.71 we get (3.226).

(ii) Similarly we can prove this part. □

Theorem 3.8.73: *Suppose that all the assumptions of Theorem 3.8.70 hold. Additionally, let* $\mathbf{a}, \mathbf{b}$ *and* $\mathbf{A}$ *be with nonnegative entries and* $\phi \in C^n([\alpha, \beta])$ *be n-convex. Let* $F : [\alpha, \beta] \to \mathbb{R}$ *be defined by*

$$\begin{aligned} F(t) &= \sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}(t-\alpha)^s \\ &+ \sum_{r=0}^{n-m-2}\sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s}}{(m+1+s)!\,(r-s)!}\phi^{(m+1+r)}(\beta)(t-\alpha)^{m+1+s}. \end{aligned} \tag{3.227}$$

(i) If (3.226) holds and F is convex, then

$$\sum_{j=1}^{l} a_j \phi(x_j) - \sum_{i=1}^{k} b_i \phi(y_i) \geq 0. \tag{3.228}$$

(ii) If the reverse of (3.226) holds and F is concave, then the reverse inequality in (3.228) holds.

Proof: (i) Let (3.226) holds. If F is convex, then by Sherman's theorem we have

$$\sum_{j=1}^{l} a_j F(x_j) - \sum_{i=1}^{k} b_i F(y_i) \geq 0,$$

which, changing the order of summation, can be written in form

$$\sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!} \left(\sum_{j=1}^{l} a_j (x_j - \alpha)^s - \sum_{i=1}^{k} b_i (y_i - \alpha)^s \right) +$$
$$\sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s}\phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!} \left(\sum_{j=1}^{l} a_j (x_j - \alpha)^{m+1+s} - \sum_{i=1}^{k} b_i (y_i - \alpha)^{m+1+s} \right)$$
$$\geq 0.$$

Therefore, the right-hand side of (3.226) is nonnegative and the inequality (3.228) immediately follows.

(ii) Similarly we can prove this part. □

Remark 3.8.29.: Note that the function $x \mapsto (x-\alpha)^p$ is convex on $[\alpha, \beta]$ for each $p = 2, ..., n-1$, i.e. $\sum_{j=1}^{l} a_j (x_j - \alpha)^p - \sum_{i=1}^{k} b_i (y_i - \alpha)^p \geq 0$, for each $p = 2, ..., n-1$.

(i) If (3.226) holds and in addition $\phi^{(s)}(\alpha) \geq 0$ for $s = 0, ..., m$ and $\phi^{(m+1+s)}(\beta) \geq 0$ if $r - s$ is even and $\phi^{(m+1+s)}(\beta) \leq 0$ if $r - s$ is odd for $s = 0, ..., r$ and $r = 0, ..., n-m-2$, then the right hand side of (3.226) is nonnegative, i.e. the inequality (3.228) holds.

(ii) If the reverse of (3.226) holds and in addition $\phi^{(s)}(\alpha) \leq 0$ for $s = 0, ..., m$ and $\phi^{(m+1+s)}(\beta) \leq 0$ if $r - s$ is even, and $\phi^{(m+1+s)}(\beta) \geq 0$ if $r - s$ is odd for $s = 0, ..., r$ and $r = 0, ..., n-m-2$, then the right hand side of (3.226) is negative, i.e. the reverse inequality in (3.228) holds.

3.8.2. Upper bound for generalized Sherman's inequality

As we noticed in Remark 3.8.29., under the special conditions, the following estimations for Sherman's difference holds:

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \phi\left(x_j\right)-\sum_{i=1}^{k} b_i \phi\left(y_i\right) \\
&\geq \sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^s-\sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^s\right) \\
&+\sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \tfrac{(-1)^{r-s}(\beta-\alpha)^{r-s} \phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^{m+1+s}-\sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^{m+1+s}\right) \\
&\geq 0 .
\end{aligned}$$

In this section we present upper bounds for the obtained generalization.

To avoid many notations, we define the function $\mathcal{F}:[\alpha, \beta] \rightarrow \mathbb{R}$ by

$$\mathcal{F}(t)=\sum_{j=1}^{l} a_j G_{mn}\left(x_j, t\right)-\sum_{i=1}^{k} b_i G_{mn}\left(y_i, t\right), \tag{3.229}$$

under assumptions of Theorem 3.8.70. We also consider the Čebyšev functional

$$T(\mathcal{F}, \mathcal{F})=\frac{1}{\beta-\alpha} \int_\alpha^\beta \mathcal{F}^2(t) dt-\left(\frac{1}{\beta-\alpha} \int_\alpha^\beta \mathcal{F}(t) dt\right)^2 .$$

To prove the following estimations we use results related to the Čebyšev functional on an analogous way as in the previous section. Therefore, we give theorem statements without proof.

Theorem 3.8.74: *Suppose that all the assumptions of Theorem 3.8.70 hold. Additionally, let $(\cdot-\alpha)(\beta-\cdot)(\phi^{(n+1)})^2 \in L_1[\alpha, \beta]$ and $\mathcal{F}$ be defined as in (3.229). Then the following* *identity holds:*

$$\sum_{j=1}^{l} a_j \phi\left(x_j\right) - \sum_{i=1}^{k} b_i \phi\left(y_i\right) \tag{3.230}$$
$$= \sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!} \left[\sum_{j=1}^{l} a_j \left(x_j - \alpha\right)^s - \sum_{i=1}^{k} b_i \left(y_i - \alpha\right)^s\right]$$
$$+ \sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s}\phi^{(m+1+r)}(\beta)}{(s+1+m)!(r-s)!} \left[\sum_{j=1}^{l} a_j \left(x_j - \alpha\right)^{m+1+s} - \sum_{i=1}^{k} b_i \left(y_i - \alpha\right)^{m+1+s}\right]$$
$$+ \frac{\phi^{(n-1)}(\beta) - \phi^{(n-1)}(\alpha)}{\beta - \alpha} \int_{\alpha}^{\beta} \mathcal{F}(t)dt + R_n(\alpha, \beta; \phi),$$

where the remainder $R_n(\alpha, \beta; \phi)$ satisfies the estimation

$$|R_n(\alpha, \beta; \phi)| \leq \sqrt{\frac{\beta - \alpha}{2}} \left[T(\mathcal{F}, \mathcal{F})\right]^{\frac{1}{2}} \left(\int_{\alpha}^{\beta} (t-\alpha)(\beta - t) \left[\phi^{(n+1)}(t)\right]^2 dt\right)^{\frac{1}{2}}.$$

The following Grüss type inequality also holds.

Theorem 3.8.75: *Suppose that all the assumptions of Theorem 3.8.70 hold. Additionally, let $\phi^{(n+1)} \geq 0$ on $[\alpha, \beta]$ and $\mathcal{F}$ be defined as in (3.229). Then the identity (3.230) holds and the remainder $R(\phi; a, b)$ satisfies the bound*

$$|R_n(\alpha, \beta; \phi)| \leq \left\|\mathcal{F}'\right\|_{\infty} \left[\frac{\phi^{(n-1)}(\beta) + \phi^{(n-1)}(\alpha)}{2} - \frac{\phi^{(n-2)}(\beta) - \phi^{(n-2)}(\alpha)}{\beta - \alpha}\right]. \tag{3.231}$$

We present an upper bound for generalized Sherman's inequality which is the Ostrowski type inequality.

Theorem 3.8.76: *Suppose that all the assumptions of Theorem 3.8.70 hold. Let (p, q) be* *a pair of conjugate exponents, that is* $1 \leq p, q \leq \infty$, $\frac{1}{p}+\frac{1}{q}=1$. *Then*

$$\begin{aligned}
&\left|\sum_{j=1}^{l} a_j \phi\left(x_j\right)-\sum_{i=1}^{k} b_i\, \phi\left(y_i\right)\right. \\
&-\sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left[\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^s-\sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^s\right] \\
&\left.-\sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s} \phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j\left(x_j-\alpha\right)^{m+1+s}-\sum_{i=1}^{k} b_i\left(y_i-\alpha\right)^{m+1+s}\right)\right| \\
&\leq\left\|\phi^{(n)}\right\|_p\left(\int_\alpha^\beta\left|\sum_{j=1}^{l} a_j G_{mn}\left(x_j, t\right)-\sum_{i=1}^{k} b_i G_{mn}\left(y_i, t\right)\right|^q dt\right)^{\frac{1}{q}} .
\end{aligned} \tag{3.232}$$

The constant on the right-hand side of (3.232) is sharp for $1<p \leq \infty$ *and the best possible for* $p=1$.

Proof: Applying the well-known Hölder inequality to the identity (3.223) we have

$$\begin{aligned}
&\left|\sum_{j=1}^{l} a_j \phi(x_j)-\sum_{i=1}^{k} b_i \phi(y_i)\right. \\
&-\sum_{s=2}^{m} \frac{\phi^{(s)}(\alpha)}{s!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^s-\sum_{i=1}^{k} b_i(y_i-\alpha)^s\right) \\
&\left.-\sum_{r=0}^{n-m-2} \sum_{s=0}^{r} \frac{(-1)^{r-s}(\beta-\alpha)^{r-s} \phi^{(m+1+r)}(\beta)}{(m+1+s)!(r-s)!}\left(\sum_{j=1}^{l} a_j(x_j-\alpha)^{m+1+s}-\sum_{i=1}^{k} b_i(y_i-\alpha)^{m+1+s}\right)\right| \\
&=\left|\int_\alpha^\beta\left(\sum_{j=1}^{l} a_j G_{mn}(x_j, t)-\sum_{i=1}^{k} b_i G_{mn}(y_i, t)\right) \phi^{(n)}(t) dt\right| \leq\left\|\phi^{(n)}\right\|_p\left(\int_\alpha^\beta|\mathcal{F}(t)|^q dt\right)^{\frac{1}{q}},
\end{aligned} \tag{3.233}$$

where we denote

$$\mathcal{F}(t)=\sum_{j=1}^{l} a_j G_{mn}\left(x_j, t\right)-\sum_{i=1}^{k} b_i G_{mn}\left(y_i, t\right). \tag{3.234}$$

To prove the sharpness of the constant $\left(\int_\alpha^\beta|\mathcal{F}(t)|^q dt\right)^{\frac{1}{q}}$ we proceed analogously as in the proof of Theorem 3.1.7. □

In the sequel we consider a particular case of Green's function $G_{mn}(u, t)$ defined by

(3.220). For $n=2$, $m=1$, we have

$$G_{12}(u,t)=\begin{cases} u-t, & \alpha\le t\le u \\ 0, & u\le t\le\beta \end{cases}, \tag{3.235}$$

and

$$\begin{aligned} &\sum_{j=1}^{l} a_j G_{12}(x_j,t)-\sum_{i=1}^{k} b_i G_{12}(y_i,t) \\ &=\sum_{\mathbf{j}} a_{\mathbf{j}}(x_{\mathbf{j}}-t)-\sum_{\mathbf{i}} b_{\mathbf{i}}(y_{\mathbf{i}}-t), \quad \mathbf{j}\in\{j;x_j\ge t\},\mathbf{i}\in\{i;y_i\ge t\}. \end{aligned} \tag{3.236}$$

As an easy consequence of Theorem 3.8.76, choosing $n=2$ and $m=1$, we get the following corollary.

Corollary 3.8.17.: *Let $\phi\in C^2([\alpha,\beta])$, $\mathbf{x}\in[\alpha,\beta]^l$, $\mathbf{y}\in[\alpha,\beta]^k$, $\mathbf{a}\in\mathbb{R}^l$ and $\mathbf{b}\in\mathbb{R}^k$ be such that (3.222) holds for some matrix $\mathbf{A}\in\mathcal{M}_{kl}(\mathbb{R})$ whose entries satisfy the condition $\sum_{j=1}^{l} a_{ij}=1$, $i=1,...,k$. Let (p,q) be a pair of conjugate exponents, that is $1\le p,q\le\infty$, $\frac{1}{p}+\frac{1}{q}=1$. Then*

$$\begin{aligned} &\left|\sum_{j=1}^{l} a_j\phi(x_j)-\sum_{i=1}^{k} b_i\phi(y_i)\right| \\ &\le \left\|\phi''\right\|_p\left(\int_\alpha^\beta\left(\sum_{j=1}^{l} a_j G_{12}(x_j,t)-\sum_{i=1}^{k} b_i G_{12}(y_i,t)\right)^q dt\right)^{\frac{1}{q}}. \end{aligned} \tag{3.237}$$

The constant on the right hand side of (3.237) is sharp for $1<p\le\infty$ and the best possible for $p=1$.

Remark 3.8.30.: If additionally suppose that vectors $\mathbf{a}$, $\mathbf{b}$ and matrix $\mathbf{A}$ are nonnegative and ϕ is convex, then the difference $\sum_{j=1}^{l} a_j\phi(x_j)-\sum_{i=1}^{k} b_i\phi(y_i)$ is nonnegative and we have

$$
\begin{aligned}
0 &\leq \sum_{j=1}^{l} a_j \phi\left(x_j\right) - \sum_{i=1}^{k} b_i \phi\left(y_i\right) \\
&\leq \left\|\phi''\right\|_p \left(\int_{\alpha}^{\beta} \left(\sum_{j=1}^{l} a_j G_{12}\left(x_j, t\right) - \sum_{i=1}^{k} b_i G_{12}\left(y_i, t\right)\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned} \tag{3.238}
$$

In sequel we consider some particular case of this result.

(i) If $k = l$ and all weights b_i and a_j are equal, we get the following estimate for weighted majorization inequality:

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{k} a_i \phi\left(x_i\right) - \sum_{i=1}^{k} a_i \phi\left(y_i\right) \\
&\leq \left\|\phi''\right\|_p \left(\int_{\alpha}^{\beta} \left(\sum_{i=1}^{k} a_i G_{12}\left(x_i, t\right) - \sum_{i=1}^{k} a_i G_{12}\left(y_i, t\right)\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned} \tag{3.239}
$$

(ii) If we denote $A_k = \sum_{i=1}^{k} a_i$ and put $y_1 = y_2 = \ldots = y_k = \frac{1}{A_k} \sum_{i=1}^{k} a_i x_i = \bar{x}$, from (3.239) as an easy consequence we get estimate for Jensen's inequality:

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{k} a_i \phi\left(x_i\right) - A_k \phi\left(\bar{x}\right) \\
&\leq \left\|\phi''\right\|_p \left(\int_{\alpha}^{\beta} \left(\sum_{i=1}^{k} a_i G_{12}\left(x_i, t\right) - A_k G_{12}\left(\bar{x}, t\right)\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned}
$$

Specially, setting $A_k = 1$, we have $\bar{x} = \sum_{i=1}^{k} a_i x_i$ and

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{k} a_i \phi\left(x_i\right) - \phi\left(\bar{x}\right) \\
&\leq \left\|\phi''\right\|_p \left(\int_{\alpha}^{\beta} \left(\sum_{i=1}^{k} a_i G_{12}\left(x_i, t\right) - G_{12}\left(\bar{x}, t\right)\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned}
$$

(iii) For $k=2$, $a_1=a_2=1$, $x_1=\alpha$ and $x_2=\beta$, we have

$$\begin{aligned}
0 &\leq \phi(\alpha)+\phi(\beta)-2\phi\left(\frac{\alpha+\beta}{2}\right) \\
&\leq \left\|\phi''\right\|_p \left(\int_\alpha^\beta \left(G_{12}(\alpha,t)+G_{12}(\beta,t)-2G_{12}\left(\frac{\alpha+\beta}{2},t\right)\right)^q dt\right)^{\frac{1}{q}},
\end{aligned}$$

where

$$G_{12}(\alpha,t)+G_{12}(\beta,t)-2G_{12}\left(\frac{\alpha+\beta}{2},t\right)=\begin{cases} t-\alpha, & \alpha\leq t\leq \frac{\alpha+\beta}{2} \\ \beta-t, & \frac{\alpha+\beta}{2}\leq t\leq\beta \end{cases}.$$

Specially, for $p=1$, $q=\infty$, we obtain

$$0\leq \frac{\phi(\alpha)+\phi(\beta)}{2}-\phi\left(\frac{\alpha+\beta}{2}\right)\leq \frac{1}{4}\left(\phi'(\beta)-\phi'(\alpha)\right)(\beta-\alpha),$$

where the constant $\frac{1}{4}$ is the best possible in the sense that it cannot be replaced by a smaller constant.

3.8.3. Applications to means

In this subsection we discuss the application of Corollary 3.8.17., i.e. the inequality (3.238), to the lower and upper bounds estimations of some relationships between well-known means.

Let $0<\alpha<\beta$, $\mathbf{x}\in[\alpha,\beta]^l$ and $\mathbf{a}\in[0,\infty)^l$ whose entries satisfy the condition $\sum_{j=1}^l a_j=1$. We consider the weighted power mean of order $s\in\mathbb{R}$ defined by

$$M_s(\mathbf{x},\mathbf{a})=\begin{cases} \left(\sum\limits_{j=1}^l a_j x_j^s\right)^{\frac{1}{s}}, & s\neq 0 \\ \prod\limits_{j=1}^l x_j^{a_j}, & s=0 \end{cases}.$$

The classical weighted means are defined as

$$M_0(\mathbf{x},\mathbf{a}){=}G(\mathbf{x},\mathbf{a}) = \prod_{j=1}^{l} x_j^{a_j}, \quad \text{the weighted geometric mean,}$$
$$M_1(\mathbf{x},\mathbf{a}){=}A(\mathbf{x},\mathbf{a}) = \sum_{j=1}^{l} a_j x_j, \quad \text{the weighted arithmetic mean,}$$
$$M_{-1}(\mathbf{x},\mathbf{a}){=}H(\mathbf{x},\mathbf{a}) = \frac{1}{\sum_{j=1}^{l} \frac{a_j}{x_j}}, \quad \text{the weighted harmonic mean.}$$

In an analogous way, for $\mathbf{y} \in [\alpha,\beta]^k$, $\mathbf{b} \in [0,\infty)^k$ with $\sum_{i=1}^k b_i = 1$, we define

$$M_s(\mathbf{y},\mathbf{b}) = \begin{cases} \left(\sum\limits_{i=1}^{k} b_i y_i^s\right)^{\frac{1}{s}}, & s \neq 0 \\ \prod\limits_{i=1}^{k} y_i^{b_i}, & s = 0 \end{cases}$$

and accordingly classical weighted means $G(\mathbf{y},\mathbf{b})$, $A(\mathbf{y},\mathbf{b})$ and $H(\mathbf{y},\mathbf{b})$.

Corollary 3.8.18.: *Let $s \geq 2$ and $0 < \alpha < \beta$. Let $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^k$, $\mathbf{a} \in [0,\infty)^l$ and $\mathbf{b} \in [0,\infty)^k$ be such that (3.222) holds for some row stochastic matrix $\mathbf{A} \in \mathcal{M}_{kl}(\mathbb{R})$ and entries of $\mathbf{a}$ and $\mathbf{b}$ satisfy the condition $\sum_{j=1}^l a_j = \sum_{i=1}^k b_i = 1$. Then*

$$0 \leq M_s^s(\mathbf{x},\mathbf{a}) - M_s^s(\mathbf{y},\mathbf{b}) \leq s(s-1)\left(\int_\alpha^\beta t^{p(s-2)}dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\,,$$
$$0 \leq \frac{M_s^s(\mathbf{x},\mathbf{a}) - M_s^s(\mathbf{y},\mathbf{b})}{s(s-1)} \leq \left(\int_\alpha^\beta t^{p(s-2)}dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\,,$$
$$1 \leq \frac{G(\mathbf{y},\mathbf{b})}{G(\mathbf{x},\mathbf{a})} \leq \exp\left[\left(\int_\alpha^\beta \left(\frac{1}{t^2}\right)^p dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\right],$$
$$0 \leq \frac{H(\mathbf{y},\mathbf{b}) - H(\mathbf{y},\mathbf{b})}{H(\mathbf{x},\mathbf{a})H(\mathbf{y},\mathbf{b})} \leq \left(\int_\alpha^\beta \left(\frac{2}{t^3}\right)^p dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\,,$$

where $\mathcal{G}(t)$ denotes the difference (3.236).

The constants on the right-hand side of inequalities are sharp for $1 < p \leq \infty$ and the best possible for $p = 1$.

Proof: Applying (3.238) to the functions $\phi(x) = x^s$, $\phi(x) = \frac{x^s}{s(s-1)}$, $\phi(x) = -\ln x$ and $\phi(x) = \frac{1}{x}$, respectively, we get required results. □

Remark 3.8.31.: Particular cases of the previous results for $p = 1$, $q = \infty$:

$$0 \le M_s^s(\mathbf{x},\mathbf{a}) - M_s^s(\mathbf{y},\mathbf{b}) \le s(\beta^{s-1} - \alpha^{s-1}) \|\mathcal{G}\|_\infty ,$$

$$0 \le \frac{M_s^s(\mathbf{x},\mathbf{a}) - M_s^s(\mathbf{y},\mathbf{b})}{s} \le (\beta^{s-1} - \alpha^{s-1}) \|\mathcal{G}\|_\infty ,$$

$$1 \le \frac{G(\mathbf{y},\mathbf{b})}{G(\mathbf{x},\mathbf{a})} \le \exp\left(\frac{\beta-\alpha}{\alpha\beta}\|\mathcal{G}\|_\infty\right),$$

$$0 \le \frac{H(\mathbf{y},\mathbf{b}) - H(\mathbf{x},\mathbf{a})}{H(\mathbf{x},\mathbf{a})H(\mathbf{y},\mathbf{b})} \le \left(\frac{\beta^2-\alpha^2}{\alpha^2\beta^2}\right)\|\mathcal{G}\|_\infty .$$

Replacing $x_j \to \frac{1}{x_j}$, $y_i \to \frac{1}{y_i}$, we have $A(\mathbf{x},\mathbf{a}) \to \frac{1}{H(\mathbf{x},\mathbf{a})}$, $A(\mathbf{y},\mathbf{b}) \to \frac{1}{H(\mathbf{y},\mathbf{b})}$, and the last inequality becomes

$$0 \le A(\mathbf{x},\mathbf{a}) - A(\mathbf{y},\mathbf{b}) \le \left(\frac{\beta-\alpha}{\alpha^2\beta^2}\right)\left\|\tilde{\mathcal{G}}\right\|_\infty ,$$

where $\tilde{\mathcal{G}}(t) = \sum_{\mathbf{j}} a_{\mathbf{j}}(\frac{1}{x_{\mathbf{j}}} - t) - \sum_{\mathbf{i}} b_{\mathbf{i}}(\frac{1}{y_{\mathbf{i}}} - t)$, $\mathbf{j} \in \{j; \frac{1}{x_j} \ge t\}$, $\mathbf{i} \in \{i; \frac{1}{y_i} \ge t\}$.

Corollary 3.8.19.: *Let $0 < \alpha < \beta$ and $\mathbf{x} \in [\alpha,\beta]^l$, $\mathbf{y} \in [\alpha,\beta]^k$, $\mathbf{a} \in [0,\infty)^l$ and $\mathbf{b} \in [0,\infty)^k$ be such that (3.222) holds for some row stochastic matrix $\mathbf{A} \in \mathcal{M}_{kl}(\mathbb{R})$ and entries of $\mathbf{a}$ and $\mathbf{b}$ satisfy the condition $\sum_{j=1}^l a_j = \sum_{i=1}^k b_i = 1$. Then*

$$1 \le \frac{\prod_{j=1}^l x_j^{a_j x_j}}{\prod_{i=1}^k y_i^{b_i y_i}} \le \exp\left[\left(\int_\alpha^\beta \left(\frac{1}{t}\right)^p dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\right],$$

$$0 \le \sum_{j=1}^l a_j e^{x_j} - \sum_{i=1}^k b_i e^{y_i} \le \left(\int_\alpha^\beta e^{tp} dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q ,$$

$$1 \le \frac{\prod_{j=1}^l (1+e^{x_j})^{a_j}}{\prod_{i=1}^k (1+e^{y_i})^{b_i}} \le \exp\left(\left(\int_\alpha^\beta \left(\frac{e^t}{(1+e^t)^2}\right)^p dt\right)^{\frac{1}{p}} \|\mathcal{G}\|_q\right),$$

where $\mathcal{G}(t)$ denotes the difference (3.236).

The constants on the right-hand side of inequalities are sharp for $1 < p \le \infty$ and the best possible for $p = 1$.

3.9. New bounds and identities for Sherman's difference

In this section we obtain refinement of Sherman's generalization for convex functions (2-convex functions). Using Green's functions we introduce new identities that include Sherman's difference, deduced from Sherman's inequality, which enable us to extend Sherman's results to the class of convex functions of higher order, i.e. to n-convex functions ($n \geq 3$). The results from this section are quoted from [71].

Denote with $\mathcal{C}_{kl}([0,\infty))$ the space of all $k \times l$ column stochastic matrices. For pairs of vectors with weights $(\mathbf{y},\mathbf{b})$ and $(\mathbf{x},\mathbf{a})$, which satisfy Sherman's conditions (2.6) for some $\mathbf{S} \in \mathcal{C}_{kl}([0,\infty))$, we use notation $(\mathbf{y},\mathbf{b}) \prec (\mathbf{x},\mathbf{a})$.

Theorem 3.9.77: *Let $\mathbf{x} \in [\alpha,\beta]^k$, $\mathbf{y} \in [\alpha,\beta]^l$, $\mathbf{a} \in [0,\infty)^k$, $\mathbf{b} \in [0,\infty)^l$ and $\mathbf{S} \in \mathcal{C}_{kl}([0,\infty))$ be such that $(\mathbf{y},\mathbf{b}) \prec (\mathbf{x},\mathbf{a})$. Then for every convex function $f\colon [\alpha,\beta] \to \mathbb{R}$, we have*

$$
\begin{aligned}
0 \leq & m\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right) \leq \sum_{i=1}^{k} a_i f(x_i) - \sum_{j=1}^{l} b_j f(y_j) \\
\leq & M\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right),
\end{aligned}
\tag{3.240}
$$

where $m = \sum_{j=1}^{l} b_j \min_{1\leq i\leq k}\{s_{ij}\}$, $M = \sum_{j=1}^{l} b_j \max_{1\leq i\leq k}\{s_{ij}\}$ and $\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i$.

Proof: Since $(\mathbf{y},\mathbf{b}) \prec (\mathbf{x},\mathbf{a})$, i.e. Sherman's conditions (2.6) hold, we have

$$
\begin{aligned}
\sum_{i=1}^{k} a_i f(x_i) - \sum_{j=1}^{l} b_j f(y_j) &= \sum_{i=1}^{k}\sum_{j=1}^{l} b_j s_{ij} f(x_i) - \sum_{j=1}^{l} b_j f\left(\sum_{i=1}^{k} x_i s_{ij}\right) \\
&= \sum_{j=1}^{l} b_j \left(\sum_{i=1}^{k} s_{ij} f(x_i) - f\left(\sum_{i=1}^{k} x_i s_{ij}\right)\right).
\end{aligned}
$$

Applying Lemma 2.3.2. to the difference $\sum\limits_{i=1}^{k} s_{ij} f(x_i) - f\left(\sum\limits_{i=1}^{k} x_i s_{ij}\right)$, for each $j = 1, ..., l$, we get

$$
\begin{aligned}
0 \leq & \min_{1\leq i\leq k}\{s_{ij}\}\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right) \\
\leq & \sum_{i=1}^{k} s_{ij} f(x_i) - f\left(\sum_{i=1}^{k} x_i s_{ij}\right) \leq \max_{1\leq i\leq k}\{s_{ij}\}\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right).
\end{aligned}
$$

Now, multiplying with b_j and summing over $j = 1, ..., l$, we get (3.240). $\square$

Remark 3.9.32.: If we take $k = l$ and $\mathbf{a} = \mathbf{b} = \mathbf{e} = (1, ..., 1) \in [0, \infty)^k$, then (3.240) becomes

$$0 \leq \tilde{m}\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right) \leq \sum_{i=1}^{k} f(x_i) - \sum_{i=1}^{k} f(y_i)$$
$$\leq \tilde{M}\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right),$$

where $\tilde{m} = \sum_{j=1}^{l} \min_{1 \leq i \leq k} \{s_{ij}\}$, $\tilde{M} = \sum_{j=1}^{l} \max_{1 \leq i \leq k} \{s_{ij}\}$, i.e. we get refinement of Majorization inequality (2.4).

To avoid many notations in further results, motivated by the inequalities (3.240), we define two functionals as follows.

Definition 3.9.6.: Let $\mathbf{x} \in [\alpha, \beta]^k$, $\mathbf{y} \in [\alpha, \beta]^l$, $\mathbf{a} \in \mathbb{R}^k$ *and* $\mathbf{b} \in \mathbb{R}^l$ be such that

$$\mathbf{y} = \mathbf{xS} \quad \text{and} \quad \mathbf{a} = \mathbf{bS}^T, \tag{3.241}$$

holds for some matrix $\mathbf{S} \in \mathcal{M}_{kl}(\mathbb{R})$ with columns sums equal to 1. Let $m = \sum_{j=1}^{l} b_j \min_{1 \leq i \leq k} \{s_{ij}\}$, $M = \sum_{j=1}^{l} b_j \max_{1 \leq i \leq k} \{s_{ij}\}$ and $\bar{x} = \frac{1}{k} \sum_{i=1}^{k} x_i$. We define the functionals F_m and F_M on $\mathbb{R}^{[\alpha,\beta]}$ as follows:

$$F_m(\varphi) = \sum_{i=1}^{k} a_i \varphi(x_i) - \sum_{j=1}^{l} b_j \varphi(y_j) - m\left(\sum_{i=1}^{k} \varphi(x_i) - k\varphi(\bar{x})\right), \tag{3.242}$$

$$F_M(\varphi) = M\left(\sum_{i=1}^{k} \varphi(x_i) - k\varphi(\bar{x})\right) - \sum_{i=1}^{k} a_i \varphi(x_i) + \sum_{j=1}^{l} b_j \varphi(y_j). \tag{3.243}$$

Remark 3.9.33.: In Definition 3.9.6. we didn't assume the nonnegativity of the entries of $\mathbf{a}, \mathbf{b}$ and $\mathbf{S}$. But if we take into account nonnegativity, i.e. under the assumptions of Theorem 3.9.77, then for every convex function $f \colon [\alpha, \beta] \to \mathbb{R}$, we have

$$F_m(f) \geq 0 \quad \text{and} \quad F_M(f) \geq 0.$$

Next we introduce two technical lemmas which give us new identities involving Green's functions defined in (3.44)-(3.48).

Lemma 3.9.4.: *Let the functionals F_m and F_M be defined as in Definition 3.9.6.. Let $s \in [\alpha, \beta]$ and $G_z(\cdot, s), z \in \{1, 2, 3, 4, 5\}$, be defined as in (3.44)-(3.48), respectively. Then for every $f\colon [\alpha, \beta] \to \mathbb{R}$, $f \in C^2([\alpha, \beta])$, the following identities hold:*

(i)

$$F_m(f) = \int_\alpha^\beta F_m(G_z(\cdot, s))f''(s)ds; \tag{3.244}$$

(ii)

$$F_M(f) = \int_\alpha^\beta F_M(G_z(\cdot, s))f''(s)ds. \tag{3.245}$$

Proof: By an easy calculation, using representations (3.49)-(3.53) and, because of (3.241),

$$\sum_{j=1}^{l} b_j - \sum_{i=1}^{k} a_i = \sum_{j=1}^{l} b_j - \sum_{i=1}^{k}\sum_{j=1}^{l} b_j s_{ij} = \sum_{j=1}^{l} b_j - \sum_{j=1}^{l} b_j \sum_{i=1}^{k} s_{ij} = 0,$$
$$\sum_{i=1}^{k} a_i x_i - \sum_{j=1}^{l} b_j y_j = \sum_{i=1}^{k}\sum_{j=1}^{l} b_j s_{ij} x_i - \sum_{j=1}^{l} b_j \sum_{i=1}^{k} x_i s_{ij} = 0, \tag{3.246}$$

we obtain the identities

$$\begin{aligned}&\sum_{i=1}^{k} a_i f(x_i) - \sum_{j=1}^{l} b_j f(y_j) - m\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right)\\ &= \int_\alpha^\beta \left[\sum_{i=1}^{k} a_i G_2(x_i, s) - \sum_{j=1}^{l} b_j G_2(y_j, s) - m\left(\sum_{i=1}^{k} G_2(x_i, s) - kG_2(\bar{x}, s)\right)\right] f''(s)ds\end{aligned}$$

and

$$\begin{aligned}&M\left(\sum_{i=1}^{k} f(x_i) - kf(\bar{x})\right) - \sum_{i=1}^{k} a_i f(x_i) + \sum_{j=1}^{l} b_j f(y_j)\\ &= \int_\alpha^\beta \left[M\left(\sum_{i=1}^{k} G_2(x_i, s) - kG_2(\bar{x}, s)\right) - \sum_{i=1}^{k} a_i G_2(x_i, s) + \sum_{j=1}^{l} b_j G_2(y_j, s)\right] f''(s)ds\end{aligned}$$

which are equivalent to (3.244) and (3.245). □

Lemma 3.9.5.: *Let the functionals F_m and F_M be defined as in Definition 3.9.6.. Let $s \in [\alpha, \beta]$ and $G_z(\cdot, s), z \in \{1, 2, 3, 4, 5\}$, be defined as in (3.44)-(3.48), respectively. Let $n \geq 3$ and $f\colon [\alpha, \beta] \to \mathbb{R}$ be such that $f^{(n-1)} \in AC([\alpha, \beta])$.*

i) For each $t \in [\alpha, \beta]$, such that $t \leqslant s$, the following identities hold

$$F_m(f) = \frac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds \right) f^{(n)}(t)dt + \sum_{w=0}^{n-3} \frac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds, \tag{3.247}$$

$$F_M(f) = \frac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds \right) f^{(n)}(t)dt + \sum_{w=0}^{n-3} \frac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds; \tag{3.248}$$

ii) For each $t \in [\alpha, \beta]$, such that $t > s$, the following identities hold

$$F_m(f) = \frac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\beta)ds \right) f^{(n)}(t)dt + \sum_{w=0}^{n-3} \frac{n-w-2}{w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds, \tag{3.249}$$

$$F_M(f) = \frac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\beta)ds \right) f^{(n)}(t)dt + \sum_{w=0}^{n-3} \frac{n-w-2}{w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds \tag{3.250}$$

Proof: i) Case $\alpha \leqslant t \leqslant s \leqslant \beta$: Applying (2.230) to f'' and to $n-2$ in place of n, we have

$$f''(s) = \frac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta (s-t)^{n-3}(t-\alpha) f^{(n)}(t)dt + \sum_{w=0}^{n-3} \frac{n-w-2}{w!} \cdot \frac{f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w}{\beta-\alpha}. \tag{3.251}$$

Using (3.251) in (3.244) and (3.245), we get

$$F_m(f) = \int_\alpha^\beta \left[F_m(G_z(\cdot,s)) \cdot \left(\tfrac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta (s-t)^{n-3}(t-\alpha)f^{(n)}(t)dt \right.\right.$$
$$\left.\left. + \sum_{w=0}^{n-3} \tfrac{n-w-2}{w!} \frac{f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w}{\beta-\alpha}\right)\right] ds$$

and

$$F_M(f) = \int_\alpha^\beta \left[F_M(G_z(\cdot,s)) \cdot \left(\tfrac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta (s-t)^{n-3}(t-\alpha)f^{(n)}(t)dt \right.\right.$$
$$\left.\left. + \sum_{w=0}^{n-3} \tfrac{n-w-2}{w!} \frac{f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w}{\beta-\alpha}\right)\right] ds$$

After interchanging the order of summation and integration and applying Fubini's theorem, we obtain (3.247) and (3.248).

ii) Similarly we can prove case $\alpha \leqslant s < t \leqslant \beta$. □

3.10. Sherman type inequalities for n-convex functions

We start this section with Sherman's type inequalities for n-convex functions ($n \geq 3$) without assumption of nonnegativity of weights $\mathbf{a}, \mathbf{b}$ and matrix $\mathbf{S}$ as given in Definition 3.9.6..

Theorem 3.10.78: *Let all the assumptions of Lemma 3.9.5. be satisfied.*

a) If in addition f is n-convex on $[\alpha, \beta]$ and

$$F_m(G_z(\cdot,s)) \geq 0, \tag{3.252}$$

then

$$F_m(f) \tag{3.253}$$
$$\geq \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w\right) ds.$$

If the reverse of (3.252) holds, then the reverse of (3.253) holds.

b) If in addition f is n-convex on $[\alpha,\beta]$ and

$$F_M(G_z(\cdot,s))\geq 0, \tag{3.254}$$

then

$$\begin{aligned} &F_M(f) \qquad (3.255)\\ &\geq \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w\right) ds. \end{aligned}$$

If the reverse of (3.254) holds, then the reverse of (3.255) holds.

Proof: a) Under the assumptions of theorem, the identities (3.247)-(3.250) hold.

Since $f^{(n-1)}$ is absolutely continuous on $[\alpha,\beta]$, then $f^{(n)}$ exists almost everywhere (see [145]). By assumption, f is n -convex on $[\alpha,\beta]$, therefore $f^{(n)}\geq 0$ on $[\alpha,\beta]$.

Now, using this fact and the assumption (3.252), from (3.247) and (3.249), we obtain (3.253).

b) Similarly we can prove this part. □

The next theorem gives Sherman's type inequalities for n-convex functions with an additional assumption of nonnegative entries of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{S}$ as in Sherman's result.

Theorem 3.10.79: *Let all the assumptions of Lemma 3.9.5. be satisfied. If in addition f is n-convex on $[\alpha,\beta]$ and $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$ are nonnegative, then (3.253) and (3.255) hold.*

Proof: Since additionally $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$ are nonnegative and $G_z(\cdot,s)$ is convex on $[\alpha,\beta]$, then by Remark 3.9.33. we have

$$F_m(G_z(\cdot,s))\geq 0 \quad \text{and} \quad F_M(G_z(\cdot,s))\geq 0,$$

i.e. the assumptions (3.252) and (3.254) are satisfied. Then by Theorem 3.10.78 the inequalities (3.253) and (3.255) hold. □

We use notation $\|\cdot\|_p$ for a standard p-norm and (p,q) denotes a pair of conjugate exponents, i.e. $1\leqslant p,q\leqslant\infty$, $1/p+1/q=1$.

Theorem 3.10.80: *Let all the assumptions of Lemma 3.9.5. be satisfied. Additionally, let $f^{(n)} \in L_p[\alpha,\beta]$.*

i) For each $t \in [\alpha,\beta]$, such that $t \leqslant s$, then

$$\left| F_m(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds \right|$$
$$\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\| f^{(n)} \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds \right)^q dt \right)^{\frac{1}{q}}$$

and

$$\left| F_M(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds \right|$$
$$\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\| f^{(n)} \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds \right)^q dt \right)^{\frac{1}{q}}$$

ii) For each $t \in [\alpha,\beta]$, such that $t > s$, then

$$\left| F_m(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds \right|$$
$$\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\| f^{(n)} \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\beta)ds \right)^q dt \right)^{\frac{1}{q}}$$

and

$$\left| F_M(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w \right) ds \right|$$
$$\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\| f^{(n)} \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\beta)ds \right)^q dt \right)^{\frac{1}{q}}.$$

Proof: (i) Under the assumptions of theorem, the identities (3.247) and (3.248) hold.

Applying Hölder's inequality to (3.247) and (3.248), we get

$$\begin{aligned}
&\left|F_m(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_m(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w\right) ds\right| \\
&= \left|\tfrac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta f^{(n)}(t) \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds\right) dt\right| \\
&\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\|f^{(n)}\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}}
\end{aligned}$$

and

$$\begin{aligned}
&\left|F_M(f) - \sum_{w=0}^{n-3} \tfrac{n-w-2}{(\beta-\alpha)w!} \int_\alpha^\beta F_M(G_z(\cdot,s)) \left(f^{(w+1)}(\beta)(s-\beta)^w - f^{(w+1)}(\alpha)(s-\alpha)^w\right) ds\right| \\
&= \left|\tfrac{1}{(n-3)!(\beta-\alpha)} \int_\alpha^\beta f^{(n)}(t) \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds\right) dt\right| \\
&\leqslant \tfrac{1}{(n-3)!(\beta-\alpha)} \left\|f^{(n)}\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(s-t)^{n-3}(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}},
\end{aligned}$$

what we need to prove.

(ii) Similarly we can prove this part. □

As direct consequence of the previous two theorems, choosing $n = 3$, we get the following corollary.

Corollary 3.10.20.: *Let the functionals F_m and F_M be defined as in Definition 3.9.6. with an additional assumption that* **a**, **b**, **S** *are nonnegative. Let $s \in [\alpha, \beta]$ and $G_z(\cdot, s), z \in \{1, 2, 3, 4, 5\}$, be defined as in (3.44)-(3.48), respectively. Let $f\colon [\alpha, \beta] \to \mathbb{R}$ be 3-convex function such that $f''\colon [\alpha, \beta] \to \mathbb{R}$ is absolutely continuous and $f''' \in L_p[\alpha, \beta]$.*

i) For each $t \in [\alpha, \beta]$, such that $t \leqslant s$, then

$$\begin{aligned}
0 &\leqslant F_m(f) - \frac{f'(\beta) - f'(\alpha)}{\beta - \alpha} \int_\alpha^\beta F_m(G_z(\cdot,s))ds \qquad (3.256)\\
&\leqslant \frac{1}{\beta-\alpha} \|f'''\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}}
\end{aligned}$$

and

$$
\begin{aligned}
0 \ &\leqslant \ F_M(f) - \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta F_M(G_z(\cdot,s))ds \qquad (3.257)\\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}};
\end{aligned}
$$

ii) For each $t \in [\alpha,\beta]$, such that $t > s$, then

$$
\begin{aligned}
0 &\leqslant F_m(f) - \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta F_m(G_z(\cdot,s))ds \qquad (3.258)\\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_m(G_z(\cdot,s))(t-\beta)ds\right)^q dt\right)^{\frac{1}{q}}
\end{aligned}
$$

and

$$
\begin{aligned}
0 &\leqslant F_M(f) - \frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta F_M(G_z(\cdot,s))ds \qquad (3.259)\\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta F_M(G_z(\cdot,s))(t-\beta)ds\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned}
$$

CHAPTER 4.

Applications in information theory

4.1. Shannon entropy and Csiszár f-divergence

Shannon ([148], [39, p. 476], [41]) introduced a statistical concept of entropy in the theory of communication and transmission of information, the measure of information that quantifies the unevenness in the probability distribution $\mathbf{p}$. It is defined as follows.

Definition 4.1.1.: Let $\mathbf{p} = (p_1, ..., p_n)$ be a positive probability distribution , i.e. $p_i > 0$, $i = 1, ..., n$, with $\sum_{i=1}^{n} p_i = 1$, for some discrete random variable X. **Shannon's entropy**of $\mathbf{p}$ is given by

$$H(\mathbf{p}) = \sum_{i=1}^{n} p_i \ln \frac{1}{p_i} = -\sum_{i=1}^{n} p_i \ln p_i. \tag{4.1}$$

With the usual agreement $0 \ln 0 = 0$, we can allow that for some i, $p_i = 0$, and (4.1) is still valid. This situation will not change anything because the value x_i, for which $p_i = 0$, can be freely removed from the set of values of X.

Shannon's entropy satisfies the following estimate

$$0 \leq H(\mathbf{p}) \leq \ln n. \tag{4.2}$$

As a slight modification of the previous formula, we get the Kullback-Leibler divergence, shorter KL-divergence, a measure of difference between two positive probability distributions $\mathbf{q}$ and $\mathbf{p}$. In statistics, it arises as the expected logarithm of difference between the probability $\mathbf{q}$ of data in the original distribution with the approximating distribution $\mathbf{p}$. We have the following definition (see [39, p. 477], [98]).

Definition 4.1.2.: Let $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ be two positive probability distributions, i.e. $p_i, q_i > 0$, $i = 1, ..., n$, with $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$. **The Kullback-Leibler divergence** or **relative entropy** of $\mathbf{q}$ with respect to $\mathbf{p}$ is given by

$$D_{KL}(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i}. \tag{4.3}$$

The Kullback-Leibler divergence satisfies the following estimates

$$0 \leq D_{KL}(\mathbf{q}, \mathbf{p}) \tag{4.4}$$

with equality iff $\mathbf{q} = \mathbf{p}$.

Shannon's entropy and the Kullback-Leibler divergence can be obtained as special cases of the more general concept known as Csiszár f-divergence (4.5) which is useful generalization of many well-known divergences widely employed in different scientific fields specially in mathematical statistics and information theory with deep connections in topics as diverse as artificial intelligence, statistical physics, and biological evolution (see [29, 40, 80, 81, 82, 83, 98, 99, 144, 148]).

Definition 4.1.3. *([33, 34])*: Let $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ be nonnegative n-tuples. For a convex function $f : [0, \infty) \to \mathbb{R}$, the **Csiszár f-divergence** or **f-divergence** is given by

$$D_f(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} p_i f\left(\frac{q_i}{p_i}\right), \tag{4.5}$$

with the convention

$$\begin{aligned} f(0) &= \lim_{t\to 0+} f(t), \quad 0f\left(\frac{0}{0}\right)=0 \\ 0f\left(\frac{c}{0}\right) &= \lim_{\varepsilon\to 0+} f\left(\frac{c}{\varepsilon}\right) = c\lim_{t\to\infty}\frac{f(t)}{t}, c>0. \end{aligned} \tag{4.6}$$

Remark 4.1.1.: a) Choosing the convex function $f(t)=\ln\frac{1}{t}=-\ln t$, we have

$$D_f(\mathbf{q},\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln\frac{q_i}{p_i} = \sum_{i=1}^{n} p_i \ln\frac{1}{q_i} - \sum_{i=1}^{n} p_i \ln\frac{1}{p_i}.$$

Specially, if $\mathbf{p}$ is positive probability distribution, then

$$D_f(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i \ln\frac{1}{q_i} - H(\mathbf{p}).$$

Further, setting $\mathbf{q}=\mathbf{e}=(1,...,1)$, we get

$$D_f(\mathbf{e},\mathbf{p}) = -H(\mathbf{p}).$$

b) Choosing the convex function $f(t)=t\ln t$, we get

$$D_f(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i \frac{q_i}{p_i}\ln\frac{q_i}{p_i} = \sum_{i=1}^{n} q_i \ln\frac{q_i}{p_i}.$$

If $\mathbf{p}$ and $\mathbf{q}$ are two positive probability distributions, then

$$D_f(\mathbf{q},\mathbf{p}) = D_{KL}(\mathbf{q},\mathbf{p}).$$

Csiszár with Körner [34] proved Jensen's type inequality for f -divergences, known as **the Csiszár-Körner inequality**, in the following form

$$\sum_{i=1}^{n} p_i f\left(\frac{\sum\limits_{i=1}^{n} q_i}{\sum\limits_{i=1}^{n} p_i}\right) \leq D_f(\mathbf{q},\mathbf{p}), \tag{4.7}$$

where $\mathbf{p}=(p_1,\ldots,p_n)$ and $\mathbf{q}=(q_1,\ldots,q_n)$ are nonnegative n-tuples with $\sum\limits_{i=1}^{n} p_i > 0$ and

$\sum_{i=1}^{n} q_i > 0$. Specially, if f is normalized, i.e. $f(1) = 0$ and $\mathbf{p}, \mathbf{q}$ are two positive probability distributions, then (4.7) reduces to

$$0 \le D_f(\mathbf{q}, \mathbf{p}) \tag{4.8}$$

with

$$D_f(\mathbf{q}, \mathbf{p}) = 0 \quad \text{iff} \quad \mathbf{p} = \mathbf{q}$$

provided f is strictly convex.

For suitable choices of the kernel f, Csiszár f-divergence (4.5) can be interpreted as a series of divergencies for two probability distributions $\mathbf{q}$ and $\mathbf{p}$:

- Hellinger divergence

$$D_{h^2}(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\sum_{i=1}^{n}(\sqrt{p_i} - \sqrt{q_i})^2, \quad f(t) = \frac{1}{2}\left(\sqrt{t} - 1\right)^2; \tag{4.9}$$

- Variational distance

$$D_V(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} |p_i - q_i|, \quad f(t) = |t - 1|;$$

- Harmonic divergence

$$D_h(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} \frac{2p_i q_i}{p_i + q_i}, \quad f(t) = \frac{2t}{1 + t}; \tag{4.10}$$

- Bhattacharya distance

$$D_B(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} \sqrt{p_i q_i}, \quad f(t) = \sqrt{t}; \tag{4.11}$$

- Triangular discrimination

$$D_\Delta(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i}, \quad f(t) = \frac{(t - 1)^2}{t + 1}; \tag{4.12}$$

- Chi square distance

$$D_{\chi^2}(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} \frac{(q_i - p_i)^2}{p_i}, \quad f(t) = (t-1)^2;$$

- Rényi divergence

$$D_{R_\lambda}(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i^{1-\lambda} q_i^{\lambda}, \quad f(t) = t^{\lambda}, \quad \lambda \in \mathbb{R}. \tag{4.13}$$

Motivated by various communication and transmission problems, Belis and Guiaşu [21] introduced the concept of weighted Shannon's entropy (see also [151]).

Definition 4.1.4.: Let $\mathbf{p} = (p_1, ..., p_n)$ be a positive probability distribution and $\mathbf{r} = (r_1, \ldots, r_n)$ nonnegative weight. **The weighted Shannon's entropy** is given by

$$H(\mathbf{p};\mathbf{r}) = \sum_{i=1}^{n} r_i p_i \ln \frac{1}{p_i}. \tag{4.14}$$

The weighted Shannon's entropy satisfied estimate [113]

$$0 \le H(\mathbf{p};\mathbf{r}) \le \sum_{i=1}^{n} r_i p_i \ln \frac{\sum_{i=1}^{n} r_i}{\sum_{i=1}^{n} r_i p_i}$$

In particular, the minimum $H(\mathbf{p};\mathbf{r}) = 0$ is reached for a constant random variable, i.e. when $p_i = 1$, for some i. The opposite extreme, the maximal $H(\mathbf{p};\mathbf{r})$ is reached for a uniform distribution, i.e. when $p_i = \frac{1}{n}$ for all $i = 1, ..., n$. In that case we have

$$0 \le H(\mathbf{p};\mathbf{r}) \le \frac{1}{n} \sum_{i=1}^{n} r_i \ln n.$$

Specially, if we ignore weights r_i, $i = 1, ..., n$, i.e. setting $\mathbf{r} = \mathbf{e} = (1, ..., 1)$, then (4.14) reduces to (4.1), i.e. $H(\mathbf{p};\mathbf{e}) = H(\mathbf{p})$ for $\mathbf{e} = (1, ..., 1) \in \mathbb{R}^n$ and the previous inequality reduces to (4.2).

The following definition of the weighted Csiszár f-divergence is given in [90].

Definition 4.1.5.: Let $\mathbf{p} = (p_1, \ldots, p_n)$, $\mathbf{q} = (q_1, \ldots, q_n)$ be nonnegative n-tuples and

$\mathbf{r} = (r_1, \ldots, r_n)$ be a nonnegative weight. For a convex function $f : [0, \infty) \to \mathbb{R}$, **the weighted Csiszár f-divergence** is given by

$$D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) = \sum_{i=1}^{n} r_i p_i f\left(\frac{q_i}{p_i}\right). \tag{4.15}$$

Remark 4.1.2.: It is obvious that $D_f(\mathbf{q}, \mathbf{p}; \mathbf{e}) = D_f(\mathbf{q}, \mathbf{p})$ for $\mathbf{e} = (1, ..., 1) \in \mathbb{R}^n$.

a) Further, for the convex function $f(t) = \ln \frac{1}{t} = -\ln t$, we have

$$D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) = -\sum_{i=1}^{n} r_i p_i \ln \frac{q_i}{p_i} = \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i} - \sum_{i=1}^{n} r_i p_i \ln \frac{1}{p_i}.$$

If $\mathbf{p}$ is positive probability distribution, then

$$D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) = \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i} - H(\mathbf{p}; \mathbf{r}).$$

Choosing $\mathbf{q} = \mathbf{e} = (1, ..., 1)$, we get

$$D_f(\mathbf{e}, \mathbf{p}; \mathbf{r}) = -H(\mathbf{p}; \mathbf{r}).$$

b) For the convex function $f(t) = t \ln t$, $t > 0$, we have

$$D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) = \sum_{i=1}^{n} r_i p_i \frac{q_i}{p_i} \ln\left(\frac{q_i}{p_i}\right) = \sum_{i=1}^{n} r_i q_i \ln\left(\frac{q_i}{p_i}\right).$$

If $\mathbf{p}$ and $\mathbf{q}$ are two positive probability distributions, then

$$D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) = D_{KL}(\mathbf{q}, \mathbf{p}; \mathbf{r}),$$

i.e. we get **the weighted Kullback-Leibler divergence**.

Below we give some weighted versions of distances for two probability distributions $\mathbf{p}, \mathbf{q}$ with weights $\mathbf{r} = (r_1, ..., r_n)$:

- The weighted Hellinger distance

$$D_{h^2}(\mathbf{q},\mathbf{p};\mathbf{r}) = \frac{1}{2}\sum_{i=1}^{n} r_i(\sqrt{p_i}-\sqrt{q_i})^2;$$

- The weighted Bhattacharya distance

$$D_B(\mathbf{q},\mathbf{p};\mathbf{r}) = \sum_{i=1}^{n} r_i\sqrt{p_i q_i}; \tag{4.16}$$

- The weighted Chi square distance

$$D_{\chi^2}(\mathbf{q},\mathbf{p};\mathbf{r}) = \sum_{i=1}^{n} r_i\frac{(p_i-q_i)^2}{p_i};$$

- The weighted variational distance

$$D_V(\mathbf{q},\mathbf{p};\mathbf{r}) = \sum_{i=1}^{n} r_i|p_i-q_i|;$$

- The weighted triangular discrimination

$$D_\Delta(\mathbf{q},\mathbf{p};\mathbf{r}) = \sum_{i=1}^{n} r_i\frac{(p_i-q_i)^2}{p_i+q_i}.$$

In this section, using results from the previous chapters, we give new estimates for f-divergence functional $D_f(\mathbf{q},\mathbf{p})$ as well as for the generalized f-divergence functional $D_f(\mathbf{p},\mathbf{q};\mathbf{r})$. As special cases of those estimates, we obtain bounds for some well known divergences. The results from this chapter are cited from [17, 71, 72].

4.1.1. Converse to the Csiszár-Körner inequality

In this subsection, applying Converse Sherman's inequality (2.11) to the concept of f-divergences, we obtain a converse to the Csiszár-Körner inequality (4.7). As easy consequences we derive some inequalities for the well-known divergences (see [17]).

In the sequel, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ will denote the sets of nonnegative and positive numbers, i.e. $\mathbb{R}_+ = [0,\infty)$ and $\mathbb{R}_{++} = (0,\infty)$, respectively. We denote with $\mathcal{M}_{nm}(\mathbb{R}_+)$ the space of $n \times m$ matrices with nonnegative entries. We use notation $\langle \cdot, \cdot \rangle$ for the standard inner product in $\mathbb{R}^n$.

We also use the following notation and assumptions:

(A) Let $\alpha, \beta \in \mathbb{R}_{++}$, $\alpha < \beta$. Let $\mathbf{p} \in \mathbb{R}^n_{++}$, $\mathbf{q} \in \mathbb{R}^n_{++}$, be such that $\frac{q_i}{p_i} \in [\alpha, \beta]$, $i = 1, ...n$. Let $\mathbf{R} = (r_{ij}) \in \mathcal{M}_{nm}(\mathbb{R}_+)$ and $\mathbf{d} \in \mathbb{R}^m_+$. We define m-tuples $\tilde{\mathbf{p}}, \tilde{\mathbf{q}} \in \mathbb{R}^m_{++}$ and n-tuple $\mathbf{c} \in \mathbb{R}^n_+$ such that

$$\tilde{\mathbf{p}} = \mathbf{p}\mathbf{R}, \quad \tilde{\mathbf{q}} = \mathbf{q}\mathbf{R} \quad \text{and} \quad \mathbf{c} = \mathbf{d}\mathbf{R}^T. \tag{4.17}$$

We denote jth column of $\mathbf{R}$ with $\mathbf{r}_j = (r_{1j}, ..., r_{nj})$, $j = 1, ..., m$.

From (4.17) it follows

$$\tilde{p}_j = \sum_{i=1}^{n} p_i r_{ij} = \langle \mathbf{p}, \mathbf{r}_j \rangle, \; j = 1, ..., m, \tag{4.18}$$

$$\tilde{q}_j = \sum_{i=1}^{n} q_i r_{ij} = \langle \mathbf{q}, \mathbf{r}_j \rangle, \; j = 1, ..., m, \tag{4.19}$$

$$c_i = \sum_{j=1}^{m} d_j r_{ij}, \; i = 1, ..., n.$$

For m-tuples $\tilde{\mathbf{p}}$, $\tilde{\mathbf{q}}$ and $\mathbf{d}$, Csiszár f-divergence (4.15) becomes

$$D_f(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) = \sum_{j=1}^{m} d_j \tilde{p}_j f\left(\frac{\tilde{q}_j}{\tilde{p}_j}\right).$$

Specially, if $\mathbf{p}$ and $\mathbf{q}$ are two probability distributions and a matrix $\mathbf{R} \in \mathcal{M}_{nm}((0, \infty))$ is row stochastic, then

$$\sum_{j=1}^{m} \tilde{p}_j = \sum_{j=1}^{m} \sum_{i=1}^{n} p_i r_{ij} = \sum_{i=1}^{n} p_i \sum_{j=1}^{m} r_{ij} = 1, \tag{4.20}$$

$$\sum_{j=1}^{m} \tilde{q}_j = \sum_{j=1}^{m} \sum_{i=1}^{n} q_i r_{ij} = \sum_{i=1}^{n} q_i \sum_{j=1}^{m} r_{ij} = 1,$$

i.e. $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{q}}$ are also probability distributions.

Applying Theorem 2.2.5 we compare two generalized Csiszár f-divergences.

Theorem 4.1.1: *Let (A) holds. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function. Then*

$$\begin{aligned} D_f(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) &\leq D_f(\mathbf{q}, \mathbf{p}; \mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{f(\alpha)\left(\beta - \frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) + f(\beta)\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle} - \alpha\right)}{\beta - \alpha}. \end{aligned} \tag{4.21}$$

Proof: According to (4.5), we need to prove the inequality

$$\begin{aligned} \sum_{j=1}^{m} d_j \tilde{p}_j f\left(\frac{\tilde{q}_j}{\tilde{p}_j}\right) &\leq \sum_{i=1}^{n} c_i p_i f\left(\frac{q_i}{p_i}\right) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{f(\alpha)\left(\beta - \frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) + f(\beta)\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle} - \alpha\right)}{\beta - \alpha}. \end{aligned} \tag{4.22}$$

Let vectors $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_m)$ be such that $x_i = \frac{q_i}{p_i}, i = 1, ..., n$, and $y_j = \frac{\tilde{q}_j}{\tilde{p}_j} = \frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}, j = 1, ..., m.$

The following equality holds

$$\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} = \frac{p_1 r_{1j}}{\sum_{i=1}^{n} p_i rij} \frac{q_1}{p_1} + ... + \frac{p_n r_{nj}}{\sum_{i=1}^{n} p_i rij} \frac{q_n}{p_n}, \quad j = 1, ..., m.$$

Hence, the following identity is valid

$$\left[\frac{\langle \mathbf{q}, \mathbf{r}_1 \rangle}{\langle \mathbf{p}, \mathbf{r}_1 \rangle}, ..., \frac{\langle \mathbf{q}, \mathbf{r}_m \rangle}{\langle \mathbf{p}, \mathbf{r}_m \rangle}\right] = \left[\frac{q_1}{p_1}, ..., \frac{q_n}{p_n}\right] \begin{bmatrix} \frac{p_1 r_{11}}{\langle \mathbf{p},\mathbf{r}_1\rangle} & \cdots & \frac{p_1 r_{1m}}{\langle \mathbf{p},\mathbf{r}_m\rangle} \\ \vdots & \ddots & \vdots \\ \frac{p_n r_{n1}}{\langle \mathbf{p},\mathbf{r}_1\rangle} & \cdots & \frac{p_n r_{nm}}{\langle \mathbf{p},\mathbf{r}_m\rangle} \end{bmatrix}. \tag{4.23}$$

Define $n \times m$ column stochastic matrix $\mathbf{S} = (s_{ij})$ with $s_{ij} = \frac{p_i r_{ij}}{\tilde{p}_j} = \frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle}$, $i = 1, ..., n, j = 1, ..., m$. Then, condition $\mathbf{y} = \mathbf{x}\mathbf{S}$ is satisfied (see (4.23)).

Consider the weight $b_j = d_j \tilde{p}_j = d_j \langle \mathbf{p}, \mathbf{r}_j \rangle$, $j = 1, .., m$, and $\mathbf{a} = (a_1, ..., a_n)$ such that $a_i = p_i c_i$, $i = 1, ..., n$.

We have

$$a_i = p_i c_i = p_i \sum_{j=1}^{m} d_j r_{ij} = \sum_{j=1}^{m} d_j p_i r_{ij} = \sum_{j=1}^{m} b_j \frac{p_i r_{ij}}{\tilde{p}_j},$$

i.e. $\mathbf{a} = \mathbf{b}\mathbf{S}^T$ is satisfied. Moreover, since (2.6) holds, applying Theorem 2.2.5 we obtain

$$\begin{aligned}
&\sum_{j=1}^{m} b_j f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle}\right) \\
&= \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j\rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle}\right) \leq \sum_{i=1}^{n} c_i p_i f\left(\frac{q_i}{p_i}\right) \\
&\leq \frac{\sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j\rangle \left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle}\right)}{\beta - \alpha} f(\alpha) + \frac{\sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j\rangle \left(\frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle} - \alpha\right)}{\beta - \alpha} f(\beta),
\end{aligned}$$

what is equivalent to (4.21). □

Corollary 4.1.1.: *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function. Let $\mathbf{R} \in \mathcal{M}_{nm}(\mathbb{R}_+)$, $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n_{++}$ and $\tilde{\mathbf{p}}, \tilde{\mathbf{q}} \in \mathbb{R}^m_{++}$ be as in (A). Further, let $\mathbf{r} = (r_1, ..., r_n)$ be n-tuple such that $r_i = \sum_{j=1}^m r_{ij}$, $i = 1, ..., n$, i.e. r_i is the ith row sum of $\mathbf{R}$. Then*

$$D_f(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}) \leq D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) \leq \sum_{j=1}^{m} \langle \mathbf{p}, \mathbf{r}_j\rangle \frac{f(\alpha)\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle}\right) + f(\beta)\left(\frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle} - \alpha\right)}{\beta - \alpha}. \tag{4.24}$$

In particular, if $\mathbf{R}$ is row stochastic matrix, then

$$D_f(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}) \leq D_f(\mathbf{q}, \mathbf{p}) \leq \sum_{j=1}^{m} \langle \mathbf{p}, \mathbf{r}_j\rangle \frac{f(\alpha)\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle}\right) + f(\beta)\left(\frac{\langle \mathbf{q}, \mathbf{r}_j\rangle}{\langle \mathbf{p}, \mathbf{r}_j\rangle} - \alpha\right)}{\beta - \alpha}. \tag{4.25}$$

Proof: By taking $\mathbf{d} = (d_1, ..., d_m) = (1, ..., 1)$ in Theorem 4.1.1, we calculate $c_i = \sum_{j=1}^m r_{ij} = r_i$ for $i = 1, ..., n$. Therefore inequality (4.21) becomes (4.24).

If additionally $\mathbf{R}$ is row stochastic, then $\mathbf{r} = (1, ..., 1) \in \mathbb{R}^n$ and (4.24) reduces to (4.25).

□

As a special case of the previous result, choosing $m = 1$, we obtain the following converse to the Csiszár-Körner inequality (4.7).

Corollary 4.1.2.: *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function. Let $\alpha, \beta \in \mathbb{R}_{++}$, $\alpha < \beta$, and $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n_{++}$, be such that $\frac{q_i}{p_i} \in [\alpha, \beta]$, $i = 1, ..., n$, and $\mathbf{r} \in \mathbb{R}^n_+$. Then*

$$\langle \mathbf{p}, \mathbf{r}\rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}\rangle}{\langle \mathbf{p}, \mathbf{r}\rangle}\right) \leq D_f(\mathbf{q}, \mathbf{p}; \mathbf{r}) \leq \langle \mathbf{p}, \mathbf{r}\rangle \frac{f(\alpha)\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}\rangle}{\langle \mathbf{p}, \mathbf{r}\rangle}\right) + f(\beta)\left(\frac{\langle \mathbf{q}, \mathbf{r}\rangle}{\langle \mathbf{p}, \mathbf{r}\rangle} - \alpha\right)}{\beta - \alpha}. \tag{4.26}$$

In particular, if $\mathbf{r} = \mathbf{e} = (1, ..., 1)$ *and* $P_n = \sum_{i=1}^{n} p_i$ *and* $Q_n = \sum_{i=1}^{n} q_i$*, then*

$$P_n f\left(\frac{Q_n}{P_n}\right) \leq D_f(\mathbf{q}, \mathbf{p}) \leq P_n \frac{f(\alpha)\left(\beta - \frac{Q_n}{P_n}\right) + f(\beta)\left(\frac{Q_n}{P_n} - \alpha\right)}{\beta - \alpha}. \tag{4.27}$$

Remark 4.1.3.: We get the improvement of the Csiszár-Körner inequality (4.7). Specially, if f is normalized function, i.e $f(1) = 0$, and $\mathbf{p}, \mathbf{q}$ are two probability distributions, then (4.27) becomes

$$0 \leq D_f(\mathbf{q}, \mathbf{p}) \leq \frac{f(\alpha)(\beta - 1) + f(\beta)(1 - \alpha)}{\beta - \alpha}.$$

4.1.2. Converse inequalities related to some divergences and Shannon's entropy

In the following results we use the following notation and assumptions:

(B) Let $\alpha, \beta \in \mathbb{R}_{++}$, $\alpha < \beta$. Let $\mathbf{p} \in \mathbb{R}^n_{++}$, $\mathbf{q} \in \mathbb{R}^n_{++}$, be two probability distributions such that $\frac{q_i}{p_i} \in [\alpha, \beta]$, $i = 1, ...n$. Let $\mathbf{R} = (r_{ij}) \in \mathcal{M}_{nm}(\mathbb{R}_{++})$ be row stochastic matrix and $\mathbf{d} \in \mathbb{R}^m_+$. We define m-tuples $\tilde{\mathbf{p}}, \tilde{\mathbf{q}} \in \mathbb{R}^m_{++}$ and n-tuple $\mathbf{c} \in \mathbb{R}^n_+$ such that

$$\tilde{\mathbf{p}} = \mathbf{pR}, \quad \tilde{\mathbf{q}} = \mathbf{qR} \quad \text{and} \quad \mathbf{c} = \mathbf{dR}^T, \tag{4.28}$$

i.e. for $\mathbf{r}_j = (r_{1j}, ..., r_{nj})$, $j$the column of $\mathbf{R}$, $j = 1, ..., m$, we have (4.18)-(4.19).
Since $\mathbf{p}$ and $\mathbf{q}$ are two probability distributions and $\mathbf{R}$ is row stochastic, then $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{q}}$ are also probability distributions (see (4.20)).

Using results from the previous subsection we get converse inequalities related to well known divergences as speciall cases of Csiszár f-divergences (4.5) and (4.15).

Corollary 4.1.3.: *Let* (B) *holds. Then:*

a) The weighted Kullback-Leibler divergences satisfy

$$\begin{aligned} D_{KL}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) &\leq D_{KL}(\mathbf{q}, \mathbf{p}; \mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{\ln \frac{1}{\alpha}\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + \ln \frac{1}{\beta}\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} - \alpha\right)}{\beta - \alpha}; \end{aligned} \tag{4.29}$$

b) The weighted Hellinger distances satisfy

$$\begin{aligned} &D_{h^2}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) \leq D_{h^2}(\mathbf{q}, \mathbf{p}; \mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{(1-\sqrt{\alpha})^2\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + (1-\sqrt{\beta})^2\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} - \alpha\right)}{2(\beta - \alpha)}; \end{aligned} \tag{4.30}$$

c) The weighted Bhattacharya distances satisfy

$$\begin{aligned} D_{B}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) &\geq D_{B}(\mathbf{q}, \mathbf{p}; \mathbf{c}) \\ &\geq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{\sqrt{\alpha}\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + \sqrt{\beta}\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} - \alpha\right)}{\beta - \alpha}; \end{aligned} \tag{4.31}$$

d) The weighted Chi square distances satisfy

$$\begin{aligned} &D_{\chi^2}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}; \mathbf{d}) \leq D_{\chi^2}(\mathbf{q}, \mathbf{p}; \mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{(1-\alpha)^2\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + (1-\beta)^2\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} - \alpha\right)}{\beta - \alpha}; \end{aligned} \tag{4.32}$$

e) The weighted variational distances satisfy

$$\begin{aligned} &D_{V}(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}; \mathbf{d}) \leq D_{V}(\mathbf{p}, \mathbf{q}; \mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p}, \mathbf{r}_j \rangle \frac{|1-\alpha|\left(\beta - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + |1-\beta|\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} - \alpha\right)}{\beta - \alpha}; \end{aligned} \tag{4.33}$$

f) The weighted triangular discriminations satisfy

$$\begin{aligned} D_{\Delta}(\tilde{\mathbf{q}},\tilde{\mathbf{p}};\mathbf{d}) &\leq D_{\Delta}(\mathbf{q},\mathbf{p};\mathbf{c}) \\ &\leq \sum_{j=1}^{m} d_j \langle \mathbf{p},\mathbf{r}_j\rangle \frac{\frac{(1-\alpha)^2}{\alpha+1}\left(\beta-\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right)+\frac{(1-\beta)^2}{\beta+1}\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}. \end{aligned} \tag{4.34}$$

Proof: We apply Theorem 4.1.1 to the different convex functions as follows:

a) for $f(t)=\ln\frac{1}{t}$, the inequality (4.29) follows from (4.21);

b) for $f(t)=\frac{1}{2}\left(1-\sqrt{t}\right)^2$, the inequality (4.30) follows from (4.21);

c) for $f(t)=-\sqrt{t}$, the inequality (4.31) follows from (4.21);

d) for $f(t)=(t-1)^2$, the inequality (4.32) follows from (4.21);

e) for $f(t)=|1-t|$, the inequality (4.34) follows from (4.21). □

By taking $\mathbf{d}=(d_1,...,d_m)=(1,...,1)$ and the fact that then $c_i=\sum_{j=1}^{m} r_{ij}=1$, $i=1,...,n$, as easy consequences of the previous corollary we get the next result.

Corollary 4.1.4.: *Let $\mathbf{R}\in\mathcal{M}_{nm}(\mathbb{R}_{++})$, $\mathbf{p},\mathbf{q}\in\mathbb{R}^n_{++}$ and $\tilde{\mathbf{p}},\tilde{\mathbf{q}}\in\mathbb{R}^m_{++}$ be as in (B). Then*

a) The Kullback-Leibler divergences satisfy

$$\begin{aligned} D_{KL}(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\leq D_{KL}(\mathbf{q},\mathbf{p}) \\ &\leq \sum_{j=1}^{m} \langle \mathbf{p},\mathbf{r}_j\rangle \frac{\ln\frac{1}{\alpha}\left(\beta-\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right)+\ln\frac{1}{\beta}\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}; \end{aligned} \tag{4.35}$$

b) Hellinger distances satisfy

$$\begin{aligned} D_{h^2}(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\leq D_{h^2}(\mathbf{q},\mathbf{p}) \\ &\leq \sum_{j=1}^{m} \langle \mathbf{p},\mathbf{r}_j\rangle \frac{(1-\sqrt{\alpha})^2\left(\beta-\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right)+(1-\sqrt{\beta})^2\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{2(\beta-\alpha)}; \end{aligned} \tag{4.36}$$

c) Bhattacharya distances satisfy

$$\begin{aligned} D_{B}(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\geq D_{B}(\mathbf{q},\mathbf{p}) \\ &\geq \sum_{j=1}^{m} \langle \mathbf{p},\mathbf{r}_j\rangle \frac{\sqrt{\alpha}\left(\beta-\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right)+\sqrt{\beta}\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}; \end{aligned} \tag{4.37}$$

d) Chi square distances satisfy

$$\begin{aligned} D_{\chi^2}(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\leq D_{\chi^2}(\mathbf{q},\mathbf{p}) \\ &\leq \sum_{j=1}^{m}\langle\mathbf{p},\mathbf{r}_j\rangle\frac{(1-\alpha)^2\left(\beta-\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)+(1-\beta)^2\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}; \end{aligned} \tag{4.38}$$

e) Variational distances satisfy

$$\begin{aligned} D_V(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\leq D_V(\mathbf{q},\mathbf{p}) \\ &\leq \sum_{j=1}^{m}\langle\mathbf{p},\mathbf{r}_j\rangle\frac{|1-\alpha|\left(\beta-\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)+|1-\beta|\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}; \end{aligned} \tag{4.39}$$

f) Triangular discriminations satisfy

$$\begin{aligned} D_\Delta(\tilde{\mathbf{q}},\tilde{\mathbf{p}}) &\leq D_\Delta(\mathbf{q},\mathbf{p}) \\ &\leq \sum_{j=1}^{m}\langle\mathbf{p},\mathbf{r}_j\rangle\frac{\frac{(1-\alpha)^2}{\alpha+1}\left(\beta-\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)+\frac{(1-\beta)^2}{\beta+1}\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}-\alpha\right)}{\beta-\alpha}. \end{aligned} \tag{4.40}$$

Next we present converse inequalities related to Shannon's entropies. We use the following notation and assumptions:

(C) Let $\alpha,\beta\in\mathbb{R}_{++}$, $\alpha<\beta$. Let $\mathbf{p}\in\mathbb{R}^n_{++}$ be probability distribution such that $p_i\in[\alpha,\beta]$, $i=1,...n$. Let $\mathbf{R}=(r_{ij})\in\mathcal{M}_{nm}(\mathbb{R}_{++})$ be doubly stochastic matrix and $\mathbf{d}\in\mathbb{R}^m_+$. We define m-tuple $\tilde{\mathbf{p}}\in\mathbb{R}^m_{++}$ and n-tuple $\mathbf{c}\in\mathbb{R}^n_+$ such that

$$\tilde{\mathbf{p}}=\mathbf{pR} \quad \text{and} \quad \mathbf{c}=\mathbf{dR}^T. \tag{4.41}$$

For $\mathbf{r}_j=(r_{1j},...,r_{nj})$, $j$the column of $\mathbf{R}$, $j=1,...,m$, we have

$$\tilde{p}_j=\sum_{i=1}^{n}p_ir_{ij}=\langle\mathbf{p},\mathbf{r}_j\rangle,\ j=1,...,m.$$

Since $\mathbf{p}$ is probability distribution and $\mathbf{R}$ is doubly stochastic, then

$$\sum_{j=1}^{m}\tilde{p}_j=\sum_{j=1}^{m}\sum_{i=1}^{n}p_ir_{ij}=\sum_{i=1}^{n}p_i\sum_{j=1}^{m}r_{ij}=1,$$

i.e. $\tilde{\mathbf{p}}$ is also probability distribution.

Corollary 4.1.5.: *Let (C) holds. Then*

$$H(\tilde{\mathbf{p}};\mathbf{d}) \geq H(\mathbf{p};\mathbf{c}) \geq \sum_{j=1}^{m} d_j \langle \mathbf{p},\mathbf{r}_j\rangle \frac{\ln\alpha\left(\beta - \frac{1}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) + \ln\beta\left(\frac{1}{\langle \mathbf{p},\mathbf{r}_j\rangle} - \alpha\right)}{\beta-\alpha}. \tag{4.42}$$

Proof: Setting $\mathbf{q} = \mathbf{e} = (1,...,1) \in \mathbb{R}^m$ and using the assumption that $\mathbf{R}$ is doubly stochastic, we have $\tilde{\mathbf{q}} = (\tilde{q}_1,...,\tilde{q}_m)$ with $\tilde{q}_j = \sum_{i=1}^{n} q_i r_{ij} = \sum_{i=1}^{n} r_{ij} = 1$, i.e. $\tilde{\mathbf{q}} = (1,...,1)$. Now applying Theorem 4.1.1 to the convex function $f(t) = \ln\frac{1}{t}$ the inequality (4.42) follows from (4.21). □

Corollary 4.1.6.: *Let $\mathbf{R} \in \mathcal{M}_{nm}(\mathbb{R}_{++})$, $\mathbf{p} \in \mathbb{R}^n_{++}$ and $\tilde{\mathbf{p}} \in \mathbb{R}^m_{++}$ be as in (C). Then*

$$H(\tilde{\mathbf{p}}) \geq H(\mathbf{p}) \geq \sum_{j=1}^{m} \langle \mathbf{p},\mathbf{r}_j\rangle \frac{\ln\alpha\left(\beta - \frac{1}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) + \ln\beta\left(\frac{1}{\langle \mathbf{p},\mathbf{r}_j\rangle} - \alpha\right)}{\beta-\alpha}. \tag{4.43}$$

Proof: Inequality (4.43) follows directly from (4.42) by taking $\mathbf{d} = (d_1,...,d_m) = (1,...,1)$ and the fact that then $c_i = \sum_{j=1}^{m} r_{ij} = 1$, for $i = 1,...,n$. □

4.1.3. New bounds for f-divergences

In this section, we present new lower and upper bounds for Csiszár f-divergences and Shannon's entropy using estimates for Sherman's functional (2.23) (see [72]).

As an easy consequence of Theorem 2.3.9 we get the next result.

Corollary 4.1.7.: *Let $f : [0,\infty) \to \mathbb{R}$ be a convex function, $\mathbf{p} = (p_1,...,p_n) \in [0,\infty)^n$, $\mathbf{q} = (q_1,...,q_n) \in [0,\infty)^n$ and $\mathbf{R} = (r_{ij}) \in \mathcal{M}_{nm}(\mathbb{R}_+)$. Let us define*

$$\begin{aligned} \langle \mathbf{p},\mathbf{r}_j\rangle &= \sum_{i=1}^{n} p_i r_{ij}, \quad \langle \mathbf{q},\mathbf{r}_j\rangle = \sum_{i=1}^{n} q_i r_{ij}, \quad j = 1,...,m, \qquad (4.44)\\ R_f(\mathbf{q},\mathbf{p};\mathbf{R}) &= \sum_{j=1}^{m}\left[\sum_{i=1}^{n} q_i r_{ij} f\left(\frac{q_i}{p_i}\right) - \langle \mathbf{q},\mathbf{r}_j\rangle f\left(\frac{1}{\langle \mathbf{q},\mathbf{r}_j\rangle}\sum_{i=1}^{n}\frac{q_i^2 r_{ij}}{p_i}\right)\right]. \end{aligned}$$

Then

$$
\begin{aligned}
0 &\leq \min_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\} R_f(\mathbf{q},\mathbf{p};\mathbf{R}) \\
&\leq \sum_{j=1}^{m}\left[\sum_{i=1}^{n} p_i r_{ij} f\left(\frac{q_i}{p_i}\right) - \langle \mathbf{p},\mathbf{r}_j\rangle f\left(\frac{1}{\langle \mathbf{p},\mathbf{r}_j\rangle}\sum_{i=1}^{n} q_i r_{ij}\right)\right] \\
&\leq \max_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\} R_f(\ \mathbf{q},\mathbf{p};\mathbf{R})
\end{aligned}
\tag{4.45}
$$

Proof: Applying (2.33) to the vector $\mathbf{x} = (x_1, ..., x_n)$ with $x_i = \frac{q_i}{p_i}$, $i = 1, ..., n$, weight $\mathbf{b} = (b_1, ..., b_m)$ with $b_j = \langle \mathbf{p}, \mathbf{r}_j\rangle$, $j = 1, ..., m$, column stochastic matrices $\mathbf{S} = (s_{ij}) \in \mathcal{C}_{nm}(\mathbb{R}_+)$ with $s_{ij} = \frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle}$, $i = 1, ..., n$, $j = 1, ..., m$ and $\mathbf{T} = (t_{ij}) \in \mathcal{C}_{nm}(\mathbb{R}_+)$ with $t_{ij} = \frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle}$, $i = 1, ..., n$, $j = 1, ..., m$, we get

$$
\begin{aligned}
0 &\leq \sum_{j=1}^{m}\langle \mathbf{p},\mathbf{r}_j\rangle \min_{1\leq i\leq n}\left\{\frac{\frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle}}{\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle}}\right\}\left(\sum_{i=1}^{n}\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle} f\left(\frac{q_i}{p_i}\right) - f\left(\sum_{i=1}^{n}\frac{q_i}{p_i}\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle}\right)\right) \\
&\leq \sum_{j=1}^{m}\langle \mathbf{p},\mathbf{r}_j\rangle \left(\sum_{i=1}^{n}\frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle} f\left(\frac{q_i}{p_i}\right) - f\left(\sum_{i=1}^{n}\frac{q_i}{p_i}\frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right)\right) \\
&\leq \sum_{j=1}^{m}\langle \mathbf{p},\mathbf{r}_j\rangle \max_{1\leq i\leq n}\left\{\frac{\frac{p_i r_{ij}}{\langle \mathbf{p},\mathbf{r}_j\rangle}}{\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle}}\right\}\left(\sum_{i=1}^{n}\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle} f\left(\frac{q_i}{p_i}\right) - f\left(\sum_{i=1}^{n}\frac{q_i}{p_i}\frac{q_i r_{ij}}{\langle \mathbf{q},\mathbf{r}_j\rangle}\right)\right).
\end{aligned}
$$

what is equivalent to (4.45). □

Specially, for $m = 1$, the previous result reduces to the next corollary.

Corollary 4.1.8.: *Let $f : [0,\infty) \to \mathbb{R}$ be a convex function, $\mathbf{p} = (p_1, ..., p_n) \in [0,\infty)^n$, $\mathbf{q} = (q_1, ..., q_n) \in [0,\infty)^n$ and $\mathbf{r} = (r_1, ..., r_n) \in [0,\infty)^n$. Let us define*

$$
\begin{aligned}
\langle \mathbf{p},\mathbf{r}\rangle &= \sum_{i=1}^{n} p_i r_i, \quad \langle \mathbf{q},\mathbf{r}\rangle = \sum_{i=1}^{n} q_i r_i, \\
R_f(\mathbf{q},\mathbf{p};\mathbf{r}) &= \sum_{i=1}^{n} q_i r_i f\left(\frac{q_i}{p_i}\right) - \langle \mathbf{q},\mathbf{r}\rangle f\left(\frac{1}{\langle \mathbf{q},\mathbf{r}\rangle}\sum_{i=1}^{n}\frac{q_i^2 r_i}{p_i}\right).
\end{aligned}
$$

Then

$$0 \leq \min_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\} R_f(\mathbf{q},\mathbf{p};\mathbf{r}) \leq D_f(\mathbf{q},\mathbf{p};\mathbf{r}) - \langle \mathbf{p},\mathbf{r}\rangle f\left(\frac{1}{\langle \mathbf{p},\mathbf{r}\rangle}\sum_{i=1}^{n} q_i r_i\right)$$
$$\leq \max_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\} R_f(\mathbf{q},\mathbf{p};\mathbf{r}). \tag{4.46}$$

If in addition $\mathbf{r}=\mathbf{e}=(1,...,1)$, *with notations* $\langle \mathbf{p},\mathbf{r}\rangle = \sum_{i=1}^{n} p_i = P_n$, $\langle \mathbf{q},\mathbf{r}\rangle = \sum_{i=1}^{n} q_i = Q_n$, *then*

$$0 \leq \min_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n} q_i f\left(\frac{q_i}{p_i}\right) - Q_n f\left(\frac{1}{Q_n}\sum_{i=1}^{n}\frac{q_i^2}{p_i}\right)\right) \tag{4.47}$$
$$\leq D_f(\mathbf{q},\mathbf{p}) - P_n f\left(\frac{Q_n}{P_n}\right)$$
$$\leq \max_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n} q_i f\left(\frac{q_i}{p_i}\right) - Q_n f\left(\frac{1}{Q_n}\sum_{i=1}^{n}\frac{q_i^2}{p_i}\right)\right).$$

Remark 4.1.4.: Note that (4.47) generalizes and refines the Csiszár-Körner inequality (4.7). Specially, applying (4.47) to normalized function f, i.e. $f(1)=0$, with assumption that $\mathbf{q}$ and $\mathbf{p}$ are two probability distributions, we get the lower and upper bound of $D_f(\mathbf{q},\mathbf{p})$ in the form

$$0 \leq \min_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n} q_i f\left(\frac{q_i}{p_i}\right) - f\left(\sum_{i=1}^{n}\frac{q_i^2}{p_i}\right)\right) \tag{4.48}$$
$$\leq D_f(\mathbf{q},\mathbf{p})$$
$$\leq \max_{1\leq i\leq n}\left\{\frac{p_i}{q_i}\right\}\left(\sum_{i=1}^{n} q_i f\left(\frac{q_i}{p_i}\right) - f\left(\sum_{i=1}^{n}\frac{q_i^2}{p_i}\right)\right).$$

In the following result, using conclusions from Remark 4.1.1. a) and Remark 4.1.2. a), we obtain new estimate for Shannon's entropies (4.1) and (4.14).

Corollary 4.1.9.: *Let* $\mathbf{p} = (p_1, ..., p_n)$ *be a probability distribution,* $\mathbf{q}, \mathbf{r} \in \mathbb{R}_+^n$. *Then*

$$\begin{aligned} & \langle \mathbf{p}, \mathbf{r} \rangle \ln \left(\frac{1}{\langle \mathbf{p}, \mathbf{r} \rangle} \sum_{i=1}^{n} q_i r_i \right) + \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i} - \max_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}; \mathbf{r}) \\ \le\ & H(\mathbf{p}; \mathbf{r}) \\ \le\ & \langle \mathbf{p}, \mathbf{r} \rangle \ln \left(\frac{1}{\langle \mathbf{p}, \mathbf{r} \rangle} \sum_{i=1}^{n} q_i r_i \right) + \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i} - \min_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}; \mathbf{r}) \\ \le\ & \langle \mathbf{p}, \mathbf{r} \rangle \ln \left(\frac{1}{\langle \mathbf{p}, \mathbf{r} \rangle} \sum_{i=1}^{n} q_i r_i \right) + \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i}, \end{aligned} \tag{4.49}$$

where

$$R_{-\ln}(\mathbf{q}, \mathbf{p}; \mathbf{r}) = \langle \mathbf{q}, \mathbf{r} \rangle \ln \left(\frac{1}{\langle \mathbf{q}, \mathbf{r} \rangle} \sum_{i=1}^{n} \frac{q_i^2 r_i}{p_i} \right) - \sum_{i=1}^{n} q_i r_i \ln \frac{q_i}{p_i}.$$

If in addition $\mathbf{r} = \mathbf{e} = (1, ..., 1)$ *with* $\sum_{i=1}^{n} q_i = n$, *then*

$$\begin{aligned} & \ln n + \sum_{i=1}^{n} p_i \ln \frac{1}{q_i} - \max_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}) \\ \le\ & H(\mathbf{p}) \\ \le\ & \ln n + \sum_{i=1}^{n} p_i \ln \frac{1}{q_i} - \min_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}) \\ \le\ & \ln n + \sum_{i=1}^{n} p_i \ln \frac{1}{q_i}, \end{aligned} \tag{4.50}$$

where

$$R_{-\ln}(\mathbf{q}, \mathbf{p}) = n \ln \left(\frac{1}{n} \sum_{i=1}^{n} \frac{q_i^2}{p_i} \right) - \sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i}.$$

Proof: Applying (4.46) to the convex function $f(t) = -\ln t$, we get

$$\begin{aligned} 0 &\le \min_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}; \mathbf{r}) \le \sum_{i=1}^{n} r_i p_i \ln \frac{1}{q_i} - H(\mathbf{p}; \mathbf{r}) + \langle \mathbf{p}, \mathbf{r} \rangle \ln \left(\frac{1}{\langle \mathbf{p}, \mathbf{r} \rangle} \sum_{i=1}^{n} q_i r_i \right) \\ &\le \max_{1 \le i \le n} \left\{ \frac{p_i}{q_i} \right\} R_{-\ln}(\mathbf{q}, \mathbf{p}; \mathbf{r}), \end{aligned}$$

which is equivalent to (4.49).

Further, let $\sum_{i=1}^{n} q_i = n$ and $\mathbf{r} = \mathbf{e} = (1, ..., 1)$. Then $\langle \mathbf{p}, \mathbf{r} \rangle = \sum_{i=1}^{n} p_i = 1$, and from (4.49) we get (4.50). $\square$

4.1.4. New bounds of f-divergences for 3-convex functions

We connect the approach from Section 3.9.5. with Csiszár f-divergence and specified divergences as the Kullback-Leibler divergence, Hellinger divergence, Harmonic divergence, Bhattacharya distance, Triangular discrimination, Renyi divergence and derive new estimates for them. The resuls from this subsection are cited from [71].

In the next results, we use the following notation and assumptions:

(A_1) Let $\mathbf{p},\mathbf{q}$ be positive k-tuples, $\mathbf{R}\in\mathcal{M}_{kl}((0,\infty))$ be a matrix with row sums equal to 1 and with columns $\mathbf{r}_j=(r_{1j},...,r_{kj})$, $j=1,...,l$. We use notation $\langle\cdot,\cdot\rangle$ for the standard inner product, i.e. $\langle\mathbf{p},\mathbf{r}_j\rangle=\sum_{i=1}^k p_i r_{ij}$ and $\langle\mathbf{q},\mathbf{r}_j\rangle=\sum_{i=1}^k q_i r_{ij}$, for $j=1,...,l$. We assume that $\frac{q_i}{p_i}\in[\alpha,\beta], i=1,...,k$ and $\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\in[\alpha,\beta]$, $j=1,...,l,$ where $[\alpha,\beta]\subset(0,\infty)$. For $s\in[\alpha,\beta]$ and $G_z(\cdot,s), z\in\{1,2,3,4,5\}$, given as in (3.44)-(3.48), respectively, we define

$$\begin{aligned}\mathcal{G}_m(G_z(\cdot,s)) &= \sum_{i=1}^k p_i G_z\left(\frac{q_i}{p_i},s\right)-\sum_{j=1}^l\langle\mathbf{p},\mathbf{r}_j\rangle G_z\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle},s\right) \qquad (4.51)\\ &\quad -\sum_{j=1}^l\min_{1\le i\le k}\{p_i r_{ij}\}\left(\sum_{i=1}^k G_z\left(\frac{q_i}{p_i},s\right)-kG_z\left(\frac{1}{k}\sum_{i=1}^k\frac{q_i}{p_i},s\right)\right),\\ \mathcal{G}_M(G_z(\cdot,s)) &= \sum_{j=1}^l\max_{1\le i\le k}\{p_i r_{ij}\}\left(\sum_{i=1}^k G_z\left(\frac{q_i}{p_i},s\right)-kG_z\left(\frac{1}{k}\sum_{i=1}^k\frac{q_i}{p_i},s\right)\right)\\ &\quad -\sum_{i=1}^k p_i G_z\left(\frac{q_i}{p_i},s\right)+\sum_{j=1}^l\langle\mathbf{p},\mathbf{r}_j\rangle G_z\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle},s\right). \qquad (4.52)\end{aligned}$$

(A_2) Specially, if we set $l=1$ and $\mathbf{r}_1=\mathbf{e}=(1,...,1)$ with $\mathbf{p}$ and $\mathbf{q}$ probability distributions, from (4.51) and (4.52) we get

$$\bar{\mathcal{G}}_m(G_z(\cdot,s)) = \sum_{i=1}^{k} p_i G_z\left(\frac{q_i}{p_i}, s\right) - G_z(1,s) - \min_{1\le i\le k}\{p_i\}\left(\sum_{i=1}^{k} G_z\left(\frac{q_i}{p_i}, s\right) - kG_z\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}, s\right)\right), \tag{4.53}$$

$$\bar{\mathcal{G}}_M(G_z(\cdot,s)) = \max_{1\le i\le k}\{p_i\}\left(\sum_{i=1}^{k} G_z\left(\frac{q_i}{p_i}, s\right) - kG_z\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}, s\right)\right) - \sum_{i=1}^{k} p_i G_z\left(\frac{q_i}{p_i}, s\right) + G_z(1,s). \tag{4.54}$$

Applying results from Section 3.9.5. we obtain new estimates of Csiszár f-divergence $D_f(\mathbf{q},\mathbf{p})$ in case when a function f is 3-convex.

Corollary 4.1.10.: *Let (A_1) holds. Let $f\colon [\alpha,\beta] \to \mathbb{R}$ be 3-convex function such that $f'' \in AC([\alpha,\beta])$ and $f''' \in L_p[\alpha,\beta]$.*

i) For each $t \in [\alpha,\beta]$, such that $t \leqslant s$, then

$$\begin{aligned} 0 &\leqslant D_f(\mathbf{q},\mathbf{p}) - \sum_{j=1}^{l}\langle \mathbf{p},\mathbf{r}_j\rangle f\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) \\ &- \sum_{j=1}^{l}\min_{1\le i\le k}\{p_i r_{ij}\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta \mathcal{G}_m(G_z(\cdot,s))ds \\ &\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta \mathcal{G}_m(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}} \end{aligned} \tag{4.55}$$

and

$$\begin{aligned} 0 &\leqslant \sum_{j=1}^{l}\langle \mathbf{p},\mathbf{r}_j\rangle f\left(\frac{\langle \mathbf{q},\mathbf{r}_j\rangle}{\langle \mathbf{p},\mathbf{r}_j\rangle}\right) - D_f(\mathbf{q},\mathbf{p}) \\ &+ \sum_{j=1}^{l}\max_{1\le i\le k}\{p_i r_{ij}\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_\alpha^\beta \mathcal{G}_M(G_z(\cdot,s))ds \\ &\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta \mathcal{G}_M(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}}; \end{aligned} \tag{4.56}$$

ii) For each $t \in [\alpha, \beta]$, such that $t > s$, then

$$
\begin{aligned}
0 &\leqslant D_f(\mathbf{q}, \mathbf{p}) - \sum_{j=1}^{l} \langle \mathbf{p}, \mathbf{r}_j \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) \\
&- \sum_{j=1}^{l} \min_{1 \le i \le k} \{p_i r_{ij}\} \left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k} \frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta) - f'(\alpha)}{\beta - \alpha} \int_\alpha^\beta \mathcal{G}_m(G_z(\cdot, s)) ds \\
&\leqslant \frac{1}{\beta - \alpha} \left\| f''' \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta \mathcal{G}_m(G_z(\cdot, s))(t - \beta) ds\right)^q dt\right)^{\frac{1}{q}},
\end{aligned}
\tag{4.57}
$$

and

$$
\begin{aligned}
0 &\leqslant \sum_{j=1}^{l} \langle \mathbf{p}, \mathbf{r}_j \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) - D_f(\mathbf{q}, \mathbf{p}) \\
&+ \sum_{j=1}^{l} \max_{1 \le i \le k} \{p_i r_{ij}\} \left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k} \frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta) - f'(\alpha)}{\beta - \alpha} \int_\alpha^\beta \mathcal{G}_M(G_z(\cdot, s)) ds \\
&\leqslant \frac{1}{\beta - \alpha} \left\| f''' \right\|_p \left(\int_\alpha^\beta \left(\int_\alpha^\beta \mathcal{G}_M(G_z(\cdot, s))(t - \beta) ds\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned}
\tag{4.58}
$$

Proof: Let us consider vectors $\mathbf{x} = (x_1, ..., x_k)$, $\mathbf{y} = (y_1, ..., y_l)$, such that $x_i = \frac{q_i}{p_i}$, $i = 1, ..., k$, and $y_j = \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}$, $j = 1, ..., l$.

Let $\mathbf{a} = (a_1, ..., a_k)$ and $\mathbf{b} = (b_1, ..., b_l)$ be weights with $a_i = \sum\limits_{j=1}^{l} b_j \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle}, i = 1, ..., k$ and $b_j = \langle \mathbf{p}, \mathbf{r}_j \rangle$, $j = 1, ..., l$. Note that $a_i = \sum\limits_{j=1}^{l} b_j \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle} = p_i \sum\limits_{j=1}^{l} r_{ij} = p_i$, for $i = 1, ..., k$.

Also, we consider a matrix $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{kl}(\mathbb{R})$, with $s_{ij} = \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle}$, $i = 1, ..., k$, $j = 1, ..., l$.

Now, applying Corollary 3.10.20., i.e. the inequalities (3.256)-(3.259), we get required results. □

Specially, for $l = 1$ and $\mathbf{r}_1 = \mathbf{e} = (1, ..., 1)$ with $\mathbf{p}$ and $\mathbf{q}$ probability distributions, the previous result reduces to the next corollary.

Corollary 4.1.11.: *Let (A_2) holds. Let $f \colon [\alpha, \beta] \to \mathbb{R}$ be 3-convex function such that $f'' \in AC([\alpha, \beta])$ and $f''' \in L_p[\alpha, \beta]$.*

i) For each $t \in [\alpha, \beta]$, such that $t \leqslant s$, then

$$\begin{aligned}
0 &\leqslant D_f(\mathbf{q},\mathbf{p}) - f(1) \\
&- \min_{1\le i\le k}\{p_i\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_{\alpha}^{\beta}\bar{\mathcal{G}}_m(G_z(\cdot,s))ds \\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p\left(\int_{\alpha}^{\beta}\left(\int_{\alpha}^{\beta}\bar{\mathcal{G}}_m(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}},
\end{aligned} \tag{4.59}$$

and

$$\begin{aligned}
0 &\leqslant f(1) - D_f(\mathbf{q},\mathbf{p}) \\
&+ \max_{1\le i\le k}\{p_i\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_{\alpha}^{\beta}\bar{\mathcal{G}}_M(G_z(\cdot,s))ds \\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p\left(\int_{\alpha}^{\beta}\left(\int_{\alpha}^{\beta}\bar{\mathcal{G}}_M(G_z(\cdot,s))(t-\alpha)ds\right)^q dt\right)^{\frac{1}{q}};
\end{aligned} \tag{4.60}$$

ii) For each $t \in [\alpha, \beta]$, such that $t > s$, then

$$\begin{aligned}
0 &\leqslant D_f(\mathbf{q},\mathbf{p}) - f(1) \\
&- \min_{1\le i\le k}\{p_i\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_{\alpha}^{\beta}\bar{\mathcal{G}}_m(G_z(\cdot,s))ds \\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p\left(\int_{\alpha}^{\beta}\left(\int_{\alpha}^{\beta}\bar{\mathcal{G}}_m(G_z(\cdot,s))(t-\beta)ds\right)^q dt\right)^{\frac{1}{q}},
\end{aligned} \tag{4.61}$$

and

$$\begin{aligned}
0 &\leqslant f(1) - D_f(\mathbf{q},\mathbf{p}) \\
&+ \max_{1\le i\le k}\{p_i\}\left[\sum_{i=1}^{k} f\left(\frac{q_i}{p_i}\right) - kf\left(\frac{1}{k}\sum_{i=1}^{k}\frac{q_i}{p_i}\right)\right] - \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\int_{\alpha}^{\beta}\bar{\mathcal{G}}_M(G_z(\cdot,s))ds \\
&\leqslant \frac{1}{\beta-\alpha}\left\|f'''\right\|_p\left(\int_{\alpha}^{\beta}\left(\int_{\alpha}^{\beta}\bar{\mathcal{G}}_M(G_z(\cdot,s))(t-\beta)ds\right)^q dt\right)^{\frac{1}{q}}.
\end{aligned} \tag{4.62}$$

In the examples below, applying the previous corollaries to suitable choices of a kernel f, we obtain new bounds for some of most commonly used divergences which are special cases of Csiszár f-divergence functional (4.5).

Example 1.: *The Kullback-Leibler divergence (4.3) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t) = t\ln t$, $t > 0$. *Since* $f'''(t) = -\frac{1}{t^2} < 0$, *then the function* $g(t) = -f(t) = -t\ln t$ *satisfies the assumptions of Corollaries 4.1.10. and 4.1.11.. So for the Kullback-Leibler divergence, inequalities (4.55)-(4.62) hold with the reversed signs of inequalities and with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha} = \frac{\ln\beta-\ln\alpha}{\beta-\alpha}, \qquad \left\|f'''\right\|_p = \left(\frac{\beta^{1-2p}-\alpha^{1-2p}}{1-2p}\right)^{\frac{1}{p}}.$$

Example 2.: *Hellinger divergence (4.9) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t) = \frac{1}{2}\left(\sqrt{t}-1\right)^2$, $t > 0$. *Since* $f'''(t) = -\frac{3}{8}t^{-\frac{5}{2}} < 0$, *then the function* $g(t) = -f(t) = -\frac{1}{2}\left(\sqrt{t}-1\right)^2$ *satisfies the assumptions of Corollaries 4.1.10. and 4.1.11.. So for Hellinger divergence, inequalities (4.55)-(4.62) hold with the reversed signs of inequalities and with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha} = \frac{\sqrt{\beta}-\sqrt{\alpha}}{2\sqrt{\alpha\beta}(\beta-\alpha)}, \qquad \left\|f'''\right\|_p = -\frac{3}{8}\left(\frac{\beta^{1-\frac{5}{2}p}-\alpha^{1-\frac{5}{2}p}}{1-\frac{5}{2}p}\right)^{\frac{1}{p}}.$$

Example 3.: *Harmonic divergence (4.10) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t) = \frac{2t}{1+t}$, $t > 0$. *Since* $f'''(t) = \frac{12}{(1+t)^4} > 0$, *then* f *satisfies the assumptions of Corollaries 4.1.10. and 4.1.11.. So for Harmonic divergence inequalities (4.55)-(4.62) hold with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha} = \frac{-2(2+\beta+\alpha)}{(1+\alpha)^2(1+\beta)^2}, \qquad \left\|f'''\right\|_p = 12\left(\frac{(1+\beta)^{1-4p}-(1+\alpha)^{1-4p}}{1-4p}\right)^{\frac{1}{p}}.$$

Example 4.: *Bhattacharya distance (4.11) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t) = -\sqrt{t}$, $t > 0$. *Since* $f'''(t) = -\frac{3}{8}t^{-\frac{5}{2}} < 0$, *then the function* $g(t) = -f(t) = \sqrt{t}$ *satisfies the assumptions of Corollaries 4.1.10. and 4.1.11.. So for Bhattacharya distance, inequalities (4.55)-(4.62) hold with the reversed signs of inequalities and with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha} = \frac{\sqrt{\beta}+\sqrt{\alpha}}{2\sqrt{\alpha\beta}(\alpha-\beta)}, \qquad \left\|f'''\right\|_p = \frac{1}{8}\left(\frac{\alpha^{1-\frac{5}{2}p}-\beta^{1-\frac{5}{2}p}}{1-\frac{5}{2}p}\right)^{\frac{1}{p}}.$$

Example 5.: *Triangular discrimination (4.12) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t)=\frac{(t-1)^2}{t+1}$, $t>0$. *Since* $f'''(t)=-\frac{24}{(1+t)^4}<0$, *then the function* $g(t)=-f(t)=-\frac{(t-1)^2}{t+1}$ *satisfies the assumptions of Corollaries 4.1.10. and 4.1.11.. So for Triangular discrimination inequalities (4.55)-(4.62) hold with the reversed signs of inequalities and with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}=\frac{4\left[(\alpha+1)^2(\beta^2+2\beta-3)-(\beta+1)^2(\alpha^2+2\alpha-3)\right]}{(\alpha+1)^2(\beta+1)^2(\beta-\alpha)},\qquad \left\|f'''\right\|_p=24\left(\frac{(1+\alpha)^{1-4p}-(1+\beta)^{1-4p}}{1-4p}\right)^{\frac{1}{p}}.$$

Example 6.: *Rényi divergence (4.13) of probability distributions* $\mathbf{q}$ *and* $\mathbf{p}$ *is defined from (4.5) with the kernel function* $f(t)=t^{\lambda}$, $t>0$, $\lambda\in\mathbb{R}$. *Since* $f'''(t)=\lambda(\lambda-1)(\lambda-2)t^{\lambda-3}$, *then* $f(t)=t^{\lambda}$ *is* 3*-convex for* $0\leqslant\lambda\leqslant 1$ *and* $\lambda\geq 2$, *and* $g(t)=-f(t)=-t^{\lambda}$ *is* 3*-convex for* $\lambda\leqslant 0$ *and* $1<\lambda<2$. *As regards, for Rényi divergence, inequalities (4.55)-(4.62) hold for* $0\leqslant\lambda\leqslant 1$ *and* $\lambda\geq 2$, *and for* $\lambda\leqslant 0$ *and* $1<\lambda<2$ *the signs of inequalities are reversed, with*

$$\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}=\frac{\lambda\beta^{\lambda-1}-\lambda\alpha^{\lambda-1}}{\beta-\alpha},\qquad \left\|f'''\right\|_p=\lambda(\lambda-1)(\lambda-2)\left(\frac{\beta^{(\lambda-3)p+1}-\alpha^{(\lambda-3)p+1}}{(\lambda-3)p+1}\right)^{\frac{1}{p}}.$$

4.2. Estimates for the Zipf-Mandelbrot entropy

Zipf's law [162], also known as the Pareto law, is experimental law for the discrete probability distribution frequently used in information science, bibliometrics, linguistics, social sciences, economy, as well as in physics, biology, computer science etc. It was in the first place established with the frequency of the words in a text in view and as such it revealed a hyperbolic relation. If words of a language are sorted in the order of decreasing frequency of usage, a word's frequency f is inversely proportional to its rank r, or sequence number in the list and the product of these is a constant: $r\cdot f=C$ (a few occur very often while many others occur rarely). The term Zipfian distribution is used to describe various types of distributions of the probability occurrences which approximately follow the mathematical form of Zipf's law (see [44], [115], [117], [150]). Benoit Mandelbrot [105] generalized Zipf's law and gave its improvement for the count of the low-rank words [115].

In this section we will consider the results from the previous section involving f-divergences in the context of Zipf's law [162] and its generalized version (see [72]), the Zipf-Mandelbrot law [105], which we define in the sequel.

Definition 4.2.6.: The Zipf-Mandelbrot law is a discrete probability distribution depending on three parameters $N \in \mathbb{N}$, $t \geq 0$ and $s > 0$ with probability mass function defined with

$$\varphi(i, N, t, s) = \frac{1}{(i+t)^s H_{N,t,s}}, \quad i = 1, 2, ..., N, \tag{4.63}$$

where

$$H_{N,t,s} = \sum_{i=1}^{N} \frac{1}{(i+t)^s}. \tag{4.64}$$

As a special case of the Zipf-Mandelbrot law, for $t = 0$, we get **Zipf's law** with probability mass function

$$\varphi(i, N, s) = \frac{1}{i^s H_{N,s}}, \quad i = 1, 2, ..., N, \tag{4.65}$$

where

$$H_{N,s} = \sum_{i=1}^{N} \frac{1}{i^s}. \tag{4.66}$$

4.2.1. Bounds for the Zipf-Mandelbrot entropy

If we put $p_i = \frac{1}{(i+t)^s H_{N,t,s}}$ in (4.1), then Shannon's entropy becomes

$$\begin{aligned}
\sum_{i=1}^{N} \frac{\ln(i+t)^s H_{N,t,s}}{(i+t)^s H_{N,t,s}} &= \sum_{i=1}^{N} \frac{\ln(i+t)^s}{(i+t)^s H_{N,t,s}} + \sum_{i=1}^{N} \frac{\ln H_{N,t,s}}{(i+t)^s H_{N,t,s}} \\
&= \frac{s}{H_{N,t,s}} \sum_{i=1}^{N} \frac{\ln(i+t)}{(i+t)^s} + \ln H_{N,t,s} \\
&= \frac{s}{H_{N,t,s}} \sum_{i=1}^{N} \frac{\ln(i+t)}{(i+t)^s} + \ln H_{N,t,s},
\end{aligned}$$

i.e. we get Shannon's entropy for the Zipf-Mandelbrot law, so called **the Zipf-Mandelbrot entropy** defined as

$$Z(H, t, s) = \frac{s}{H_{N,t,s}} \sum_{i=1}^{N} \frac{\ln(i+t)}{(i+t)^s} + \ln H_{N,t,s}. \tag{4.67}$$

Using results for Sherman's functional from the second chapter we obtain lower and upper bounds for the Zipf-Mandelbrot entropy $Z(H,t,s)$ as follows.

Corollary 4.2.12.: *Let $N \in \mathbb{N}$, $t \geq 0$, $s > 0$ and $\mathbf{q} = (q_1, ..., q_N) \in [0,\infty)^N$ with $\sum_{i=1}^N q_i = N$. Let $H_{N,t,s}$ and $Z(H,t,s)$ be defined by (4.64) and (4.67), respectively. Then*

$$\begin{aligned} & \ln N + \frac{1}{H_{N,t,s}} \sum_{i=1}^{N} \frac{1}{(i+t)^s} \ln \frac{1}{q_i} - \frac{1}{H_{N,t,s}} \max_{1\leq i\leq N} \left\{ \frac{1}{(i+t)^s q_i} \right\} S(N,t,s,\mathbf{q}) \\ \leq\ & Z(H,t,s) \qquad (4.68) \\ \leq\ & \ln N + \frac{1}{H_{N,t,s}} \sum_{i=1}^{N} \frac{1}{(i+t)^s} \ln \frac{1}{q_i} - \frac{1}{H_{N,t,s}} \min_{1\leq i\leq N} \left\{ \frac{1}{(i+t)^s q_i} \right\} S(N,t,s,\mathbf{q}) \\ \leq\ & \ln N + \frac{1}{H_{N,t,s}} \sum_{i=1}^{N} \frac{1}{(i+t)^s} \ln \frac{1}{q_i}, \end{aligned}$$

where

$$S(N,t,s,\mathbf{q}) = N \ln \left(\frac{H_{N,t,s}}{N} \sum_{i=1}^{N} (i+t)^s q_i^2 \right) - \sum_{i=1}^{N} q_i \ln \left(q_i (i+t)^s H_{N,t,s} \right).$$

Proof: Using (4.50) we get

$$\begin{aligned} & \ln N + \sum_{i=1}^{N} \frac{1}{(i+t)^s H_{N,t,s}} \ln \frac{1}{q_i} - \max_{1\leq i\leq N} \left\{ \frac{\frac{1}{(i+t)^s H_{N,t,s}}}{q_i} \right\} S(N,t,s,\mathbf{q}) \\ \leq\ & Z(H,t,s) \\ \leq\ & \ln N + \sum_{i=1}^{N} \frac{1}{(i+t)^s H_{N,t,s}} \ln \frac{1}{q_i} - \min_{1\leq i\leq N} \left\{ \frac{\frac{1}{(i+t)^s H_{N,t,s}}}{q_i} \right\} S(N,t,s,\mathbf{q}) \\ \leq\ & \ln N + \sum_{i=1}^{N} \frac{1}{(i+t)^s H_{N,t,s}} \ln \frac{1}{q_i}, \end{aligned}$$

which is equivalent to (4.68). □

Corollary 4.2.13.: *Let $N \in \mathbb{N}$, $t \geq 0$, $s > 0$, $H_{N,t,s}$ and $Z(H,t,s)$ be defined by (4.64)* *and (4.67), respectively. Then*

$$\begin{aligned} \ln N - \frac{S(N,t,s)}{H_{N,t,s}(1+t)^s} &\leq Z(H,t,s) \qquad (4.69)\\ &\leq \ln N - \frac{S(N,t,s)}{H_{N,t,s}(N+t)^s} \\ &\leq \ln N, \end{aligned}$$

where

$$S(N,t,s) = N \ln\left(\frac{H_{N,t,s}}{N}\sum_{i=1}^{N}(i+t)^s\right) - \sum_{i=1}^{N}\ln\left((i+t)^s H_{N,t,s}\right).$$

Proof: If we choose $\mathbf{q} = \mathbf{e} = (1, ..., 1)$, then (4.68) becomes

$$\begin{aligned} \ln N - \frac{1}{H_{N,t,s}}\max_{1\leq i\leq N}\left\{\frac{1}{(i+t)^s}\right\} S(N,t,s) &\leq Z(H,t,s) \\ &\leq \ln N - \frac{1}{H_{N,t,s}}\min_{1\leq i\leq N}\left\{\frac{1}{(i+t)^s}\right\} S(N,t,s) \\ &\leq \ln N, \end{aligned}$$

which is equivalent to (4.69). □

4.2.2. On the Zipf–Mandelbrot entropy and 3-convex functions

In this subsection we connect the approach from Subsection 4.1.4. with the Zipf-Mandelbrot law and its special form Zipf's law (see [71]).

We define $\mathbf{p}$ and $\mathbf{q}$ via (4.63) as the Zipf-Mandelbrot law N-tuples, i.e. we put $p_i = \frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}$ and $q_i = \frac{1}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}}$, for $i = 1, ..., N$, $N \in \mathbb{N}$, and $\tilde{t}_1, \tilde{t}_2 \geq 0$, $\tilde{s}_1, \tilde{s}_2 > 0$. Then by (4.5), Csiszár f-divergence becomes

$$D_f(N, \tilde{t}_1, \tilde{t}_2, \tilde{s}_1, \tilde{s}_2) = \sum_{i=1}^{N} \frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}} f\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}}\right),$$

for $f\colon [\alpha,\beta] \to \mathbb{R}$, $0 < \alpha < \beta$, and $\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \in [\alpha,\beta]$, $i = 1, ..., N$.

If $\mathbf{p}$ and $\mathbf{q}$ are defined via Zipf's law (4.65), i.e. for $\tilde{t}_1 = \tilde{t}_2 = 0$, then Csiszár f-divergence functional assumes the following form

$$D_f(N,\tilde{s}_1,\tilde{s}_2)=\sum_{i=1}^{N}\frac{1}{i^{\tilde{s}_2}H_{N,\tilde{s}_2}}f\left(i^{\tilde{s}_2-\tilde{s}_1}\frac{H_{N,\tilde{s}_2}}{H_{N,\tilde{s}_1}}\right). \tag{4.70}$$

In the next results we also suppose that $\frac{\sum_{i=1}^{N}(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{\sum_{i=1}^{N}(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}}\in[\alpha,\beta]$. Further, for some $s\in[\alpha,\beta]$, $G_z(\cdot,s), z\in\{1,2,3,4,5\}$, given as in (3.44)-(3.44), respectively, we define

$$\begin{aligned}
&\tilde{\mathcal{G}}_m(G_z(\cdot,s)) \qquad (4.71)\\
&=\sum_{i=1}^{N}\frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}G_z\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)-G_z(1,s)\\
&-m_N\left[\sum_{i=1}^{N}G_z\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)-NG_z\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)\right],
\end{aligned}$$

where $m_N=\min_{1\le i\le N}\left\{\frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}\right\}$, and

$$\begin{aligned}
&\tilde{\mathcal{G}}_M(G_z(\cdot,s)) \qquad (4.72)\\
&=M_N\left[\sum_{i=1}^{N}G_z\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)-NG_z\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)\right]\\
&-\sum_{i=1}^{N}\frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}G_z\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1}H_{N,\tilde{t}_1,\tilde{s}_1}},s\right)+G_z(1,s)
\end{aligned}$$

where $M_N=\max_{1\le i\le N}\left\{\frac{1}{(i+\tilde{t}_2)^{\tilde{s}_2}H_{N,\tilde{t}_2,\tilde{s}_2}}\right\}$.

Taking into account these remarks, as direct consequence of Corollary 4.1.11., we get the following result.

Corollary 4.2.14.: *Let $f\colon[\alpha,\beta]\to\mathbb{R}$ be 3-convex function such that $f''\in AC([\alpha,\beta])$ and $f'''\in L_p[\alpha,\beta]$.*

i) For each $t \in [\alpha, \beta]$*, such that* $t \leqslant s$*, then*

$$\begin{aligned}
0 &\leqslant D_f(N, \tilde{t}_1, \tilde{t}_2, \tilde{s}_1, \tilde{s}_2) - f(1) \\
&- m_N \left[\sum_{i=1}^{N} f\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) - N f\left(\frac{1}{N} \sum_{i=1}^{N} \frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) \right] \\
&- \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha} \int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot, s)) ds \\
&\leqslant \frac{1}{\beta-\alpha} \left\| f''' \right\|_p \left(\int_{\alpha}^{\beta} \left(\int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot, s))(t-\alpha) ds \right)^q dt \right)^{\frac{1}{q}}
\end{aligned} \tag{4.73}$$

and

$$\begin{aligned}
0 &\leqslant f(1) - D_f((N, \tilde{t}_1, \tilde{t}_2, \tilde{s}_1, \tilde{s}_2) \\
&+ M_N \left[\sum_{i=1}^{N} f\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) - N f\left(\frac{1}{N} \sum_{i=1}^{N} \frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) \right] \\
&- \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha} \int_{\alpha}^{\beta} \tilde{\mathcal{G}}_M(G_z(\cdot, s)) ds \\
&\leqslant \frac{1}{\beta-\alpha} \left\| f''' \right\|_p \left(\int_{\alpha}^{\beta} \left(\int_{\alpha}^{\beta} \tilde{\mathcal{G}}_M(G_z(\cdot, s))(t-\alpha) ds \right)^q dt \right)^{\frac{1}{q}};
\end{aligned} \tag{4.74}$$

ii) For each $t \in [\alpha, \beta]$*, such that* $t > s$*, then*

$$\begin{aligned}
0 &\leqslant D_f(N, \tilde{t}_1, \tilde{t}_2, \tilde{s}_1, \tilde{s}_2) - f(1) \\
&- m_N \left[\sum_{i=1}^{N} f\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) - N f\left(\frac{1}{N} \sum_{i=1}^{N} \frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) \right] \\
&- \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha} \int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot, s)) ds \\
&\leqslant \frac{1}{\beta-\alpha} \left\| f''' \right\|_p \left(\int_{\alpha}^{\beta} \left(\int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot, s))(t-\beta) ds \right)^q dt \right)^{\frac{1}{q}}
\end{aligned} \tag{4.75}$$

and

$$
\begin{aligned}
0 &\leqslant f(1) - D_f((N, \tilde{t}_1, \tilde{t}_2, \tilde{s}_1, \tilde{s}_2) \qquad (4.76)\\
&+ M_N \left[\sum_{i=1}^{N} f\left(\frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) - Nf\left(\frac{1}{N} \sum_{i=1}^{N} \frac{(i+\tilde{t}_2)^{\tilde{s}_2} H_{N,\tilde{t}_2,\tilde{s}_2}}{(i+\tilde{t}_1)^{\tilde{s}_1} H_{N,\tilde{t}_1,\tilde{s}_1}} \right) \right]\\
&- \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha} \int_{\alpha}^{\beta} \tilde{\mathcal{G}}_M(G_z(\cdot,s))ds\\
&\leqslant \frac{1}{\beta-\alpha} \left\| f''' \right\|_p \left(\int_{\alpha}^{\beta} \left(\int_{\alpha}^{\beta} \tilde{\mathcal{G}}_M(G_z(\cdot,s))(t-\beta)ds \right)^q dt \right)^{\frac{1}{q}}.
\end{aligned}
$$

Remark 4.2.5.: If $\tilde{t}_1, \tilde{t}_2 = 0$, then we can get analogous results as in (4.73) and (4.74) observed in the context of Zipf's law.

At the end we give one linguistic example using experimentally obtained values of coefficients from Zipf's law assigned to different languages.

Example 7.: *Gelbukh and Sidorov [55] were analyzed the coefficients s_1 and s_2 from Zipf's law assigned to the Russian and English languages. Calculating coefficients s_1 and s_2 and their difference for each of the* 39 *literature texts in both languages, with more than* 10000 *running words inside of each of them, they obtained the average $s_1 = 0.973863$ for the English and $s_2 = 0.892869$ for the Russian language. In this context, using experimentally obtained values of s_1 and s_2, we present the bounds for f-divergence functional (4.70) which depend only on the parameter N and the value of chosen Green's function G_z.*

Namely, for $\mathbf{p} = (p_1, ..., p_N)$ and $\mathbf{q} = (q_1, ..., q_N)$ probabilities associated to the Russian and English languages, respectively, and $N \in \mathbb{N}$, $\tilde{t}_1, \tilde{t}_2 = 0$, from (4.73) we have

$$
\begin{aligned}
0 &\leqslant \sum_{i=1}^{N} \frac{1}{i^{0.973863} H_{N,0.973863}} f\left(i^{0.080994} \frac{H_{N,0.973863}}{H_{N,0.892869}} \right) - f(1)\\
&- \bar{m}_N \left[\sum_{i=1}^{N} f\left(i^{0.080994} \frac{H_{N,0.973863}}{H_{N,0.892869}} \right) - Nf\left(\frac{1}{N} \sum_{i=1}^{N} i^{0.080994} \frac{H_{N,0.973863}}{H_{N,0.892869}} \right) \right]\\
&- \tfrac{f'(\beta)-f'(\alpha)}{\beta-\alpha} \int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot,s))ds\\
&\leqslant \frac{1}{\beta-\alpha} \left\| f''' \right\|_p \left(\int_{\alpha}^{\beta} \left(\int_{\alpha}^{\beta} \tilde{\mathcal{G}}_m(G_z(\cdot,s))(t-\alpha)ds \right)^q dt \right)^{\frac{1}{q}},
\end{aligned}
$$

where

$$H_{N,0.973863}=\sum_{i=1}^{N}\frac{1}{i^{0.973863}},\quad H_{N,0.892869}=\sum_{i=1}^{N}\frac{1}{i^{0.892869}},\quad \bar{m}_N=\frac{1}{i^{0.973863}\sum\limits_{i=1}^{N}\frac{1}{i^{0.973863}}}$$

and $\tilde{\mathcal{G}}_m(G_z(\cdot,s))$ *is defined by*

$$\begin{aligned}
&\sum_{i=1}^{N}\frac{1}{i^{0.973863}H_{N,0.973863}}G_z\left(i^{0.080994}\frac{H_{N,0.973863}}{H_{N,0.892869}},s\right)\\
&-\sum_{i=1}^{N}\frac{1}{i^{0.973863}H_{N,0.973863}}G_z\left(\frac{\sum\limits_{i=1}^{N}i^{0.973863}H_{N,0.973863}}{\sum\limits_{i=1}^{N}i^{0.892869}H_{N,0.892869}},s\right)\\
&-\bar{m}_N\left[\sum_{i=1}^{N}G_z\left(i^{0.080994}\frac{H_{N,0.973863}}{H_{N,0.892869}},s\right)-NG_z\left(\frac{1}{N}\sum_{i=1}^{N}i^{0.080994}\frac{H_{N,0.973863}}{H_{N,0.892869}},s\right)\right].
\end{aligned}$$

Analogously, we can apply other results from the Corollary 4.2.14..

CHAPTER 5.

Sherman's type inequalities for operators on a Hilbert space with applications

In this chapter, we give extensions of Sherman's inequality (2.10) to self-adjoint operators and positive linear maps. We present general version of Sherman's operator inequality and its complement, and as easy consequences we derive Jensen's operator inequality as well as majorization operator inequalities and their conversions. We discuss about applications to some operator inequalities related to connections, solidarities and multidimensional weighted geometric mean. Further, we extend the results for Csiszár's f-divergence functional (4.5) to operator Csiszár f-divergence. In consequence, we get the Shannon like inequality for operators. At the end, we investigate vectorial joint majorization with applications to operator Csiszár f-divergence. Material of this chapter is quoted mainly from [70, 125, 128, 132].

5.1. Sherman's operator inequality

In this section we deal with Sherman's operator inequality and its complementary inequalities with applications. We deal with Sherman's inequality whose real arguments are substituted with the bounded self-adjoint operators acting on a Hilbert space. We give extensions of Sherman's inequality (2.10) to self-adjoint operators and positive linear maps. We also get upper bounds for obtained inequalities, so called converse inequalities. As easy consequences of obtained results we get Jensen's operator inequality (5.1) and the majorization operator inequalities as well as their conversions. We introduce so called Sherman's operator, deduced from Sherman's operator inequality, and study its properties. Using extended idea of convexity to operator functions of several variables, we obtain multidimensional Sherman's operator inequality which reduces to multidimensional Jensen's and the majorization operator inequalities. Moreover, we define multidimensional Sherman's operator and study its properties. At the end, we discuss about applications to some operator inequalities related to connections, solidarities and multidimensional weighted geometric mean. Results in this section are cited from [70].

First of all, we recall a version of Choi-Davis-Jensen inequality (1.26) for multiple operators and mappings, which is known as Jensen's operator inequality.

Theorem 5.1.1: *Let $(X_1, X_2, \ldots, X_n)$ be $n-$tuple of self-adjoint operators in $\mathcal{B}_h(\mathcal{H})$ with spectra in an interval J, $(\Phi_1, \Phi_2, \ldots, \Phi_n)$ be an $n-$tuple positive linear mappings $\Phi_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, $i = 1, 2, \ldots, n$, and $(a_1, a_2, \ldots, a_n) \in [0, \infty)^n$ be a vector of non-negative numbers. If $f \in \mathcal{C}(J)$ is an operator convex function and one of the following conditions is true:*

(A_1) $\Phi = \sum_{i=1}^{n} a_i \Phi_i$ *is a unital mapping,*

(A_2) $\sum_{i=1}^{n} a_i = 1$ *and* Φ_i *is a unital mapping* $(i = 1, 2, \ldots, n)$,

(A_3) $\sum_{i=1}^{n} a_i = 1$, *J containing* 0, $f(0) = 0$ *and* $0 < \Phi_i(1_{\mathcal{H}}) < 1_{\mathcal{K}}$ $(i = 1, 2, \ldots, n)$,

(A_4) $\sum_{i=1}^{n} a_i = 1$, *J containing* 0, $f(0) = 0$ *and* $0 < \sum_{i=1}^{n} \Phi_i(1_{\mathcal{H}}) \leqslant 1_{\mathcal{K}}$,

then

$$f\left(\sum_{i=1}^{n} a_i \Phi_i(X_i)\right) \leqslant \sum_{i=1}^{n} a_i \Phi_i\big(f(X_i)\big). \tag{5.1}$$

If f is operator concave, then the reverse inequality is valid in (5.1).

Theorem 5.1.1 follows from [51], [54] and [86]. We present the proof to interested readers.

Proof: We will prove only the operator convex case.

(A_1) Since $a_i \geq 0$, then $\Psi_i = a_i \Phi_i$ is positive linear mapping and $\sum\limits_{i=1}^{n} \Psi_i(1_H) = 1_K$. So, using the Jensen type operator inequality (see [51, Theorem 8.9]):

$$f\left(\sum_{i=1}^{n} \Psi_i(X_i)\right) \leqslant \sum_{i=1}^{n} \Psi_i\big(f(X_i)\big), \tag{5.2}$$

we obtain (5.1).

(A_2) Let $B = (B_1, B_2, \ldots, B_n)$ be $n-$tuple of self-adjoint operators with spectra in an interval J. If $\sum\limits_{i=1}^{n} a_i = 1$, then $\Phi(B) = \sum\limits_{i=1}^{n} a_i B_i$, is unital positive linear mapping and using (A_1) gives the following Jensen operator inequality holds

$$f\left(\sum_{j=1}^{n} a_i B_i\right) \leqslant \sum_{j=1}^{n} a_i f(B_i). \tag{5.3}$$

Now, applying (5.3) and (1.26), we have

$$f\left(\sum_{i=1}^{n} a_i \Phi_i(X_i)\right) \leqslant \sum_{i=1}^{n} a_i f\big(\Phi_i(X_i)\big) \leqslant \sum_{i=1}^{n} a_i \Phi_i\big(f(X_i)\big), \tag{5.4}$$

which gives the desired inequality (5.1).

(A_3) We obtain (5.4) again, if we applying (1.26) to a positive linear mapping Φ_i such that $0 < \Phi_i(1_H) \leqslant 1_K$.

(A_4) Since $0 < \sum\limits_{i=1}^{n} \Phi_i(1_H) \leqslant 1_K$ imply $0 < \Phi_i(1_H) \leqslant 1_K$, $i = 1, 2, \ldots, n$, for positive mappings, then applying (A_3) we obtain (5.1). □

Now we give a general version of Sherman's operator inequality.

Theorem 5.1.2 ***(Sherman's operator inequality):*** *Let $(X_1, X_2, \ldots, X_l)$ be $l-$tuple of self-adjoint operators in $\mathcal{B}_h(H)$ with spectra in J, $(\Phi_1, \Phi_2, \ldots, \Phi_l)$ of $l-$tuple positive linear mappings $\Phi_j : \mathcal{B}(H) \to \mathcal{B}(K)$, $j = 1, 2, \ldots, l$, and $(a_1, a_2, \ldots, a_l) \in [0, \infty)^l$, $(b_1, b_2 \ldots, b_n) \in [0, \infty)^n$, $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nl}(\mathbb{R})$ be matrix with non negative entries such that $\mathbf{a} = \mathbf{bS}$.*

If $f \in \mathcal{C}(J)$ is an operator convex function and one of the following conditions is true:

(i) $\Psi_i = \sum\limits_{j=1}^{l} s_{ij} \Phi_j$ *is unital mapping,* $i = 1, 2, \ldots, n$,

(ii) $\mathbf{S}$ *is row stochastic and* Φ_j *is unital,* $j = 1, 2, \ldots, l$,

(iii) $\mathbf{S}$ *is row stochastic,* $0 \in J$, $f(0) = 0$ *and* $0 < \Phi_j(1_H) < 1_K$, $j = 1, 2, \ldots, l$,

(*iv*) $\mathbf{S}$ *is row stochastic,* $0 \in J$, $f(0)=0$ *and* $0 < \sum_{j=1}^{l} \Phi_j(1_H) \leqslant 1_K$,

then

$$\sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right) \leqslant \sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right). \tag{5.5}$$

If f is operator concave, then the reverse inequality is valid in (5.5).

Proof: We will prove only the operator convex case. Applying Theorem 5.1.1 and since $\mathbf{a}=\mathbf{bS}$ we have

$$\begin{aligned}\sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right) &\leqslant \sum_{i=1}^{n} b_i \left(\sum_{j=1}^{l} s_{ij}\Phi_j\left(f(X_j)\right)\right)\\ &= \sum_{j=1}^{l}\left(\sum_{i=1}^{n} b_i s_{ij}\right)\Phi_j\left(f(X_j)\right) = \sum_{j=1}^{l} a_j\Phi_j\left(f(X_j)\right).\end{aligned}$$

□

Remark 5.1.1.: Setting $n=1$ and $\mathbf{b}=(1)$ in Theorem 5.1.2 we get Theorem 5.1.1, i.e. Sherman's operator inequality (5.5) reduces to Jensen's operator inequality (5.1). Moreover, setting $\mathbf{a}=(1,1,\ldots,1)$ the inequality (5.5) reduces to the form of (5.2). Also, setting $\Phi_j(A)=A$ in (5.5) we get Jensen's operator inequality (5.3).

Let $\mathbf{X}=(X_1,X_2,\ldots,X_l)$ be $l-$tuple and $\mathbf{Y}=(Y_1,Y_2,\ldots,Y_n)$ be $n-$tuple of self-adjoint operators in $\mathcal{B}_h(H)$, and $\mathbf{a}=(a_1,a_2,\ldots,a_l)\in[0,\infty)^l$, $\mathbf{b}=(b_1,b_2,\ldots,b_n)\in[0,\infty)^n$ be nonnegative weights. Inspired by (2.1) and (2.7), we introduce the following two symbols:

- We will use a symbol $\mathbf{Y}\prec_\circ \mathbf{X}$ if $n=l$ and $\mathbf{Y}=\mathbf{XS}$, i.e. $Y_i=\sum_{j=1}^{n} s_{ij}X_j$ $(i=1,2,\ldots,n)$ for some doubly stochastic $n\times n$ matrix $\mathbf{S}=(s_{ij})\in\mathcal{M}_{nn}(\mathbb{R})$.
- We will use a symbol $\langle\mathbf{a},\mathbf{Y}\rangle \prec_\circ \langle\mathbf{b},\mathbf{X}\rangle$ if $\mathbf{Y}=\mathbf{XS}^T$ and $\mathbf{a}=\mathbf{bS}$, i.e. $Y_i=\sum_{j=1}^{l} s_{ij}X_j$ $(i=1,2,\ldots,n)$ and $a_j=\sum_{i=1}^{n} b_i s_{ij}$ $(j=1,2,\ldots,l)$ for some row stochastic matrix $\mathbf{S}=(s_{ij})\in\mathcal{M}_{nl}(\mathbb{R})$.

By applying Theorem 5.1.2 on the identity mappings we get the operator version of Sherman's inequality and the weighted majorization operator inequality.

Corollary 5.1.1.: *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_l)$ *be* $l-$*tuple and* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be* $n-$*tuple of self-adjoint operators in* $\mathcal{B}_h(H)$ *with spectra in* J, $\mathbf{a} = (a_1, a_2, \ldots, a_l) \in [0, \infty)^l$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0, \infty)^n$ *be vectors and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nl}(\mathbb{R})$ *be a row stochastic matrix such that* $\langle \mathbf{a}, \mathbf{Y} \rangle \prec_\circ \langle \mathbf{b}, \mathbf{X} \rangle$. *If* $f \in \mathcal{C}(J)$ *is an operator convex function then*

$$\sum_{i=1}^{n} b_i f(Y_i) \leqslant \sum_{j=1}^{l} a_j f(X_j). \tag{5.6}$$

Especially, if $l = n$ *and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ *is a doubly stochastic matrix, then the weighted multidimensional majorization inequality holds, so that*

$$\mathbf{Y} \prec_\circ \mathbf{X} \qquad \textit{implies} \qquad \sum_{i=1}^{n} a_i f(Y_i) \leqslant \sum_{i=1}^{n} a_i f(X_i). \tag{5.7}$$

If f *is operator concave, then the reverse inequality is valid in* (5.6) *and* (5.7).

Proof: By applying Theorem 5.1.2 on the identity mappings, i.e. $\Phi_j(A) = A$, for all $A \in \mathcal{B}(H)$, $j = 1, 2, \ldots, l$, we obtain (5.6). Next, setting $n = l$ in (5.6) and all weights a_i and b_j are equal and nonnegative, the condition $\mathbf{a} = \mathbf{bS}$ assures stochasticity on columns of $\mathbf{S}$, so in that case we deal with doubly stochastic matrix and (5.7) holds. □

Setting $l = n$, $\mathbf{b} = (1, 1, \ldots, 1)$ and if $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ is doubly stochastic, we obtain that $\mathbf{a} = (1, 1, \ldots, 1)$. By applying Theorem 5.1.2 we get the following versions of the majorization operator inequality with mappings.

Corollary 5.1.2.: *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, *and* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be* $n-$*tuples of self-adjoint operators in* $\mathcal{B}_h(H)$ *with spectra in* J, $(\Phi_1, \Phi_2, \ldots, \Phi_n)$ *of* $n-$*tuple positive linear mappings* $\Phi_j : \mathcal{B}(H) \to \mathcal{B}(K)$, $j = 1, 2, \ldots, n$, $\mathbf{\Phi}(\mathbf{X}) = (\Phi_1(X_1), \Phi_2(X_2), \ldots, \Phi_n(X_n))$ *and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ *be a doubly stochastic matrix such that* $\mathbf{Y} \prec_\circ \mathbf{\Phi}(\mathbf{X})$.

If $f \in \mathcal{C}(J)$ *is an operator convex function and one of the following conditions is true:*

(*a*) $\Psi_i = \sum_{j=1}^{n} s_{ij}\Phi_j$ *is unital mapping,* $i = 1, 2, \ldots, n$,

(*b*) Φ_j *is unital,* $j = 1, 2, \ldots, n$,

(*c*) $0 \in J$, $f(0) = 0$ *and* $0 < \Phi_j(1_H) < 1_K$, $j = 1, 2, \ldots, n$,

(*d*) $0 \in J$, $f(0) = 0$ *and* $0 < \sum_{j=1}^{n} \Phi_j(1_H) \leqslant 1_K$,

then

$$\sum_{i=1}^{n} f\left(Y_i\right) \leqslant \sum_{i=1}^{n} \Phi_i\left(f(X_i)\right). \tag{5.8}$$

If f is operator concave, then the reverse inequalities are valid in (5.8).

5.1.1. Inequalities complementary to Sherman's operator inequality

Now we give a general version of complementary to Sherman's operator inequality (5.5).

Theorem 5.1.3: *Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$ be as in Theorem 5.1.2 and $J=[m,M]$, $m<M$. Let $m_Y=\min\limits_{1\le i\le n}\{m_i\}$, $M_Y=\max\limits_{1\le i\le n}\{M_i\}$, where m_i and M_i, $m_i\le M_i$, are the bounds of the self-adjoint operator $Y_i=\sum\limits_{j=1}^{l} s_{ij}\Phi_j(X_j)$ $(i=1,2,\ldots,n)$. Let $f\in\mathcal{C}([m,M])$, $g\in\mathcal{C}([m_Y,M_Y])$, $F:U\times V\to\mathbb{R}$ be bounded, where $f([m,M])\subseteq U$, $g([m_Y,M_Y])\subseteq V$ and one of the conditions (i) or (ii) in Theorem 5.1.2 is true.*

If f is convex and F is operator monotone in the first variable, then

$$\begin{aligned}
&F\left[\sum_{j=1}^{l} a_j\Phi_j\left(f(X_j)\right), \sum_{i=1}^{n} b_i g\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right)\right]\\
\leqslant\;& \max_{m_Y\le t\le M_Y} F\left[\left(\frac{M-t}{M-m}f(m)+\frac{t-m}{M-m}f(M)\right)\|\mathbf{b}\|_1, g\left(t\right)\|\mathbf{b}\|_1\right]1_K \\
\leqslant\;& \max_{m\le t\le M} F\left[\left(\frac{M-t}{M-m}f(m)+\frac{t-m}{M-m}f(M)\right)\|\mathbf{b}\|_1, g\left(t\right)\|\mathbf{b}\|_1\right]1_K,
\end{aligned} \tag{5.9}$$

where $\|\cdot\|_1$ denotes l_1 norm.

If f is concave, then the reverse inequality is valid in (5.9) *with* min *instead of* max.

Proof: We only prove the convex case. Since $m\,\Phi_j(1_H)\le\Phi_j(X_j)\le M\,\Phi_j(1_H)$ and $\sum\limits_{j=1}^{l} s_{ij}\Phi_j(1_H)=1_K$ $(i=1,2,\ldots,n)$, then $m1_K\le\sum\limits_{j=1}^{l} s_{ij}\Phi_j(X_j)\le M1_K$. Furthermore $m_Y 1_K\le\sum\limits_{j=1}^{l} s_{ij}\Phi_j(X_j)\le M_Y 1_K$, so $[m_Y,M_Y]\subseteq[m,M]$.

By using convexity of f and functional calculus, we obtain

$$\begin{aligned}
&\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right)\\
\leq\ &\sum_{j=1}^{l} a_j \left(\frac{M\Phi_j(1_H)-\Phi_j(X_j)}{M-m} f(m)+\frac{\Phi_j(X_j)-m\Phi_j(1_H)}{M-m} f(M)\right)\\
=\ &\sum_{j=1}^{l} \left(\sum_{i=1}^{n} b_i s_{ij}\right) \left(\frac{M\Phi_j(1_H)-\Phi_j(X_j)}{M-m} f(m)+\frac{\Phi_j(X_j)-m\Phi_j(1_H)}{M-m} f(M)\right)\\
=\ &\sum_{i=1}^{n} b_i \left(\frac{M1_K-\sum\limits_{j=1}^{l} s_{ij}\Phi_j(X_j)}{M-m} f(m)+\frac{\sum\limits_{j=1}^{l} s_{ij}\Phi_j(X_j)-m1_K}{M-m} f(M)\right).
\end{aligned}$$

Using operator monotonicity of $u \mapsto F(u,v)$ and boundedness of F, it follows

$$\begin{aligned}
&F\left[\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right), \sum_{i=1}^{n} b_i g\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right)\right]\\
\leq\ &F\left[\sum_{i=1}^{n} b_i \left(\frac{M1_K-\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)}{M-m} f(m)+\frac{\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)-m1_K}{M-m} f(M)\right), \sum_{i=1}^{n} b_i g\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right)\right]\\
\leq\ &\max_{m_Y\leq t\leq M_Y} F\left[\sum_{i=1}^{n} b_i \left(\frac{M-t}{M-m} f(m)+\frac{t-m}{M-m} f(M)\right), \sum_{i=1}^{n} b_i g(t)\right] 1_K\\
=\ &\max_{m_Y\leq t\leq M_Y} F\left[\left(\frac{M-t}{M-m} f(m)+\frac{t-m}{M-m} f(M)\right)\|\mathbf{b}\|_1, g\left(t\right)\|\mathbf{b}\|_1\right] 1_K\\
\leq\ &\max_{m\leq t\leq M} F\left[\left(\frac{M-t}{M-m} f(m)+\frac{t-m}{M-m} f(M)\right)\|\mathbf{b}\|_1, g\left(t\right)\|\mathbf{b}\|_1\right] 1_K.
\end{aligned}$$

□

5.1.2. Difference type complementary inequalities

In this subsection we consider the difference type complements to the inequality (5.5).

For convenience we introduce some abbreviations. Let $f : [m, M] \to \mathbb{R}$, $m < M$, be a convex or a concave function. We denote a linear function through $(m, f(m))$ and $(M, f(M))$ by $f^{cho}_{[m,M]}$, i.e.

$$f^{cho}_{[m,M]}(t) = \frac{M-t}{M-m} f(m)+\frac{t-m}{M-m} f(M), \qquad t\in\mathbb{R}$$

and the slope and the intercept by k_f and l_f, respectively, i.e.

$$k_f = \frac{f(M) - f(m)}{M - m} \qquad \text{and} \qquad l_f = \frac{Mf(m) - mf(M)}{M - m}.$$

Putting $F(u,v) = u - \alpha v$, $\alpha \in \mathbb{R}$, in Theorem 5.1.3 we obtain the following corollary.

Corollary 5.1.3.: *Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$, m, M, m_Y, M_Y and g be as in Theorem 5.1.3 and let one of the conditions (i) or (ii) in Theorem 5.1.2 be true.*

If $f \in \mathcal{C}([m,M])$ is convex, then

$$\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right) \leqslant \alpha \sum_{i=1}^{n} b_i g\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right) + \max_{m_Y \le t \le M_Y} \left\{k_f t + l_f - \alpha g\left(t\right)\right\} \|\mathbf{b}\|_1 1_K. \tag{5.10}$$

If f is concave, then the reverse inequality is valid in (5.9) *with* min *instead of* max.

Remark 5.1.2.: Now we give a way of determining the bound

$$C_\alpha := \max_{m_Y \le t \le M_Y} \left\{k_f t + l_f - \alpha g(t)\right\}$$

placed in Corollary 5.1.3. (see [109, Corollary 3.2]). Let f is convex.

If αg is concave, then

$$C_\alpha = \max_{m_Y \le t \le M_Y} \left\{ f^{cho}_{[m,M]}(m_Y) - \alpha g(m_Y)\, ,\, f^{cho}_{[m,M]}(M_Y) - \alpha g(M_Y) \right\}.$$

But, if αg is convex, then

$$C_\alpha = \begin{cases} f^{cho}_{[m,M]}(m_Y) - \alpha g(m_Y) & \text{if} \quad \alpha g'_-(t) \ge k_f \ \text{ for every } t \in (m_Y, M_Y), \\ f^{cho}_{[m,M]}(t_0) - \alpha g(t_0) & \text{if} \quad \alpha g'_-(t_0) \le k_f \le \alpha g'_+(t_0) \ \text{ for some } t_0 \in (m_Y, M_Y), \\ f^{cho}_{[m,M]}(M_Y) - \alpha g(M_Y) & \text{if} \quad \alpha g'_+(t) \le k_f \ \text{ for every } t \in (m_Y, M_Y). \end{cases} \tag{5.11}$$

Putting $g \equiv f$ and $\alpha = 1$ in Corollary 5.1.3. we obtain the following result.

Corollary 5.1.4.: *Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$, m, M and m_Y, M_Y be as in Theorem 5.1.3 and* *let one of the conditions (i) or (ii) in Theorem 5.1.2 be true.*

If $f \in \mathcal{C}([m,M])$ is convex, then

$$\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right) \leqslant \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right) + \bar{C}\|\mathbf{b}\|_1 1_K, \tag{5.12}$$

where the value of the constant $\bar{C} := \max_{m_Y \le t \le M_Y} \{k_f t + l_f - f(t)\}$ *determined as in* (5.11) *with* $g = f$, *i.e.*

$$\bar{C} = \begin{cases} f^{cho}_{[m,M]}(m_Y) - f(m_Y) & \text{if} \quad f'_-(t) \ge k_f \ \text{ for every } t \in (m_Y, M_Y), \\ f^{cho}_{[m,M]}(t_0) - f(t_0) & \text{if} \quad f'_-(t_0) \le k_f \le \ f'_+(t_0) \ \text{ for some } t_0 \in (m_Y, M_Y), \\ f^{cho}_{[m,M]}(M_Y) - f(M_Y) & \text{if} \quad f'_+(t) \le k_f \ \text{ for every } t \in (m_Y, M_Y). \end{cases} \tag{5.13}$$

If f is concave, then the reverse inequality is valid in (5.12) *with constant* $\bar{c} := \min_{m_Y \le t \le M_Y} \{k_f t + l_f - f(t)\}$ *can be determined as in the right side in* (5.13) *with reverse inequality signs.*

Remark 5.1.3.: If f is a strictly convex differentiable function on $[m_Y, M_Y]$, then

$$\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right) \leqslant \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij}\Phi_j(X_j)\right) + \left(k_f\, t_0 + l_f - f(t_0)\right)\|\mathbf{b}\|_1 1_K,$$

where $t_0 = m_Y$ if $f'(m_Y) \ge k_f$, $t_0 = M_Y$ if $f'(M_Y) \le k_f$ and $t_0 = f'^{-1}(k_f)$ if $f'(m_Y) \le k_f \le f'(M_Y)$.

Now, applying Corollary 5.1.4. on the identity mappings, we obtain complements to the inequalities (5.6) and (5.7).

Corollary 5.1.5.: *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_l)$ *be* $l-$*tuple and* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be* $n-$*tuple of self-adjoint operators in* $\mathcal{B}_h(H)$ *with spectra of* X_j *in* $[m, M]$, $m < M$ *and* m_i *and* M_i, $m_i \le M_i$ *are the bounds of* Y_i, $\mathbf{a} = (a_1, a_2, \ldots, a_l) \in [0,\infty)^l$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0,\infty)^n$ *be vectors and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nl}(\mathbb{R})$ *be a row stochastic matrix such that* $\langle \mathbf{a}, \mathbf{Y}\rangle \prec_\circ$

$\langle \mathbf{b}, \mathbf{X}\rangle$. *If* $f \in \mathcal{C}([m, M])$ *is convex, then*

$$\sum_{j=1}^{l} a_j f(X_j) \leqslant \sum_{i=1}^{n} b_i f\left(Y_i\right) + \bar{C}\|\mathbf{b}\|_1 1_K, \tag{5.14}$$

where the constant $\bar{C}$ *is defined by* (5.13) *with* $m_Y = \min_{1\leq i\leq n}\{m_i\}$ *and* $M_Y = \max_{1\leq i\leq n}\{M_i\}$.

Especially, if $l = n$ *and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ *is a doubly stochastic matrix, then the weighted multidimensional majorization inequality holds, so that*

$$\mathbf{Y} \prec_{\circ} \mathbf{X} \qquad \text{implies} \qquad \sum_{i=1}^{n} a_i f(X_i) \leqslant \sum_{i=1}^{n} a_i f\left(Y_i\right) + \bar{C}\|\mathbf{a}\|_1 1_K. \tag{5.15}$$

If f *is concave, the reverse inequality is valid in* (5.14) *and* (5.15), *with constant* $\bar{c} := \min_{m_Y \leq t\leq M_Y}\{k_f t + l_f - f(t)\}$ *can be determined as in the right side in* (5.13) *with reverse inequality signs.*

Finally, setting $l = n$, $\mathbf{b} = \mathbf{a} = (1, 1, \ldots, 1)$ in Corollary 5.1.4., we obtain complements to (5.8).

Corollary 5.1.6.: *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *and* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be* $n-$*tuples of self-adjoint operators in* $\mathcal{B}_h(H)$ *with spectra of* X_j *in* $[m, M]$, $m < M$ *and* m_i *and* M_i, $m_i \leq M_i$ *are the bounds of* Y_i. *Let* $(\Phi_1, \Phi_2, \ldots, \Phi_n)$ *of* $n-$*tuple positive linear mappings* $\Phi_j : \mathcal{B}(H) \to \mathcal{B}(K)$, $\mathbf{\Phi}(\mathbf{X}) = (\Phi_1(X_1), \Phi_2(X_2), \ldots, \Phi_n(X_n))$ *and* $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ *be a doubly stochastic matrix such that* $\mathbf{Y} \prec_{\circ} \mathbf{\Phi}(\mathbf{X})$.

If $f \in \mathcal{C}(J)$ *is convex and one of the conditions (a) or (b) in Corollary 5.1.2. is true, then*

$$\sum_{i=1}^{n} \Phi_i\left(f(X_i)\right) \leqslant \sum_{i=1}^{n} f\left(Y_i\right) + n\cdot \bar{C}\; 1_K \tag{5.16}$$

where the constant $\bar{C}$ *is defined by* (5.13) *with* $m_Y = \min_{1\leq i\leq n}\{m_i\}$ *and* $M_Y = \max_{1\leq i\leq n}\{M_i\}$.

If f *is concave, then the reverse inequalities are valid in* (5.16), *with constant* $\bar{c} := \min_{m_Y \leq t\leq M_Y}\{k_f t + l_f - f(t)\}$ *can be determined as in the right side in* (5.13) *with reverse inequality signs.*

5.1.3. Ratio type complementary inequalities

In this subsection we consider the ratio type complements to the inequality (5.5).

Corollary 5.1.7.: *Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$, m, M and m_Y, M_Y be as in Theorem 5.1.3 and let one of the conditions (i) or (ii) in Theorem 5.1.2 be true.*

If $f \in \mathcal{C}([m, M])$ is convex and $g \in \mathcal{C}([m_Y, M_Y])$ is strictly positive, then

$$\sum_{j=1}^{l} a_j \Phi_j \left(f(X_j)\right) \leqslant \max_{m_Y \le t \le M_Y} \left\{ \frac{k_f t + l_f}{g\left(t\right)} \right\} \sum_{i=1}^{n} b_i g\left(\sum_{j=1}^{l} s_{ij} \Phi_j(X_j) \right). \tag{5.17}$$

If f is concave, then the reverse inequality is valid in (5.17) *with* min *instead of* max.

Proof: We only prove the cases when f is convex. We choose α such that $\beta = 0$ in Corollary 5.1.3., where $\beta := \max\limits_{m_Y \le t \le M_Y} \{k_f t + l_f - \alpha g\left(t\right)\}$. Then $k_f t + l_f - \alpha g\left(t\right) \le \beta = 0$ for all $t \in [m_Y, M_Y]$. Since $g > 0$ it follows $\alpha \ge \frac{k_f t + l_f}{g(t)}$ for all $t \in [m_Y, M_Y]$, so $\alpha \ge \max\limits_{m_Y \le t \le M_Y} \left\{ \frac{k_f t + l_f}{g(t)} \right\}$. Because a function $t \mapsto k_f t + l_f - \alpha g\left(t\right)$ is continuous on $[m_Y, M_Y]$, then there is exactly one point $t_0 \in [m_Y, M_Y]$ which achieves the global maximum: $k_f t_0 + l_f - \alpha g\left(t_0\right) = 0$. It follows that $\alpha = \frac{k_f t_0 + l_f}{g(t_0)}$. So $\alpha = \max\limits_{m_Y \le t \le M_Y} \left\{ \frac{k_f t + l_f}{g(t)} \right\}$. □

Remark 5.1.4.: Now we give a way of determining the bound $K := \max\limits_{m_Y \le t \le M_Y} \left\{ \frac{k_f t + l_f}{g(t)} \right\}$ placed in Corollary 5.1.7. (see [109, Corollary 4.3]). Let f is convex.

If g is concave, then

$$K = \max \left\{ \frac{f^{cho}_{[m,M]}(m_Y)}{g(m_Y)}, \frac{f^{cho}_{[m,M]}(M_Y)}{g(M_Y)} \right\}.$$

But, if g is convex, then

$$K = \begin{cases} \frac{f^{cho}_{[m,M]}(m_Y)}{g(m_Y)} & \text{if } \; g'_{-}(t) \ge \frac{k_f g(t)}{k_f t + l_f} \; \text{ for every } t \in (m_Y, M_Y), \\ \frac{f^{cho}_{[m,M]}(t_0)}{g(t_0)} & \text{if } \; g'_{-}(t_0) \le \frac{k_f g(t_0)}{k_f t_0 + l_f} \le g'_{+}(t_0) \; \text{ for some } t_0 \in (m_Y, M_Y), \\ \frac{f^{cho}_{[m,M]}(M_Y)}{g(M_Y)} & \text{if } \; g'_{+}(t) \le \frac{k_f g(t)}{k_f t + l_f} \; \text{ for every } t \in (m_Y, M_Y). \end{cases} \tag{5.18}$$

Putting $g \equiv f$ in Corollary 5.1.7. we obtain the following result.

Corollary 5.1.8.: *Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$, m, M and m_Y, M_Y be as in Theorem 5.1.3. Let $f \in \mathcal{C}([m,M])$ be strictly positive on $[m_Y, M_Y]$ and let one of the conditions (i) or (ii) in Theorem 5.1.2 be true. If f is convex, then*

$$\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right) \leqslant \bar{K} \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij} \Phi_j(X_j)\right), \tag{5.19}$$

where the value of the constant
$\bar{K} := \max\limits_{m_Y \le t \le M_Y} \left\{\frac{k_f t + l_f}{f(t)}\right\}$ *determined as in* (5.18) *with* $g = f$, *i.e.*

$$\bar{K} = \begin{cases} \frac{f^{cho}_{[m,M]}(m_Y)}{f(m_Y)} & \text{if } \; f'_-(t) \ge \frac{k_f f(t)}{k_f t + l_f} \; \text{ for every } t \in (m_Y, M_Y), \\ \frac{f^{cho}_{[m,M]}(t_0)}{f(t_0)} & \text{if } \; f'_-(t_0) \le \frac{k_f f(t_0)}{k_f t_0 + l_f} \le f'_+(t_0) \; \text{ for some } t_0 \in (m_Y, M_Y), \\ \frac{f^{cho}_{[m,M]}(M_Y)}{f(M_Y)} & \text{if } \; f'_+(t) \le \frac{k_f f(t)}{k_f t + l_f} \; \text{ for every } t \in (m_Y, M_Y). \end{cases} \tag{5.20}$$

But, if f is concave, then the reverse inequality is valid in (5.19) *with constant* $\bar{k} := \min\limits_{m_Y \le t \le M_Y} \left\{\frac{k_f t + l_f}{f(t)}\right\}$ *can be determined as in the right side in* (5.20) *with reverse inequality signs.*

Remark 5.1.5.: Let X_j, Φ_j, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{S}$, m, M and m_Y, M_Y be as in Theorem 5.1.3. Let $f\colon [m,M] \to \mathbb{R}$ be a continuous function and $f(m), f(M) > 0$. If f is strictly positive and strictly convex twice differentiable on $[m_Y, M_Y]$, then

$$\sum_{j=1}^{l} a_j \Phi_j\left(f(X_j)\right) \leqslant \frac{k_f\, t_0 + l_f}{f(t_0)} \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij} \Phi_j(X_j)\right),$$

where $t_0 \in (m_Y, M_Y)$ is defined as the unique solution of $k_f f(t) = (k_f t + l_f) f'(t)$ provided $(k_f m_Y + l_f) f'(m_Y)/f(m_Y) \le k_f \le (k_f M_Y + l_f) f'(M_Y)/f(M_Y)$, otherwise t_0 is defined as m_Y or M_Y provided $k_f \le (k_f m_Y + l_f) f'(m_Y)/f(m_Y)$ or $k_f \ge (k_f M_Y + l_f) f'(M_Y)/f(M_Y)$, respectively.

Now, applying Corollary 5.1.8. on the identity mappings, we obtain another complements to (5.6) and (5.7).

Corollary 5.1.9.: *Let $\mathbf{X} = (X_1, X_2, \ldots, X_l)$ be $l-$tuple and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ be*

$n-$tuple of self-adjoint operators in $\mathcal{B}_h(H)$ with spectra of X_j in $[m,M]$, $m<M$ and m_i and M_i, $m_i \leq M_i$ are the bounds of Y_i, and $m_Y = \min_{1\leq i\leq n}\{m_i\}$, $M_Y = \max_{1\leq i\leq n}\{M_i\}$. Let $\mathbf{a} = (a_1,a_2,\dots,a_l) \in [0,\infty)^l$, $\mathbf{b} = (b_1,b_2,\dots,b_n) \in [0,\infty)^n$ be vectors and $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nl}(\mathbb{R})$ be a row stochastic matrix such that $\langle \mathbf{a}, \mathbf{Y}\rangle \prec_\circ \langle \mathbf{b}, \mathbf{X}\rangle$.

If $f \in \mathcal{C}([m,M])$ is convex and strictly positive on $[m_Y, M_Y]$, then

$$\sum_{j=1}^{l} a_j f(X_j) \leqslant \bar{K} \sum_{i=1}^{n} b_i f(Y_i), \tag{5.21}$$

where the constant $\bar{K}$ is defined by (5.20).

Especially, if $l=n$ and $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ is a doubly stochastic matrix, then the weighted multidimensional majorization inequality holds, so that

$$\mathbf{Y} \prec_\circ \mathbf{X} \qquad \textit{implies} \qquad \sum_{i=1}^{n} a_i f(X_i) \leqslant \bar{K} \sum_{i=1}^{n} a_i f(Y_i). \tag{5.22}$$

If f is concave, the reverse inequality is valid in (5.21) *and* (5.22) *with constant $\bar{k} := \min_{m_Y\leq t\leq M_Y}\left\{\frac{k_f t + l_f}{f(t)}\right\}$ can be determined as in the right side in* (5.20) *with reverse inequality signs.*

Finally, setting $l=n$, $\mathbf{b} = \mathbf{a} = (1,1,\dots,1)$ in Corollary 5.1.4. we obtain complements to the inequality (5.8).

Corollary 5.1.10.: *Let $\mathbf{X} = (X_1,X_2,\dots,X_n)$ and $\mathbf{Y} = (Y_1,Y_2,\dots,Y_n)$ be $n-$tuples of self-adjoint operators in $\mathcal{B}_h(H)$ with spectra of X_j in $[m,M]$, $m<M$ and m_i and M_i, $m_i \leq M_i$ are the bounds of Y_i, and $m_Y = \min_{1\leq i\leq n}\{m_i\}$, $M_Y = \max_{1\leq i\leq n}\{M_i\}$. Let $(\Phi_1,\Phi_2,\dots,\Phi_n)$ of $n-$tuple positive linear mappings $\Phi_j : \mathcal{B}(H) \to \mathcal{B}(K)$, $\mathbf{\Phi}(\mathbf{X}) = (\Phi_1(X_1),\Phi_2(X_2),\dots,\Phi_n(X_n))$ and $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ be a doubly stochastic matrix such that $\mathbf{Y} \prec_\circ \mathbf{\Phi}(\mathbf{X})$.*

If $f \in \mathcal{C}(J)$ is convex and strictly positive on $[m_Y, M_Y]$ and if one of the conditions (a) or (b) in Corollary 5.1.2. is true, then

$$\sum_{i=1}^{n} \Phi_i(f(X_i)) \leqslant \bar{K} \sum_{i=1}^{n} f(Y_i), \tag{5.23}$$

where the constant $\bar{K}$ is defined by is defined by (5.20).

If f is concave, then the reverse inequalities are valid in (5.23) *with constant* $\bar{k} := \min_{m_Y \le t \le M_Y} \left\{ \frac{k_f t + l_f}{f(t)} \right\}$ *can be determined as in the right side in* (5.20) *with reverse inequality signs.*

5.1.4. Sherman's operator and its properties

In this subsection we define Sherman's operator, deduced from Sherman's operator inequality (5.6).

Let $f \in \mathcal{C}(J)$ be an operator convex function and let $\mathcal{F}_o(J)$ denote the set of all operator convex functions on interval J. Let $\mathcal{B}_h^J(H)$ denotes the convex set of bounded self-adjoint operators on the Hilbert space H with spectra in J. The set of all $n \times l$ row stochastic matrices is denoted by $\mathbb{S}_{nl}(\mathbb{R})$.

We define **Sherman's operator** $\mathcal{S} : \mathcal{F}_o(J) \times [\mathcal{B}_h^J(H)]^l \times [0, \infty)^n \times \mathbb{S}_{nl}(\mathbb{R}) \to \mathcal{B}^+(H)$ as

$$\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) = \sum_{j=1}^{l} \sum_{i=1}^{n} b_i s_{ij} f(X_j) - \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij} X_j \right), \tag{5.24}$$

where $\mathbf{X} = (X_1, X_2, \ldots, X_l)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ and $\mathbf{S} = (s_{ij})$.

Note that operator $\mathcal{S}$ is well-defined. Namely, positivity of operator $\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S})$ follows from Sherman's operator inequality (5.6).

Now, we state and prove our first result that shows the properties of concavity and monotonicity of Sherman's operator.

Theorem 5.1.4: *Suppose $\mathcal{S}$ is an operator defined by* (5.24)*. Then it satisfies the following properties:*

(i) $\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \cdot)$ *is concave on* $\mathbb{S}_{nl}(\mathbb{R})$*, that is*

$$\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \lambda \mathbf{S} + (1 - \lambda)\ \mathbf{T}) \ge \lambda\, \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) + (1 - \lambda)\, \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{T}). \tag{5.25}$$

(ii) If $\mathbf{b}, \mathbf{c} \in [0, \infty)^n$ *with* $\mathbf{b} \ge \mathbf{c}$ *(i.e.* $b_i \ge c_i$, $i = 1, 2, \ldots, n$*), then*

$$\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \ge \mathcal{S}(f,\ \mathbf{X}, \mathbf{c}, \mathbf{S}) \ge 0, \tag{5.26}$$

i.e. $\mathcal{S}(f, \mathbf{X}, \cdot, \mathbf{S})$ *is increasing on* $[0, \infty)^n$.

But, if $f \in \mathcal{C}(J)$ *is operator concave, then* $-\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \in \mathcal{B}^+(H)$, $\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \cdot)$ *is convex on* $\mathbb{S}_{nl}(\mathbb{R})$ *and* $\mathcal{S}(f, \mathbf{X}, \cdot, \mathbf{S})$ *is decreasing on* $[0, \infty)^n$.

Proof: (i) We start with Sherman's operator (5.24) equipped with a $\lambda\mathbf{S} + (1-\lambda)\mathbf{T}$. Clearly, $\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \lambda\mathbf{S} + (1-\lambda)\mathbf{T})$ can be rewritten in the following form:

$$\begin{aligned}
&\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \lambda\mathbf{S} + (1-\lambda)\ \mathbf{T}) \\
&\quad = \sum_{j=1}^{l}\sum_{i=1}^{n} b_i(\lambda s_{ij} + (1-\lambda)t_{ij}) f(X_j) - \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l}(\lambda s_{ij} + (1-\lambda)t_{ij})X_j\right) \\
&\quad = \lambda \sum_{j=1}^{l}\sum_{i=1}^{n} b_i s_{ij} f(X_j) + (1-\lambda)\sum_{j=1}^{l}\sum_{i=1}^{n} b_i t_{ij} f(X_j) \\
&\qquad - \sum_{i=1}^{n} b_i f\left(\lambda \sum_{j=1}^{l} s_{ij}X_j + (1-\lambda)\sum_{j=1}^{l} t_{ij}X_j\right).
\end{aligned} \tag{5.27}$$

On the other hand, operator convexity of the function f and non-negativity of weights provide inequality

$$\begin{aligned}
&\sum_{i=1}^{n} b_i f\left(\lambda \sum_{j=1}^{l} s_{ij}X_j + (1-\lambda)\sum_{j=1}^{l} t_{ij}X_j\right) \\
&\qquad \leq \lambda \sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} s_{ij}X_j\right) + (1-\lambda)\sum_{i=1}^{n} b_i f\left(\sum_{j=1}^{l} t_{ij}X_j\right).
\end{aligned} \tag{5.28}$$

Now, considering (5.27) and (5.28), we get operator concave property (5.25), due to definition (5.24) .

(ii) First, we note that $\mathcal{S}(f, \mathbf{X}, \cdot, \mathbf{S})$ is linear, that is

$$\mathcal{S}(f, \mathbf{X}, \alpha\mathbf{b} + \beta\mathbf{c}, \mathbf{S}) = \alpha\,\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) + \beta\,\mathcal{S}(f, \mathbf{X}, \mathbf{c}, \mathbf{S}) \tag{5.29}$$

for every $\alpha, \beta \in \mathbb{R}$.

If $\mathbf{b} = \mathbf{c}$, then relation (5.26) holds trivially. If $\mathbf{b} > \mathbf{c}$, then $\mathbf{b} \in [0, \infty)^n$ can be represented as a sum $\mathbf{b} = (\mathbf{b} - \mathbf{c}) + \mathbf{c}$, where $\mathbf{b} - \mathbf{c}, \mathbf{c} \in [0, \infty)^n$. So, from linear property

(5.29) and positive property $\mathcal{S}(f, \mathbf{X}, \cdot, \mathbf{S}) \geq 0$ we get

$$\begin{aligned}\mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) &= \mathcal{S}(f, \mathbf{X}, (\mathbf{b}-\mathbf{c})+\mathbf{c}, \mathbf{S}) \\ &= \mathcal{S}(f, \mathbf{X}, \mathbf{b}-\mathbf{c}, \mathbf{S})+\mathcal{S}(f, \mathbf{X}, \mathbf{c}, \mathbf{S}) \\ &\geq \mathcal{S}(f, \mathbf{X}, \mathbf{c}, \mathbf{S}) \geq 0,\end{aligned}$$

and consequently (5.26) holds. □

The concavity and monotonic properties of Sherman's operator are very important, because of the numerous applications that will follow from them. First, regarding the monotonicity property (5.26), we give the consequence of Theorem 5.1.4, which includes the lower and upper bound for Sherman's operator, which are expressed in terms of associated non-weighted Jensen's operator.

Theorem 5.1.5: *Suppose $\mathcal{S}$ is an operator defined by* (5.24)*. Then*

$$0 \leq \min_{\substack{1\leq i\leq n\\ 1\leq j\leq l}} \{b_i s_{ij}\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}) \leq \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \leq \max_{\substack{1\leq i\leq n\\ 1\leq j\leq l}} \{b_i s_{ij}\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}), \tag{5.30}$$

where $\mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl})$ is non-weighted Sherman's operator

$$\mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}) = \sum_{j=1}^{l}\sum_{i=1}^{n} f(X_j) - l\sum_{i=1}^{n} f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right) = n\cdot \mathcal{J}_{\mathcal{N}}(f, \mathbf{X}),$$

and $\mathcal{J}_{\mathcal{N}}(f, \mathbf{X})$ is non-weighted Jensen's operator

$$\mathcal{J}_{\mathcal{N}}(f, \mathbf{X}) = \sum_{j=1}^{l} f(X_j) - l\, f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right). \tag{5.31}$$

Proof:

- First let us prove that

$$\min_{1\leq i\leq n}\{b_i\}\ \mathcal{S}(f, \mathbf{X},\ \mathbf{1}_n, \mathbf{S}) \leq \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \leq \max_{1\leq i\leq n}\{b_i\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, \mathbf{S}) \tag{5.32}$$

holds, where $\mathbf{1}_n = (1, 1, \ldots, 1)$ is constant ordered n-tuple, that is $\mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, \mathbf{S}) = \sum_{j=1}^{l}\sum_{i=1}^{n} s_{ij} f(X_j) -$

$\sum\limits_{i=1}^{n} f\left(\sum\limits_{j=1}^{l} s_{ij}X_j\right)$.

We compare an ordered n-tuple $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0,\infty)^n$ with constant ordered n-tuples $\mathbf{b}_{\min} = \left(\min\limits_{1\le i\le n}\{b_i\}, \ldots, \min\limits_{1\le i\le n}\{b_i\}\right)$ and $\mathbf{b}_{\max} = \left(\max\limits_{1\le i\le n}\{b_i\}, \ldots, \max\limits_{1\le i\le n}\{b_i\}\right)$. Clearly, $\mathbf{b}_{\max} \ge \mathbf{b} \ge \mathbf{b}_{\min}$. Hence use by the monotonicity property (5.26) of $\mathcal{S}(f, \mathbf{X}, \cdot, \mathbf{S})$ yields series of inequalities:

$$\mathcal{S}(f, \mathbf{X}, \mathbf{b}_{\min}, \mathbf{S}) \le \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \le \mathcal{S}(f, \mathbf{X}, \mathbf{b}_{\max}, \mathbf{S}).$$

Finally, considering $\mathcal{S}(f, \mathbf{X}, \mathbf{b}_{\min}, \mathbf{S}) = \min\limits_{1\le i\le n}\{b_i\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, \mathbf{S})$ and $\mathcal{S}(f, \mathbf{X}, \mathbf{b}_{\max}, \mathbf{S}) = \max\limits_{1\le i\le n}\{b_i\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, \mathbf{S})$ we get the series of inequalities in (5.32).

- Next, let us prove that

$$\begin{aligned} m_i\left(\sum_{j=1}^{l} f(X_j) - l\, f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right)\right) &\le \sum_{j=1}^{l} s_{ij} f(X_j) - f\left(\sum_{j=1}^{l} s_{ij}X_j\right) \\ &\le M_i\left(\sum_{j=1}^{l} f(X_j) - l\, f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right)\right) \end{aligned} \tag{5.33}$$

holds for every $i = 1, 2, \ldots, n$, where $m_i = \min\limits_{1\le j\le l}\{s_{ij}\}$ and $M_i = \max\limits_{1\le j\le l}\{s_{ij}\}$.
Since $l \cdot m_i + \sum\limits_{j=1}^{l}(s_{ij} - m_i) = 1$, we have

$$\begin{aligned} & f\left(\sum_{j=1}^{l} s_{ij}X_j\right) = f\left(l\cdot m_i\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right) + \sum_{j=1}^{l}(s_{ij} - m_i)\, X_j\right) \\ \le\ & l\cdot m_i f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right) + \sum_{j=1}^{l}(s_{ij} - m_i)\, f(X_j) \qquad \text{(by Jensen's operator inequality)} \\ =\ & \sum_{j=1}^{l} s_{ij} f(X_j) - m_i\left(\sum_{j=1}^{l} f(X_j) - l\, f\left(\frac{1}{l}\sum_{j=1}^{l} X_j\right)\right), \end{aligned}$$

which give the LHS inequality in (5.33).

Also, since $M_i > 0$ $(i = 1, 2, \ldots, n)$ and $\frac{1}{M_i}\sum_{j=1}^{l}\left(\frac{M_i}{l} - \frac{s_{ij}}{l}\right) + \frac{1}{lM_i} = 1$, we have

$$\begin{aligned}
& f\left(\tfrac{1}{l}\sum_{j=1}^{l} X_j\right) = f\left(\tfrac{1}{M_i}\sum_{j=1}^{l}\left(\tfrac{M_i}{l} - \tfrac{s_{ij}}{l}\right) X_j + \tfrac{1}{lM_i}\left(\sum_{j=1}^{l} s_{ij}X_j\right)\right) \\
\leq\ & \tfrac{1}{M_i}\sum_{j=1}^{l}\left(\tfrac{M_i}{l} - \tfrac{s_{ij}}{l}\right) f(X_j) + \tfrac{1}{lM_i} f\left(\sum_{j=1}^{l} s_{ij}X_j\right) \qquad \text{(by Jensen's operator inequality)} \\
=\ & \tfrac{1}{l}\sum_{j=1}^{l} f(X_j) - \tfrac{1}{lM_i}\left(\sum_{j=1}^{l} s_{ij} f(X_j) - f\left(\sum_{j=1}^{l} s_{ij}X_j\right)\right),
\end{aligned}$$

which give the RHS inequality in (5.33).

- Finally we prove (5.30).

Multiplying the series of inequalities in (5.33) with $b_i \geq 0$ and summing over $i = 1, ..., n$, we get

$$\begin{aligned}
\sum_{i=1}^{n} b_i m_i\left(\sum_{j=1}^{l} f(X_j) - l f\left(\tfrac{1}{l}\sum_{j=1}^{l} X_j\right)\right) \ &\leq\ \sum_{i=1}^{n} b_i\left(\sum_{j=1}^{l} s_{ij} f(X_j) - f\left(\sum_{j=1}^{l} s_{ij}X_j\right)\right) \\
&\leq\ \sum_{i=1}^{n} b_i M_i\left(\sum_{j=1}^{l} f(X_j) - l f\left(\tfrac{1}{l}\sum_{j=1}^{l} X_j\right)\right)
\end{aligned}$$

i.e.

$$\mathcal{S}(f, \mathbf{X}, \underline{\mathbf{b}}, (1/l)_{nl}) \leq \mathcal{S}(f, \mathbf{X}, \mathbf{b}, \mathbf{S}) \leq \mathcal{S}(f, \mathbf{X}, \overline{\mathbf{b}}, (1/l)_{nl}), \tag{5.34}$$

where $\underline{\mathbf{b}} = (b_1m_1, b_2m_2, \ldots, b_nm_n)$, $\overline{\mathbf{b}} = (b_1M_1, b_2M_2, \ldots, b_nM_n)$ and $(1/l)_{nl} \in \mathbb{S}_{nl}(\mathbb{R})$ is row stochastic matrix with constant entries. By using the LHS and the RHS inequality in (5.32), we obtain

$$\min_{1\leq i\leq n}\{b_im_i\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}) \leq \mathcal{S}(f, \mathbf{X}, \underline{\mathbf{b}}, (1/l)_{nl}), \tag{5.35}$$

$$\mathcal{S}(f, \mathbf{X}, \overline{\mathbf{b}}, (1/l)_{nl}) \leq \max_{1\leq i\leq n}\{b_im_i\}\ \mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}). \tag{5.36}$$

Finally, we get (5.30) by combining (5.34), (5.35) and (5.36) and using that

$$\min_{1\leq i\leq n}\{b_im_i\} = \min_{\substack{1\leq i\leq n\\ 1\leq j\leq l}}\{b_is_{ij}\}, \qquad \max_{1\leq i\leq n}\{b_im_i\} = \max_{\substack{1\leq i\leq n\\ 1\leq j\leq l}}\{b_is_{ij}\},$$
$$\mathcal{S}(f, \mathbf{X}, \mathbf{1}_n, (1/l)_{nl}) = n\cdot\mathcal{J}_{\mathcal{N}}(f, \mathbf{X}).$$

□

Remark 5.1.6.: 1) Dragomir et al. in [42] investigated the properties of discrete Jensen's functional

$$J_l(f,\mathbf{x},\mathbf{p})=\sum_{i=1}^{l}p_if(x_i)-Pf\left(\frac{\sum_{i=1}^{l}p_if(x_i)}{P}\right), \tag{5.37}$$

where $f:J\subset\mathbb{R}\to\mathbb{R}$ is a convex function, $\mathbf{x}=(x_1,x_2,\ldots,x_l)\in J^l$ and $\mathbf{p}=(p_1,p_2,\ldots,p_l)$ is positive l-tuple of real numbers with $P=\sum_{i=1}^{l}p_i$.

Setting $n=1$, $b_1=P$, $s_{1j}=\frac{p_j}{P}$ $(j=1,2,\ldots,l)$ in Sherman's operator (5.24) we obtain Jensen's operator that corresponds to (5.37):

$$\mathcal{J}_l(f,\mathbf{X},\mathbf{p})=\sum_{i=1}^{l}p_if(X_i)-Pf\left(\sum_{i=1}^{l}\frac{p_i}{P}f(X_i)\right), \tag{5.38}$$

where $f\in\mathcal{F}_o(J)$ and $\mathbf{X}=(X_1,X_2,\ldots,X_l)\in[\mathcal{B}_h^J(H)]^l$.

Applying Theorem 5.1.4, we obtain that $\mathcal{J}_l$ has the properties of superadditivity and monotonicity, since putting $\lambda=\frac{1}{2}$ in (5.25) we obtain

$$\mathcal{J}_l(f,\mathbf{X},\mathbf{p}+\mathbf{q})\geq\mathcal{J}_l(f,\mathbf{X},\mathbf{p})+\mathcal{J}_l(f,\mathbf{X},\mathbf{q}).$$

Also, applying Theorem 5.1.5 we obtain the bounds of $\mathcal{J}_l$, which are expressed in terms of non-weighted Jensen's operator (5.31):

$$0\leq\min_{1\leq j\leq l}\{p_j\}\ \mathcal{J}_{\mathcal{N}}(f,\mathbf{X})\leq\mathcal{J}_l(f,\mathbf{X},\mathbf{p})\leq\max_{1\leq j\leq l}\{p_j\}\ \mathcal{J}_{\mathcal{N}}(f,\mathbf{X}). \tag{5.39}$$

2) Setting $l=2$ and $X_1=D$, $X_2=\delta 1_H$ in Jensen's operator $\mathcal{J}_l$ defined by (5.38), we obtain **Jensen's operator** consider in [93, (14)]:

$$\mathcal{J}_2(f,D,\delta,\mathbf{p})=p_1f(D)+p_2(\delta)1_H-(p_1+p_2)f\left(\frac{p_1D+p_2\delta 1_H}{p_1+p_2}\right),$$

where $\mathbf{p}=(p_1,p_2)$, $a1_H\leq D\leq b1_H$, $a\leq\delta\leq b$ and $f\in\mathcal{F}([a,b])$. We remark that f is not operator convex in this case, because $\mathcal{J}_2(f,D,\delta,\mathbf{p})$ we can obtain directly from Jensen's functional $J_2(f,\mathbf{x},\mathbf{p})$ defined by (5.37). We obtain that $\mathcal{J}_2$ has the properties of superadditivity and monotonicity as immediately proved in [93, Theorem 1]. Also, we obtained that

$$0 \leq \min\{p_1, p_2\} \mathcal{J}_{\mathcal{N}}(f, D, \delta) \leq \mathcal{J}_2(f, D, \delta, \mathbf{p}) \leq \max\{p_1, p_2\} \mathcal{J}_{\mathcal{N}}(f, D, \delta)$$

holds where $\mathcal{J}_{\mathcal{N}}(f, D, \delta) = f(D) + f(\delta)1_H - 2 \cdot f\left(\frac{D+\delta 1_H}{2}\right)$, see [93, Corollary 1].

The obtained results we can applied to operator means. E.g. we can get refinements and conversions of numerous mean inequalities for Hilbert space operators, see [93, §4].

5.1.5. Multidimensional Sherman's operator and applications

The main objective of this subsection is to obtain an unified treatment for the examples of joint concave mappings. With that intention we are going to extend the concept of Sherman's operator to several variables.

Let $k \in \mathbb{N}$ and H_j be the Hilbert space ($j = 1, 2, \ldots, k$). Suppose $A_j \in \mathcal{B}_h(H_j)$ ($j = 1, 2, \ldots, k$) and let $\mathbf{A} = (A_1, A_2, \ldots, A_k)$ denotes a k-tuple of self-adjoint operators. Let $\mathcal{D} \subseteq \prod_{j=1}^{k} \mathcal{B}_h(H_j)$ be a convex set. Further, let F be a mapping that every $\mathbf{A} \in \mathcal{D}$ assigns the self-adjoint operator on a Hilbert space H', that is, we suppose the function $F : \mathcal{D} \to \mathcal{B}_h(H')$ is well-defined.

With the above assumptions, we can establish the definition of operator convexity in multiple variables. A function $F : \mathcal{D} \to \mathcal{B}_h(H')$ is said to be **operator convex in k variables**, if the operator inequality

$$F(\lambda \mathbf{A} + (1-\lambda)\mathbf{B}) \leqslant \lambda F(\mathbf{A}) + (1-\lambda)F(\mathbf{B}) \tag{5.40}$$

holds for all $\mathbf{A}, \mathbf{B} \in \mathcal{D}$ and $\lambda \in [0, 1]$, with respect to operator order in H'. A function F is called **operator concave in k variables**, if the reverse inequality holds in (5.40) for all $\mathbf{A}, \mathbf{B} \in \mathcal{D}$ and $\lambda \in [0, 1]$.

Similarly as by the classical Jensen's inequality, inequality (5.40) can be easily extended to **multidimensional Jensen's operator inequality** (see [94, Proposition 1]) as follows: If $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is n-tuple of nonnegative real numbers, $\sum_{i=1}^{n} a_i = 1$, and $\mathbf{X}_i \in \mathcal{D}$

$(i = 1, 2, \ldots, n)$ then
$$F\left(\sum_{i=1}^{n} a_i \mathbf{X}_i\right) \leqslant \sum_{i=1}^{n} a_i F(\mathbf{X}_i) \tag{5.41}$$
holds for any operator convex function in k variables, $F : \mathcal{D} \to \mathcal{B}_h(H')$. If F is an operator concave function in k variables, then the reverse inequality is valid in (5.41).

Now we extend Sherman's operator inequality (5.1.1.) to multidimensional operator inequality.

In the following, $\mathcal{D}$ denoted a convex set $\mathcal{D} \subseteq \prod_{j=1}^{k} \mathcal{B}_h(H_j)$.

Proposition 5.1.6 ***(Multidimensional Sherman's operator inequality)*****:** Let $(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_l)$ be $l-$tuple and $(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n)$ be $n-$tuple of operators $\mathbf{X}_j = (X_{j1}, X_{j2}, \ldots, X_{jk}) \in \mathcal{D}$ $(j = 1, 2, \ldots l)$, $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{ik}) \in \mathcal{D}$ $(i = 1, 2, \ldots n)$, let $\mathbf{a} = (a_1, a_2, \ldots, a_l) \in [0, \infty)^l$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0, \infty)^n$ be vectors and $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nl}(\mathbb{R})$ be a row stochastic matrix such that
$$\mathbf{Y}_i = \sum_{j=1}^{l} s_{ij} \mathbf{X}_j \ \ (i = 1, 2, \ldots, n) \quad \text{and} \quad a_j = \sum_{i=1}^{n} b_i s_{ij} \ \ (j = 1, 2, \ldots, l).$$
If $F : \mathcal{D} \to \mathcal{B}_h(H')$ is an operator convex function in k variables, then
$$\sum_{i=1}^{n} b_i F\left(\mathbf{Y}_i\right) \leqslant \sum_{j=1}^{l} a_j F(\mathbf{X}_j). \tag{5.42}$$

Especially, if $l = n$ and $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ is a doubly stochastic matrix, then the weighted multidimensional majorization inequality holds, so that
$$\mathbf{Y}_i = \sum_{j=1}^{n} s_{ij} \mathbf{X}_j \ \ (i = 1, 2, \ldots, n) \qquad \text{implies} \qquad \sum_{i=1}^{n} a_i F\left(\mathbf{Y}_i\right) \leqslant \sum_{i=1}^{n} a_i F(\mathbf{X}_\mathbf{i}). \tag{5.43}$$

If F is operator concave in k variables, then the reverse inequality is valid in (5.42) and (5.43).

Proof: Since F is an operator convex function in k variables and $\mathbf{S}$ is row stochastic, then

$$\sum_{i=1}^{n} b_i F\left(\mathbf{Y}_i\right) = \sum_{i=1}^{n} b_i F\left(\sum_{j=1}^{l} s_{ij}\mathbf{X}_j\right) \leqslant \sum_{i=1}^{n} b_i \left(\sum_{j=1}^{l} s_{ij} F(\mathbf{X}_j)\right)$$
$$= \sum_{j=1}^{l}\left(\sum_{i=1}^{n} b_i s_{ij}\right) F(\mathbf{X}_j) = \sum_{j=1}^{l} a_j F(\mathbf{X}_j),$$

which gives (5.42).

Setting $n = l$ in (5.42) and all weights a_i and b_j are equal and nonnegative, the condition $\mathbf{a} = \mathbf{bS}$ assures stochasticity on columns of $\mathbf{S}$, so in that case we deal with doubly stochastic matrix. Then (5.43) holds. □

Remark 5.1.7.: a) Setting $n = 1$ and $\mathbf{b} = (1)$ the multidimensional Sherman operator inequality (5.42) reduces to the multidimensional Jensen operator inequality (5.40).

b) Setting $\mathbf{a} = (1, 1, \ldots, 1)$ in the inequality (5.43), we get the multidimensional majorization inequality in the form:

$$\mathbf{Y}_i = \sum_{j=1}^{n} s_{ij}\mathbf{X}_j \;\; (i = 1, 2, \ldots, n) \qquad \text{implies} \qquad \sum_{i=1}^{n} F\left(\mathbf{Y}_i\right) \leqslant \sum_{i=1}^{n} F(\mathbf{X_i}),$$

where $\mathbf{S} = (s_{ij}) \in \mathcal{M}_{nn}(\mathbb{R})$ is a doubly stochastic matrix.

If F is operator concave in k variables, then the reverse inequality is valid in previous inequalities.

Now, we define multidimensional Sherman's operator, deduced from multidimensional Sherman's operator inequality.

Let $F : \mathcal{D} \to \mathcal{B}_h(H')$ is an operator convex function in k variables and let $\mathcal{F}_k(\mathcal{D})$ denote the set of all operator convex functions in k variables on $\mathcal{D}$.

Considering relation (5.42), we define **multidimensional Sherman's operator** $\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{S}, \mathbf{b}) : \mathcal{F}_k(\mathcal{D}) \times \mathcal{D} \times [0, \infty)^n \times \mathbb{S}_{nl}(\mathbb{R}) \to \mathcal{B}_h(H')$ as

$$\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S}) = \sum_{j=1}^{l}\sum_{i=1}^{n} b_i s_{ij} F(\mathbf{X}_j) - \sum_{i=1}^{n} b_i F\left(\sum_{j=1}^{l} s_{ij}\mathbf{X}_j\right), \tag{5.44}$$

where $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_l)$ is $l-$tuple of operators $\mathbf{X}_j = (X_{j1}, X_{j2}, \ldots, X_{jk}) \in \mathcal{D}$ $(j =$

$1, 2, \ldots, l)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ and $\mathbf{S} = (s_{ij})$.

Note that operator $\mathcal{S}_{\mathcal{M}}$ is well-defined and positive. Positivity of operator $\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S})$ follows from multidimensional Sherman's operator inequality (5.42). Now, we prove the properties of concavity and monotonicity of this operator.

Theorem 5.1.7: *Suppose that $\mathcal{S}_{\mathcal{M}}$ is an operator defined by* (5.44)*. Then it satisfies the following properties:*

(i) $\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \cdot)$ *is concave on* $\mathbb{S}_{nl}(\mathbb{R})$*, that is*

$$\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \lambda\mathbf{S} + (1-\lambda)\mathbf{T}) \geq \lambda\, \mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S}) + (1-\lambda)\, \mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{T}). \tag{5.45}$$

(ii) If $\mathbf{b}, \mathbf{c} \in [0, \infty)^n$ *with* $\mathbf{b} \geq \mathbf{c}$ *(i.e.* $b_i \geq c_i$, $i = 1, 2, \ldots, n$*), then*

$$\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S}) \geq \mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{c}, \mathbf{S}) \geq 0, \tag{5.46}$$

i.e. $\mathcal{S}_{\mathcal{M}}(f, \mathbf{X}, \cdot, \mathbf{S})$ *is increasing on* $[0, \infty)^n$*.*

But, if F is operator concave in k variables, then $-\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S}) \in \mathcal{B}^+(H')$, $\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \cdot)$ *is convex on* $\mathbb{S}_{nl}(\mathbb{R})$ *and* $\mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \cdot, \mathbf{S})$ *is decreasing on* $[0, \infty)^n$*.*

Proof: We use the same technique as in the proof of Theorem 5.1.4. We omit the details. □

The concavity and monotonic properties of multidimensional Sherman's operator are very significant properties, considering the numerous applications that will follow from them. As a first application of Theorem 5.1.7 we establish the lower and upper bounds for multidimensional operator (5.44), which are expressed in terms of associated non-weighted Jensen's operator.

Corollary 5.1.11.: *Suppose that $\mathcal{S}_{\mathcal{M}}$ is an operator defined by* (5.44)*. Then*

$$0 \leq n \cdot \min_{\substack{1\leq i\leq n \\ 1\leq j\leq l}} \{b_i s_{ij}\}\ \mathcal{J}_{\mathcal{N}}(F, \mathbf{X}) \leq \mathcal{S}_{\mathcal{M}}(F, \mathbf{X}, \mathbf{b}, \mathbf{S}) \leq n \cdot \max_{\substack{1\leq i\leq n \\ 1\leq j\leq l}} \{b_i s_{ij}\}\ \mathcal{J}_{\mathcal{N}}(F, \mathbf{X}), \tag{5.47}$$

where $\mathcal{J}_{\mathcal{N}}$ *is non-weighted Jensen's operator defined by* (5.31), *i.e.*

$$\mathcal{J}_{\mathcal{N}}(F,\mathbf{X}) = \sum_{j=1}^{l} F(\mathbf{X}_j) - lF\left(\sum_{j=1}^{l} \frac{1}{l}\mathbf{X}_j\right).$$

Proof: We use the same technique as in the proof of Theorem 5.1.5. We omit the details.

□

Remark 5.1.8.: 1) Setting $n = 1$, $b_1 = P$, $s_{1j} = \frac{p_j}{P}$ $(j = 1, 2, \ldots, l)$ in (5.44), we obtain Jensen's operator consider in [94, (6)]:

$$\mathcal{J}_{\mathcal{M}}(F,\mathbf{X},\mathbf{p}) = \sum_{j=1}^{l} p_j F(\mathbf{X}_j) - PF\left(\frac{1}{P}\sum_{j=1}^{l} p_j\mathbf{X}_j\right), \tag{5.48}$$

where $\mathbf{p} = (p_1, p_2, \ldots, p_l)$. By applying Theorem 5.1.7 and Corollary 5.1.11., we obtain properties and bounds of operator $\mathcal{J}_{\mathcal{M}}$ as in [94, Theorem 1, Corolary 1]. We omit the details.

2) Let $\mathcal{D} \subseteq \mathcal{B}_h(H_1) \times \mathcal{B}_h(H_2)$, $F : \mathcal{D} \to \mathcal{B}_h(H')$ be an operator function in two variable. Putting $\mathbf{X}_j = (A_j, B_j)$ $(j = 1, 2, \ldots, l)$ in Sherman's operator (5.44) we obtain

$$\mathcal{S}_{\mathcal{M}}(F,\mathbf{A},\mathbf{B},\mathbf{b},\mathbf{S}) = \sum_{j=1}^{l}\sum_{i=1}^{n} b_i s_{ij} F(A_j, B_j) - \sum_{i=1}^{n} b_i F\left(\sum_{j=1}^{l} s_{ij}A_j, \sum_{j=1}^{l} s_{ij}B_j\right), \tag{5.49}$$

where $\mathbf{A} = (A_1, A_2, \ldots, A_l)$, $\mathbf{B} = (B_1, B_2, \ldots, B_l)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0, \infty)^n$ and $\mathbf{S} = (s_{ij}) \in \mathbb{S}_{nl}(\mathbb{R})$. Also, Jensen's operator $\mathcal{J}_{\mathcal{M}}$ (5.48) reduces to (see [94, (12]):

$$\mathcal{J}_{\mathcal{M}_2}(F,\mathbf{A},\mathbf{B},\mathbf{p}) = \sum_{j=1}^{l} p_j F(A_j, B_j) - P \cdot F\left(\frac{1}{P}\sum_{j=1}^{l} p_jA_j, \frac{1}{P}\sum_{j=1}^{l} p_jB_j\right).$$

a) Replacing the function F with connection $\sigma : \mathcal{B}^+(H)\times\mathcal{B}^+(H) \to \mathcal{B}^+(H)$ in the operator (5.49) and using that every connection is positive homogeneous, i.e. $\alpha(A\sigma B) = (\alpha A)\sigma(\alpha B)$ for all $\alpha > 0$ and $A, B \in \mathcal{B}^+(H)$ (see [94]), then we obtain the operator

$$\mathcal{S}_{\sigma}(\mathbf{A},\mathbf{B},\mathbf{a}) = \sum_{j=1}^{l} a_j(A_j\sigma B_j) - \left(\sum_{j=1}^{l} a_jA_j\right)\sigma\left(\sum_{j=1}^{l} a_jB_j\right), \tag{5.50}$$

since $\mathbf{a} = \mathbf{bS}$. So, applying Theorem 5.1.7 and Corollary 5.1.11., we obtain results as in [94, §4.1]. Connections are the weighted arithmetic mean $A\nabla_s B$, the weighted harmonic mean $A!_s B$ and the weighted geometric mean $A\sharp_s B$ (see [54]). In these cases the connection operator (5.50) reduces to appropriate operators, see [94, Remark 3].

b) Replacing the function F with the solidarity $s : \mathcal{B}^{++}(H) \times \mathcal{B}^{++}(H) \to \mathcal{B}^{++}(H)$ in the operator (5.49) and using that every solidarity is also positive homogeneous, i.e. $\alpha(AsB) = (\alpha A)s(\alpha B)$ for all $\alpha > 0$ and $A, B \in \mathcal{B}^{++}(H)$ (see [50]), then we obtain the operator

$$\mathcal{S}_s(\mathbf{A}, \mathbf{B}, \mathbf{a}) = \sum_{j=1}^{l} a_j(A_j s B_j) - \left(\sum_{j=1}^{l} a_j A_j\right) s \left(\sum_{j=1}^{l} a_j B_j\right). \tag{5.51}$$

Applying Theorem 5.1.7 and Corollary 5.1.11., we obtain results as in [94, §4.2]. Examples of solidarity are the well-known relative operator entropy $S(A|B)$ and the Tsallis relative operator entropy $T_\lambda(A|B)$, see [52]. In these cases the solidarity operator (5.51) reduces to appropriate operators, see [94, Remark 4].

5.1.6. Multidimensional weighted geometric mean

Fujii et.al. in [49] established the weighted geometric mean $G[n,t]$, $0 \leq t \leq 1$, for an n-tuple of positive invertible operators $A_1, A_2, \ldots, A_n$. Let $G[2,t](A_1, A_2) = A_1 \sharp_t A_2$. For $n \geq 3$, $G[n,t]$ is defined inductively as follows: Define $A_i^{(1)} = A_i$ $(i = 1, 2, \cdots, n)$ and

$$A_i^{(r)} = G[n-1,t]\big(A_i^{(r-1)}, \ldots, A_{i-1}^{(r-1)}, A_{i+1}^{(r-1)}, \ldots, A_n^{(r-1)}\big),$$

inductively for r. Then, there exist the limit $\lim_{r\to\infty} A_i^{(r)}$ in the Thompson metric and it does not depend on i (see [49]). Thus, the above mentioned weighted geometric mean is defined as $G[n,t](A_1, A_2, \ldots, A_n) = \lim_{r\to\infty} A_i^{(r)}$. Such defined geometric mean is also jointly concave, i.e.

$$G[n,t]\left(\sum_{i=1}^{n} \lambda_i \mathbf{A}_i\right) \geq \sum_{i=1}^{n} \lambda_i G[n,t](\mathbf{A}_i), \tag{5.52}$$

where $\lambda_i \geq 0$, $\sum_{i=1}^{n} \lambda_i = 1$, and every $\mathrm{A}_i = (A_1, A_2, \ldots, A_n)$ $(i = 1, 2, ..., n)$ is an ordered n-tuple of positive invertible operators on a Hilbert space.

Now, we define a multidimensional geometric operator of Sherman type, deduced from multidimensional Sherman's operator (5.44)

$$\mathcal{S}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \mathbf{b}, \mathbf{S}) = \sum_{j=1}^{n} \sum_{i=1}^{n} b_i s_{ij} G[n,t](\mathbf{X}_j) - \sum_{i=1}^{n} b_i G[n,t]\left(\sum_{j=1}^{n} s_{ij} \mathbf{X}_j\right), \quad (5.53)$$

where $G[n,t]$ is multidimensional weighted geometric mean, $(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$ is $n-$tuple of operators $\mathbf{X}_j = (X_{j1}, X_{j2}, \ldots, X_{jn}) \in \mathcal{D} \subseteq [\mathcal{B}^{++}(H)]^n$ $(j = 1, 2, \ldots, n)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in [0, \infty)^n$ and $\mathbf{S} = (s_{ij}) \in \mathbb{S}_{nn}(\mathbb{R})$.

The operator (5.53) is a generalisation of the multidimensional geometric operator of Jensen type consider in [94, §4.3]:

$$\mathcal{J}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \mathbf{p}) := \sum_{j=1}^{n} p_j G[n,t](\mathbf{X}_j) - G[n,t]\left(\sum_{j=1}^{n} p_j \mathbf{X}_j\right).$$

Since the weighted geometric mean is jointly concave, then similarly as in Corollary 5.1.11., we obtain that the operator (5.53) can mutually be bounded by the simpler operator of the same type. Such estimates can be regarded as both refinement and converse of inequality (5.52). This result is also a generalisation of the result given in [94, Corollary 4].

Corollary 5.1.12.: *Suppose that $\mathcal{S}_{G[n,t]}$ is a geometric operator of Sherman type defined by* (5.53)*. Then*

$$\begin{aligned} & n \cdot \min_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \{b_i s_{ij}\}\ \mathcal{S}^{\mathcal{N}}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n) \\ & \leq \mathcal{S}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \mathbf{b}, \mathbf{S}) \leq n \cdot \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \{b_i s_{ij}\}\ \mathcal{S}^{\mathcal{N}}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n), \end{aligned}$$

where $\mathcal{S}^{\mathcal{N}}$ is non-weighted geometric operator:

$$\mathcal{S}^{\mathcal{N}}_{G[n,t]}(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n) = \sum_{j=1}^{n} G[n,t](\mathbf{X}_j) - G[n,t]\left(\sum_{j=1}^{n} \mathbf{X}_j\right).$$

Proof: Follows from Corollary 5.1.11.. □

5.2. Shannon like inequalities for f-connections of positive linear maps and operators

In this section, for an operator convex function f, we study maps of the form $(T, B) \to TfT^{-}(B)$, called *f-connections* (see Definition 5.2.3.), where T is a positive linear map between unital C^*-algebras $\mathcal{A}$ and $\mathcal{B}$ of Hilbert space operators, T^- denotes a reflexive generalized inverse of T and $B \in \mathcal{B}$. We extend the results for Csiszár's f-divergence functional (4.5) to the operator Csiszár f-divergence, the general case when p_k and q_k are replaced by a positive linear map T_k and a positive operator B_k, $k = 1, 2, \ldots, n$, respectively. We provide conditions under which the inequality $f(1)I \leq \sum_{k=1}^{n} T_k f T_k^{-}(B_k)$ holds, where $T_k : \mathcal{A} \to \mathcal{B}$ are positive linear maps and $B_k \in \operatorname{Ran}(T_k)$ (see Definition 5.2.2.). In consequence, the following Shannon like inequality $0 \leq \sum_{k=1}^{n} T_k f T_k^{-}(B_k)$ is fulfilled, whenever $f(1) = 0$. The obtained results are applied to operator Csiszár f-divergence. In particular, recent results by Isa et al. [75] are recovered. Results in this section are cited from [125] and [128].

5.2.1. Jensen type operator inequalities

By $\mathcal{B}(H)$ we denote the C^*-algebra of all bounded linear operators on Hilbert space H with inner product $\langle \cdot, \cdot \rangle$.

As a special case of Theorem 1.6.17 we get the following theorem with an operator Jensen's inequality (see [114], cf. also [59]).

Theorem 5.2.8 ***([114])*****:** *If f is an operator convex function on an interval J, and $T_1, \ldots, T_n$ are positive linear maps on $\mathcal{B}(H)$ such that $\sum_{k=1}^{n} T_k(I) = I$, then*

$$f\left(\sum_{k=1}^{n} T_k(A_k)\right) \leq \sum_{k=1}^{n} T_k(f(A_k)) \tag{5.54}$$

for all self-adjoint operators $A_k \in \mathcal{B}(H)$, $k = 1, 2, \ldots, n$, with spectra contained in J.

In the next result we show a refinement of the operator Jensen's inequality [86].

Theorem 5.2.9*([86])*: *Let unital C^*-algebras $\mathcal{A}$ and $\mathcal{B}$ be closed *-subalgebras of $\mathcal{B}(H)$ and $\mathcal{B}(K)$, respectively, for some Hilbert spaces H and K. Let $T_1, \ldots, T_n$ be strictly positive linear maps from $\mathcal{A}$ into $\mathcal{B}$, and let $T = \sum_{k=1}^{n} T_k$ be unital.*
If f is an operator convex function on an interval J, then

$$f(T(A)) \leq \sum_{k=1}^{n} T_k(I)^{1/2} f\left(T_k(I)^{-1/2} T_k(A) T_k(I)^{-1/2}\right) T_k(I)^{1/2} \leq T(f(A)) \tag{5.55}$$

for every self-adjoint operator $A \in \mathcal{A}$ with spectrum contained in J.

We now present the notion of generalized inverse of a linear operator.

Definition 5.2.1.*(see [143], [30, pp. 819-820])*: A **generalized inverse**(in short, **g.i.**) of a linear map $T : V \to W$ between linear spaces V and W is a linear map $T^- : W \to V$ satisfying $TT^-T = T$. If in addition $T^-TT^- = T^-$ then T^- is called a **reflexive generalized inverse** of T.

For example, let T be a tripotent linear map (i.e., $T^3 = T$) with $V = W$. It is readily seen that $T^- = T$ is a reflexive generalized inverse of T [125].

On the other hand, if V and W are Hilbert spaces and T is a *partial isometry* (i.e. $TT^*T = T$) then $T^- = T^*$ is a reflexive generalized inverse of T [125].

Example 1.*(Cf. [125, Example 2.2])*: *We consider a $p \times p$ complex matrix T. By making use of Singular Value Decomposition, we get $T = U_2 \, diag \, s(T) U_1^*$ with unitary U_1 and U_2 and with the vector $s(T) = (s_1(T), s_2(T), \ldots, s_p(T))$ of singular values of T [63, p. 144]. We now introduce $T^- = U_1 \, diag \, \sigma(T) U_2^*$, where $\sigma(T) = (\sigma_1(T), \sigma_2(T), \ldots, \sigma_p(T))$, with $\sigma_i(T) = \frac{1}{s_i(T)}$ if $s_i(T) \neq 0$, and $\sigma_i(T) = 0$ if $s_i(T) = 0$, $i = 1, 2, \ldots, p$. It is easy to check that T^- is a reflexive generalized inverse of T. Furthermore, T is positive iff T^- is so.*

Unless stated otherwise, in the remaining of this subsection, we assume that $\mathcal{A} = \mathcal{B}(H)$ with a Hilbert space H and $\mathcal{B}$ is a unital C^*-algebra. Additionally, we assume that there exists a generalized inverse T^- of a linear map T, whenever the symbol T^- is used.

In the sequel, the symbol I denotes for the unity in a unital C^*-algebra.

Definition 5.2.2.: For a linear map $T : \mathcal{A} \to \mathcal{B}$, by Ran(T) we denote the **range**of T, that is

$$\operatorname{Ran}(T) = \{TA \in \mathcal{B} \colon A \in \mathcal{A}\}.$$

In Subsection 5.2.2. we shall explore the following result (see [125], cf. also [129]).

Theorem 5.2.10*([125])*: *Let $T : \mathcal{A} \to \mathcal{B}$ and $T_k : \mathcal{A} \to \mathcal{B}$, $k = 1, 2, \ldots, n$, be positive linear maps and $T^- : \mathcal{B} \to \mathcal{A}$ be a positive reflexive g.i. of T. Assume that*

$$\begin{aligned} T(I) &= \sum_{k=1}^{n} T_k(I), && (5.56) \\ \operatorname{Ran}(T_k) &\subset \operatorname{Ran}(T), \quad k = 1, 2, \ldots, n, && (5.57) \\ I &\in \operatorname{Ran}(T^-). && (5.58) \end{aligned}$$

If f is an operator convex function on an interval J then

$$TfT^-\left(\sum_{k=1}^{n} T_k A_k\right) \leq \sum_{k=1}^{n} T_k f(A_k) \tag{5.59}$$

for all self-adjoint operators $A_1, \ldots, A_n \in \mathcal{A}$ with spectra contained in J.

Definition 5.2.3.*(cf. [85, p. 637])*: The binary map of type

$$(T, B) \to TfT^-(B) \quad \text{for} \quad T \in \mathcal{B}(\mathcal{A}, \mathcal{B}) \quad \text{and} \quad B \in \mathcal{B}, \tag{5.60}$$

is said to be f**-connection.**

If f is an (operator) convex function and $\mathcal{A} = \mathcal{B}$ and $T := A^{1/2}(\cdot)A^{1/2}$ with strictly positive operator A, then (5.60) becomes the *generalized perspective function* defined as follows (see [43, 62, 116]).

Definition 5.2.4. ***([85, p. 637]):*** The map

$$(A, B) \to A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} \quad \text{for } 0 < A, B \in \mathcal{A} = \mathcal{B} \tag{5.61}$$

is called the **generalized perspective function**of f.

It is not hard to verify that for the operator concave functions $f = (\cdot)^{1/2}$ and $f = \log(\cdot)$, the specifications of (5.61) yield *operator geometric mean* [85, 97] and of *relative operator entropy* [18, 48], respectively. Furthermore, some sums (integrals) of maps of type (5.61) lead to *operator Csiszár f-divergence* [38, 116] (see also (5.75)).

5.2.2. Shannon like inequalities for operators

In this subsection $\mathcal{A} = \mathcal{B}(H)$ with a Hilbert space H, and $\mathcal{B}$ is a unital C^*-algebra.

In Theorem 5.2.11 we extend a result by Csiszár and Körner [34] given for n-tuples $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ (see (4.7) or [38, Theorem 2 and Corollary 1]). Namely, the numbers p_k and q_k will be replaced by a positive linear map T_k and a positive operator B_k, $k = 1, 2, \ldots, n$, respectively.

Theorem 5.2.11 ***([128]):*** *Let f be an operator convex function on an interval J with $1 \in J$. Let $T_k : \mathcal{A} \to \mathcal{B}$, $k = 1, 2, \ldots, n$, be positive linear maps such that $\sum\limits_{k=1}^{n} T_k(I) = I$. Let $A_1, \ldots, A_n \in \mathcal{A}$ be self-adjoint operators with spectra contained in J such that $\sum\limits_{k=1}^{n} T_k A_k = I$.*
If there exists a unital positive linear map $T : \mathcal{A} \to \mathcal{B}$ with a positive reflexive g.i. $T^- : \mathcal{B} \to \mathcal{A}$ such that

$$\operatorname{Ran}(T_k) \subset \operatorname{Ran}(T), \quad k = 1, 2, \ldots, n, \tag{5.62}$$

$$I \in \operatorname{Ran}(T^-), \tag{5.63}$$

then

$$f(1) I \leq \sum_{k=1}^{n} T_k f(A_k). \tag{5.64}$$

If in addition $f(1) = 0$, then (5.64) reduces to Shannon like inequality

$$0 \leq \sum_{k=1}^{n} T_k f(A_k). \tag{5.65}$$

Proof: It follows from Theorem 5.2.10 that

$$TfT^{-}\left(\sum_{k=1}^{n} T_k A_k\right) \leq \sum_{k=1}^{n} T_k f(A_k), \tag{5.66}$$

because

$$I = T(I) = \sum_{k=1}^{n} T_k(I). \tag{5.67}$$

But $\sum\limits_{k=1}^{n} T_k A_k = I$, so (5.66) gives

$$TfT^{-}(I) \leq \sum_{k=1}^{n} T_k f(A_k). \tag{5.68}$$

We shall prove that

$$T^{-}(I) = I. \tag{5.69}$$

In fact, by (5.62) we have $I = T^{-}(B_0)$ for some $B_0 \in \mathcal{B}$. It now follows that $T^{-}TT^{-} = T^{-}$, since T^{-} is a reflexive generalized inverse of T. Next, $T^{-}TT^{-}(B_0) = T^{-}(B_0)$, and further $T^{-}T(I) = I$. In addition, $T(I) = I$ by (5.67). So, we obtain $T^{-}(I) = I$, as desired.

Now, (5.68) and (5.69) yields

$$Tf(I) \leq \sum_{k=1}^{n} T_k f(A_k). \tag{5.70}$$

Simultaneously, $f(I) = f(1)I$. Therefore $Tf(I) = f(1)T(I) = f(1)I$ by (5.67). So, (5.70) guarantees that

$$f(1)I \leq \sum_{k=1}^{n} T_k f(A_k).$$

This finishes the proof of (5.64).

Finally, it is evident that if $f(1) = 0$ then (5.64) forces (5.65). □

Remark 5.2.9.: The conditions (5.62)-(5.63) in Theorem 5.2.11 are present because of the (potential) non-invertibility of the maps T and T_k.

Theorem 5.2.11 can be rewritten in terms of $B_k = T_k A_k$, as follows.

Corollary 5.2.13. *([128])*: *Let f be an operator convex function on an interval J with $1 \in J$. Let $T_k : \mathcal{A} \to \mathcal{B}$, $k = 1, 2, \ldots, n$, be positive linear maps such that $\sum_{k=1}^{n} T_k(I) = I$ with positive T_k^-. Let $B_k \in (T_k)$, $k = 1, 2, \ldots, n$, be positive operators such that the spectrum of $T_k^-(B_k)$ is contained in J, and $\sum_{k=1}^{n} B_k = I$.*
If there exists a unital positive linear map $T : \mathcal{A} \to \mathcal{B}$ with a positive reflexive g.i. $T^- : \mathcal{B} \to \mathcal{A}$ such that

$$\operatorname{Ran}(T_k) \subset \operatorname{Ran}(T), \quad k = 1, 2, \ldots, n, \tag{5.71}$$

$$I \in \operatorname{Ran}(T^-), \tag{5.72}$$

then

$$f(1)I \leq \sum_{k=1}^{n} T_k f T_k^-(B_k). \tag{5.73}$$

If in addition $f(1) = 0$ then (5.73) reduces to Shannon like inequality

$$0 \leq \sum_{k=1}^{n} T_k f T_k^-(B_k). \tag{5.74}$$

Proof: Since $B_k \in \operatorname{Ran}(T_k)$ for $k = 1, 2, \ldots, n$, we have $T_k T_k^-(B_k) = B_k$. By setting $A_k = T_k^-(B_k)$, we get $T_k(A_k) = B_k$. By the positivity of T_k^- and $B_k \geq 0$, we also obtain $A_k \geq 0$. Hence A_k is self-adjoint with spectrum contained in J and $\sum_{k=1}^{n} T_k A_k = I$. Now, due to inequality (5.64) in Theorem 5.2.11, we obtain (5.73), as wanted. □

We now reduce the conditions (5.71)-(5.72) of Corollary 5.2.13. by the invertibility of T.

Corollary 5.2.14. *([128])*: *Let f be an operator convex function on an interval J with $1 \in J$. Let $T_k : \mathcal{A} \to \mathcal{B}$, $k = 1, 2, \ldots, n$, be positive linear maps such that $\sum_{k=1}^{n} T_k(I) = I$ with positive T_k^-. Let $B_k \in (T_k)$, $k = 1, 2, \ldots, n$, be positive operators such that the spectrum of $T_k^-(B_k)$ is contained in J, and $\sum_{k=1}^{n} B_k = I$.*
If there exists an invertible unital positive linear map $T : \mathcal{A} \to \mathcal{B}$ with positive inverse $T^{-1} : \mathcal{B} \to \mathcal{A}$, then inequality (5.73) holds.

If in addition $f(1) = 0$ then Shannon like inequality (5.74) holds.

Proof: From $T^{-} = T^{-1}$, we obtain $\mathrm{Ran}(T) = \mathcal{B}$ and $\mathrm{Ran}(T^{-}) = \mathcal{A}$. So, conditions (5.71) and (5.72) are fulfilled automatically. Now, Corollary 5.2.14. gives (5.73) and (5.74), as claimed. $\square$

In the case $\mathcal{A} = \mathcal{B}$, we get the following simplification of Corollary 5.2.14..

Corollary 5.2.15. *([128])*: *Let f be an operator convex function on an interval J with $1 \in J$. Let $T_k : \mathcal{A} \to \mathcal{A}$, $k = 1, 2, \ldots, n$, be positive linear maps such that $\sum\limits_{k=1}^{n} T_k(I) = I$ with positive T_k^{-}. Let $B_k \in (T_k)$, $k = 1, 2, \ldots, n$, be positive operators such that the spectrum of $T_k^{-}(B_k)$ is contained in J, and $\sum\limits_{k=1}^{n} B_k = I$.*
Then inequality (5.73) holds.

If in addition $f(1) = 0$ then Shannon like inequality (5.74) holds.

Proof: By taking T to be the identity map id on $\mathcal{A}$, we have that $T : \mathcal{A} \to \mathcal{A}$ is invertible, unital and positive, and the inverse $T^{-1} = id$ is positive, as well. So, by using Corollary 5.2.14. we establish (5.73) and (5.74), as desired. $\square$

5.2.3. Applications to operator divergences, means and entropies

It is assumed in this subsection that $\mathcal{A} = \mathcal{B} = \mathcal{B}(H)$ with a Hilbert space H.

For operators $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \ldots, n$, we put

$$\mathbf{A} = (A_1, A_2, \ldots, A_n) \quad \text{and} \quad \mathbf{B} = (B_1, B_2, \ldots, B_n).$$

Definition 5.2.5. *([33, 116])*: For an operator convex function f defined on $J = (0, \infty)$, **the operator Csiszár f-divergence** of $\mathbf{A}$ and $\mathbf{B}$ is defined by

$$D_f(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{n} A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}. \tag{5.75}$$

Theorem 5.2.12 *([128])*: *Let f be an operator convex function on $J = (0, \infty)$. Let $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \ldots, n$, be such that $\sum\limits_{k=1}^{n} A_k = I$ and $\sum\limits_{k=1}^{n} B_k = I$.*

Then

$$f(1)I \leq \sum_{k=1}^{n} A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}. \tag{5.76}$$

If in addition $f(1) = 0$, *then*

$$0 \leq \sum_{k=1}^{n} A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}. \tag{5.77}$$

Proof: We set $T_k = A_k^{1/2}(\cdot)A_k^{1/2}$, $k = 1, 2, \ldots, n$. Clearly, $T_k : \mathcal{A} \to \mathcal{A}$, $k = 1, 2, \ldots, n$, are invertible positive linear maps. Moreover $T_k^{-1} = A_k^{-1/2}(\cdot)A_k^{-1/2}$ is positive.

It is obvious that

$$\sum_{k=1}^{n} T_k(I) = \sum_{k=1}^{n} A_k^{1/2} I A_k^{1/2} = \sum_{k=1}^{n} A_k = I. \tag{5.78}$$

It is not hard to check that the spectra of $T_k^{-1}(B_k) = A_k^{-1/2} B_k A_k^{-1/2}$, $k = 1, 2, \ldots, n$, are contained in the interval $J = (0, \infty)$.

On account of Corollary 5.2.15.,

$$f(1)I \leq \sum_{k=1}^{n} T_k f T_k^{-1}(B_k).$$

We also have

$$T_k f T_k^{-1}(B_k) = A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}.$$

Summarizing all of this we deduce that

$$f(1)I \leq \sum_{k=1}^{n} A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2},$$

and therefore

$$0 \leq \sum_{k=1}^{n} A_k^{1/2} f(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}, \quad \text{whenever } f(1) = 0.$$

This finishes the proof of (5.76) and (5.77). □

We extend operator means from Chapter 1 as follows.

Definition 5.2.6.*([75, pp. 3149-3950])*: For n-tuples of positive operators

$$\mathbf{A} = (A_1, A_2, \dots, A_n) \quad \text{and} \quad \mathbf{B} = (B_1, B_2, \dots, B_n)$$

with $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \dots, n$, we define

$$S(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{n} S(A_k, B_k) = \sum_{k=1}^{n} A_k^{1/2} \log(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}, \tag{5.79}$$

$$S_p(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{n} S_p(A_k, B_k) = \sum_{k=1}^{n} A_k^{1/2} (A_k^{-1/2} B_k A_k^{-1/2})^p \log(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}, \tag{5.80}$$

$$T_p(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{n} T_p(A_k, B_k) = \sum_{k=1}^{n} A_k^{1/2} \ln_p(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2}, \tag{5.81}$$

where $p \in (0, 1]$ and $\ln_p t = \frac{t^p - 1}{p}$, $t > 0$,

$$T_{s,r}(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{n} T_{s,r}(A_k, B_k) = \sum_{k=1}^{n} A_k^{1/2} \frac{[(1-s)I + s(A_k^{-1/2} B_k A_k^{-1/2})^r]^{1/r} - I}{s} A_k^{1/2}, \tag{5.82}$$

where $s \in (0, 1]$ and $r \in [-1, 1]$.

We now demonstrate some specifications of Theorem 5.2.12.

Corollary 5.2.16.*(Furuta [53, Corollary 2.4])*: *Let $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \dots, n$, be such that $\sum\limits_{k=1}^{n} A_k = I$ and $\sum\limits_{k=1}^{n} B_k = I$.*
Then

$$0 \geq \sum_{k=1}^{n} S(A_k, B_k). \tag{5.83}$$

Proof: It is enough to apply Theorem 5.2.12 to the function operator convex $f(t) = -\log t$, $t > 0$, with $f(1) = 0$. Therefore we get

$$0 \geq \sum_{k=1}^{n} A_k^{1/2} \log(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2} = \sum_{k=1}^{n} S(A_k, B_k),$$

completing the proof. □

Corollary 5.2.17.*(Yanagi et al. [160, Theorem 1])*: *Let $0 < A_k \in \mathcal{A}$ and $0 < B_k \in$*

$\mathcal{A}$, $k = 1, 2, \ldots, n$, be such that $\sum_{k=1}^{n} A_k = I$ and $\sum_{k=1}^{n} B_k = I$.
Then for $0 < p \leq 1$,

$$0 \geq \sum_{k=1}^{n} T_p(A_k, B_k). \tag{5.84}$$

Proof: It follows from [23, Theorem V.1.9] that the function $t \to t^p$, $0 < p \leq 1$, is operator monotone on $[0, \infty)$. Therefore this function is operator concave on $[0, \infty)$ (see [23, Theorem V.2.5]). Consequently, the function $g(t) = \ln_p t = \frac{t^p - 1}{p}$, $t \geq 0$, is operator concave on $[0, \infty)$. Now, according to Theorem 5.2.12 applied to the function $f(t) = -\ln_p t$, $t > 0$, we infer that

$$0 \geq \sum_{k=1}^{n} A_k^{1/2} \ln_p(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2} = \sum_{k=1}^{n} T_p(A_k, B_k),$$

as was to be proved. □

Corollary 5.2.18. ***(Furuta [53, Corollary 2.4]):*** *Let $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \ldots, n$, be such that $\sum_{k=1}^{n} A_k = I$ and $\sum_{k=1}^{n} B_k = I$.*
Then

$$0 \leq \sum_{k=1}^{n} S_1(A_k, B_k). \tag{5.85}$$

Proof: According to Theorem 5.2.12 applied to for the operator convex function $f(t) = t \log t$, $t > 0$, with $f(1) = 0$, one has

$$0 \leq \sum_{k=1}^{n} A_k^{1/2} (A_k^{-1/2} B_k A_k^{-1/2}) \log(A_k^{-1/2} B_k A_k^{-1/2}) A_k^{1/2} = \sum_{k=1}^{n} S_1(A_k, B_k),$$

which implies (5.85). □

Corollary 5.2.18. states that a Shannon type inequality is satisfied for the *Kullback-Leibler distance* $D_{KL}(\mathbf{A}, \mathbf{B}) = S_1(\mathbf{A}, \mathbf{B})$.

Corollary 5.2.19. ***(Isa et al. [75, Theorems 2.1 and 2.2]):*** *Let $0 < A_k \in \mathcal{A}$ and $0 < B_k \in \mathcal{A}$, $k = 1, 2, \ldots, n$, be such that $\sum_{k=1}^{n} A_k = I$ and $\sum_{k=1}^{n} B_k = I$.*

Then for $0 < s < 1$ *and* $0 < r \leq 1$,

$$0 \geq \sum_{k=1}^{n} T_{s,r}(A_k, B_k). \tag{5.86}$$

Proof: The parametric Tsallis relative operator entropy $T_{s,r}(A_k, B_k)$ (see (1.39)) is generated by operator concave function

$$f(t) = \frac{(1 - s + st^r)^{1/r} - 1}{s}, \quad t > 0.$$

Clearly, here we have $f(1) = 0$. In light of Theorem 5.2.12, one sees that

$$0 \geq \sum_{k=1}^{n} A_k^{1/2} \frac{[(1-s)I + s(A_k^{-1/2} B_k A_k^{-1/2})^r]^{1/r} - I}{s} A_k^{1/2} = \sum_{k=1}^{n} T_{s,r}(A_k, B_k).$$

This completes the proof. □

The *α-order Rényi entropy* is induced by the power function $f(t) = t^\alpha$ with $\alpha \in (0,1)$. It is obvious that $f(1) = 1 \neq 0$.

Corollary 5.2.20. ***([128]):*** *Let* $0 < A_k \in \mathcal{A}$ *and* $0 < B_k \in \mathcal{A}$, $k = 1, 2, \ldots, n$, *be such that* $\sum_{k=1}^{n} A_k = I$ *and* $\sum_{k=1}^{n} B_k = I$.
Then for $\alpha \in (0,1)$,

$$I \geq \sum_{k=1}^{n} A_k \sharp_\alpha B_k. \tag{5.87}$$

Proof: It is well known that the function $f(t) = t^\alpha$, $t > 0$, $0 < \alpha < 1$, is operator concave. Due to Theorem 5.2.12 we have

$$I \geq \sum_{k=1}^{n} A_k^{1/2} \left(A_k^{-1/2} B_k A_k^{-1/2} \right)^\alpha A_k^{1/2} = \sum_{k=1}^{n} A_k \sharp_\alpha B_k, \tag{5.88}$$

as required. □

The operator Csiszár inequalities (5.76)-(5.77) embrace a number of important results.

Example 2. ***([128]):*** *Inequality (5.87) with* $\alpha = \frac{1}{2}$ *is connected with A-G inequality. Namely, let* $0 < A_k \in A$ *and* $0 < B_k \in A$, $k = 1, 2, \ldots, n$, *be such that* $\sum_{k=1}^{n} A_k = I$ *and*

$\sum\limits_{k=1}^{n} B_k = I$. *Using Theorem 5.2.12 with the operator convex Hellinger distance function*

$$f(t) = \frac{1}{2}(1-\sqrt{t})^2, \quad t>0,$$

we derive

$$\begin{aligned} 0 &\le \sum_{k=1}^{n} A_k^{1/2}\left(I-(A_k^{-1/2}B_kA_k^{-1/2})^{1/2}\right)^2 A_k^{1/2} \\ &= \sum_{k=1}^{n}\left(A_k - 2A_k^{1/2}(A_k^{-1/2}B_kA_k^{-1/2})^{1/2}A_k^{1/2}+B_k\right) \\ &= \sum_{k=1}^{n}(A_k - 2A_k\sharp B_k + B_k) = 2I - 2\sum_{k=1}^{n}(A_k\sharp B_k)\,. \end{aligned}$$

Therefore the inequality (5.87) holds with $\alpha = \frac{1}{2}$.

On the other hand, the above estimation is valid by the following A-G operator inequality: $A_k\sharp B_k \le \frac{1}{2}(A_k+B_k)$, $k=1,\dots,n$ *(see [163]).*

Example 3. *([128]):* *Let* $0 < A_k \in A$ *and* $0 < B_k \in A$, $k=1,2,\dots,n$, *with* $\sum\limits_{k=1}^{n} A_k = I$ *and* $\sum\limits_{k=1}^{n} B_k = I$. *Then*

$$I \le \sum_{k=1}^{n} B_kA_k^{-1}B_k = \sum_{k=1}^{n} A_k\sharp_2 B_k. \tag{5.89}$$

To prove inequality (5.89), we utilize Theorem 5.2.12 with the χ^2*-distance function* $f(t) = (t-1)^2$, $t\in J=(0,\infty)$. *In fact,* f *is operator convex and* $f(1)=0$. *Thus we obtain*

$$\begin{aligned} 0 &\le \sum_{k=1}^{n} A_k^{1/2}\left(A_k^{-1/2}B_kA_k^{-1/2}-I\right)^2 A_k^{1/2} \\ &= \sum_{k=1}^{n} A_k^{1/2}\left((A_k^{-1/2}B_kA_k^{-1/2})^2 - 2A_k^{-1/2}B_kA_k^{-1/2}+I\right)A_k^{1/2} \\ &= \sum_{k=1}^{n} B_kA_k^{-1}B_k - 2\sum_{k=1}^{n}B_k + \sum_{k=1}^{n}A_k = \sum_{k=1}^{n} B_kA_k^{-1}B_k - 2\sum_{k=1}^{n}B_k + \sum_{k=1}^{n}A_k \\ &= \sum_{k=1}^{n} B_kA_k^{-1}B_k - I = \sum_{k=1}^{n} A_k\sharp_2 B_k - I, \end{aligned}$$

which gives (5.89).

The scalar version of (5.89) is

$$1 \le \sum_{k=1}^{n} \frac{b_k^2}{a_k} \tag{5.90}$$

with the assumptions $\sum_{k=1}^{n} a_k = 1 = \sum_{k=1}^{n} b_k$ *and* $a_k, b_k > 0$, $k = 1, 2, \ldots, n$.
In the case $a_k = \frac{1}{n}$, $k = 1, 2, \ldots, n$, *(5.90) asserts that if* $\sum_{k=1}^{n} b_k = 1$ *then* $\frac{1}{n} \le \sum_{k=1}^{n} b_k^2$.
Another way to see this is possible by using the Theorem 2.1.1. Indeed, the function $t \mapsto t^2$, $t > 0$, *is convex, and the vector* $(\frac{1}{n}, \ldots, \frac{1}{n})$ *is majorized by* $(b_1, \ldots, b_n)$.

5.3. Generalization of the Csiszár-Körner inequality for f -divergence

In this section we investigate vectorial joint majorization for comparing m-tuple and n-tuple of pairs in the Cartesian product of two real linear spaces. We derive a Sherman type inequality for a vector-valued $\le_C$-convex function f, where $\le_C$ is a cone preordering. We also show the Hardy-Littlewood-Pólya-Karamata type theorem. As applications, we generalize the Csiszár-Körner inequality for f-divergence of two n-tuples of positive numbers. To do so, we apply n-tuples of pairs of positive linear maps and positive operators on a Hilbert space. Material of this section is cited from [125] and [132].

5.3.1. Motivation

The following **Csiszár-Körner operator type inequality** is fulfilled [125].

Theorem 5.3.13 ***([125])*****:** *Let* $T_k : \mathcal{A} \to \mathcal{B}$, $k = 1, 2, \ldots, n$, *be positive linear maps, and* $T_k^- : \mathcal{B} \to \mathcal{A}$ *be a positive g.i. of* T_k, *and* $\left(\sum_{k=1}^{n} T_k\right)^- : \mathcal{B} \to \mathcal{A}$ *be a positive reflexive g.i. of* $\sum_{k=1}^{n} T_k$. *Assume that*

$$I \in \operatorname{Ran}\left(\sum_{k=1}^{n} T_k\right)^-, \tag{5.91}$$

$$\operatorname{Ran}(T_k) \subset \operatorname{Ran}\left(\sum_{k=1}^{n} T_k\right), \quad k = 1, 2, \ldots, n. \tag{5.92}$$

If f is an operator convex function on $[0,\infty)$ then

$$\left(\sum_{k=1}^{n} T_k\right) f \left(\sum_{k=1}^{n} T_k\right)^{-} \left(\sum_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} T_k f T_k^{-}(B_k) \tag{5.93}$$

for every positive operators $0 \leq B_k \in \operatorname{Ran}(T_k)$, $k = 1, 2, \ldots, n$.

We are going to develop the above result to majorization form. To do so, in Subsection 5.3.2. we introduce vectorial joint majorization prepared to comparing m-tuple and n-tuple of (column) pairs in the Cartesian product of two arbitrary linear spaces (see [31, 35, 154]). Next we prove a Sherman type inequality for the vectorial joint majorization. We also derive a Hardy-Littlewood-P ´olya-Karamata like theorem for vector-valued $\leq_C$-convex functions.

In Subsection 5.3.3. we deal with applications for Csiszár operator f-divergence. We employ the results of Subsection 5.3.2. to n-tuples of pairs of positive linear maps and positive operators. We show that the Csiszár-Körner inequality corresponds to the simplest joint majorization $(\binom{\bar{T}}{\bar{B}}, \ldots, \binom{\bar{T}}{\bar{B}}) \prec (\binom{T_1}{B_1}, \ldots, \binom{T_n}{B_n})$, where $\bar{T} = \frac{T_1+\ldots+T_n}{n}$ and $\bar{B} = \frac{B_1+\ldots+B_n}{n}$. Furthermore, we extend the C-K inequality to the case of a general joint majorization $(\binom{R_1}{A_1}, \ldots, \binom{R_n}{A_n}) \prec (\binom{T_1}{B_1}, \ldots, \binom{T_n}{B_n})$ with positive operators $A_1, \ldots, A_n$, $B_1, \ldots, B_n$ in $\mathcal{B}$, and with positive linear maps $R_1, \ldots, R_n$, $T_1, \ldots, T_n$ from $\mathcal{A}$ to $\mathcal{B}$, where $\mathcal{A} = \mathcal{B} = \mathcal{B}(H)$ with a Hilbert space H.

5.3.2. Vectorial joint majorization

In the sequel we utilize the Cartesian product space $V = V_1 \times V_2$ with arbitrary real linear spaces V_1 and V_2. For the notational convenience, we write the elements of $V = V_1 \times V_2$ as column pairs $\binom{\mathbf{x}}{\mathbf{w}}$ with $\mathbf{x} \in V_1$ and $\mathbf{w} \in V_2$.

Here the symbol $\mathbf{e}$ stands for the vector $(1, \ldots, 1)$ of all ones of appropriate dimension.

Definition 5.3.7. ***([132]):*** An n-tuple of (column) pairs

$$\begin{pmatrix}\mathbf{Y}\\ \mathbf{U}\end{pmatrix} = \left(\begin{pmatrix}\mathbf{y}_1\\ \mathbf{u}_1\end{pmatrix}, \begin{pmatrix}\mathbf{y}_2\\ \mathbf{u}_2\end{pmatrix}, \ldots, \begin{pmatrix}\mathbf{y}_n\\ \mathbf{u}_n\end{pmatrix}\right) \in (V_1 \times V_2)^n$$

is said to be **jointly pre-majorized** by an m-tuple of (column) pairs

$$\binom{\mathbf{X}}{\mathbf{W}} = \left(\binom{\mathbf{x}_1}{\mathbf{w}_1}, \binom{\mathbf{x}_2}{\mathbf{w}_2}, \ldots, \binom{\mathbf{x}_m}{\mathbf{w}_m} \right) \in (V_1 \times V_2)^m,$$

written as $\binom{\mathbf{Y}}{\mathbf{U}} \prec_p \binom{\mathbf{X}}{\mathbf{W}}$, if there exists an $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij})$ such that

$$\left(\binom{\mathbf{y}_1}{\mathbf{u}_1}, \binom{\mathbf{y}_2}{\mathbf{u}_2}, \ldots, \binom{\mathbf{y}_n}{\mathbf{u}_n} \right) = \left(\binom{\mathbf{x}_1}{\mathbf{w}_1}, \binom{\mathbf{x}_2}{\mathbf{w}_2}, \ldots, \binom{\mathbf{x}_m}{\mathbf{w}_m} \right) \mathbf{S} \tag{5.94}$$

in the sense that

$$\binom{\mathbf{y}_j}{\mathbf{u}_j} = s_{1j}\binom{\mathbf{x}_1}{\mathbf{w}_1} + s_{2j}\binom{\mathbf{x}_2}{\mathbf{w}_2} + \ldots + s_{mj}\binom{\mathbf{x}_m}{\mathbf{w}_m}, \quad \text{for } j = 1, 2, \ldots, n,$$

that is, for $j = 1, 2, \ldots, n$,

$$\mathbf{y}_j = s_{1j}\mathbf{x}_1 + s_{2j}\mathbf{x}_2 + \ldots + s_{mj}\mathbf{x}_m \quad \text{and} \quad \mathbf{u}_j = s_{1j}\mathbf{w}_1 + s_{2j}\mathbf{w}_2 + \ldots + s_{mj}\mathbf{w}_m.$$

Definition 5.3.8. *([132])***:** An n-tuple of pairs

$$\binom{\mathbf{Y}}{\mathbf{U}} = \left(\binom{\mathbf{y}_1}{\mathbf{u}_1}, \binom{\mathbf{y}_2}{\mathbf{u}_2}, \ldots, \binom{\mathbf{y}_n}{\mathbf{u}_n} \right) \in (V_1 \times V_2)^n$$

is said to be **jointly majorized**by n-tuple of pairs

$$\binom{\mathbf{X}}{\mathbf{W}} = \left(\binom{\mathbf{x}_1}{\mathbf{w}_1}, \binom{\mathbf{x}_2}{\mathbf{w}_2}, \ldots, \binom{\mathbf{x}_n}{\mathbf{w}_n} \right) \in (V_1 \times V_2)^n,$$

written as $\binom{\mathbf{Y}}{\mathbf{U}} \prec \binom{\mathbf{X}}{\mathbf{W}}$, if there exists an $n \times n$ doubly stochastic matrix $\mathbf{S} = (s_{ij})$ such that

$$\left(\binom{\mathbf{y}_1}{\mathbf{u}_1}, \binom{\mathbf{y}_2}{\mathbf{u}_2}, \ldots, \binom{\mathbf{y}_n}{\mathbf{u}_n} \right) = \left(\binom{\mathbf{x}_1}{\mathbf{w}_1}, \binom{\mathbf{x}_2}{\mathbf{w}_2}, \ldots, \binom{\mathbf{x}_n}{\mathbf{w}_n} \right) \mathbf{S}. \tag{5.95}$$

From now on we assume that W is a real linear space endowed with a convex cone $C \subset W$ inducing a cone preordering $\leq_C$ on W.

Definition 5.3.9. *([132])***:** A function $f : J_1 \times J_2 \to W$, with convex sets $J_1 \subset V_1$ and

$J_2 \subset V_2$, is called **jointly $\leq_C$-convex** on $J_1 \times J_2$, if

$$f\left(\alpha\begin{pmatrix}\mathbf{x}\\\mathbf{w}\end{pmatrix}+(1-\alpha)\begin{pmatrix}\mathbf{y}\\\mathbf{u}\end{pmatrix}\right)\leq_C \alpha f\begin{pmatrix}\mathbf{x}\\\mathbf{w}\end{pmatrix}+(1-\alpha)f\begin{pmatrix}\mathbf{y}\\\mathbf{u}\end{pmatrix}$$

for $\alpha \in [0,1]$, and $x, y \in J_1$ and $w, u \in J_2$.

A function $f : J_1 \times J_2 \to W$ with convex sets $J_1 \subset V_1$ and $J_2 \subset V_2$, is called **jointly $\leq_C$-concave** if $-f$ is jointly $\leq_C$-convex (see [23, p. 271]).

We are now ready to prove a Sherman type inequality (5.97) for vector-valued function f defined on a set of pairs of vectors (cf. [149], see also [25, 31]).

Theorem 5.3.14*([132])*: *Let $f : J_1 \times J_2 \to W$ be a jointly $\leq_C$-convex function with convex sets $J_1 \subset V_1$ and $J_2 \subset V_2$. Suppose that*

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) \in J_1^m \quad \textit{and} \quad \mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in J_1^n,$$

$$\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m) \in J_2^m \quad \textit{and} \quad \mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \in J_2^n,$$

$$\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{R}_+^m \quad \textit{and} \quad \mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}_+^n.$$

If

$$\begin{pmatrix}\mathbf{Y}\\\mathbf{U}\end{pmatrix} = \begin{pmatrix}\mathbf{X}\\\mathbf{W}\end{pmatrix}\mathbf{S} \quad \textit{and} \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{5.96}$$

for some $m \times n$ column stochastic matrix $\mathbf{S} = (s_{ij})$, then

$$\sum_{j=1}^{n} b_j f\begin{pmatrix}\mathbf{y}_j\\\mathbf{u}_j\end{pmatrix} \leq_C \sum_{i=1}^{m} a_i f\begin{pmatrix}\mathbf{x}_i\\\mathbf{w}_i\end{pmatrix}. \tag{5.97}$$

If f is jointly $\leq_C$-concave, then the inequality (5.97) is reversed.

Proof: It follows from (5.96) that

$$\left(\begin{pmatrix}\mathbf{y}_1\\\mathbf{u}_1\end{pmatrix}, \begin{pmatrix}\mathbf{y}_2\\\mathbf{u}_2\end{pmatrix}, \ldots, \begin{pmatrix}\mathbf{y}_n\\\mathbf{u}_n\end{pmatrix}\right) = \left(\begin{pmatrix}\mathbf{x}_1\\\mathbf{w}_1\end{pmatrix}, \begin{pmatrix}\mathbf{x}_2\\\mathbf{w}_2\end{pmatrix}, \ldots, \begin{pmatrix}\mathbf{x}_m\\\mathbf{w}_m\end{pmatrix}\right)\mathbf{S}.$$

Hence we get $\mathbf{y}_j = \sum_{i=1}^{m} a_{ij}\mathbf{x}_i$ and $\mathbf{u}_j = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i$ with $\sum_{i=1}^{m} a_{ij} = 1$, $j = 1, 2, \ldots, n$, and

$a_{i,j} \geq 0$. By using the joint $\leq_C$-convexity of f we have

$$f\begin{pmatrix}\mathbf{y}_j\\ \mathbf{u}_j\end{pmatrix} = f\begin{pmatrix}\sum\limits_{i=1}^{m} a_{ij}\mathbf{x}_i\\ \sum\limits_{i=1}^{m} a_{ij}\ \mathbf{w}_i\end{pmatrix}$$

$$= f\left(\sum_{i=1}^{m} a_{ij}\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix}\right) \leq_C \sum_{i=1}^{m} a_{ij} f\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix} \quad \text{for } j = 1, 2, \ldots, n.$$

Therefore we obtain

$$\sum_{j=1}^{n} b_j f\begin{pmatrix}\mathbf{y}_j\\ \mathbf{u}_j\end{pmatrix} \leq_C \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij} f\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix}.$$

This gives

$$\sum_{j=1}^{n} b_j f\begin{pmatrix}\mathbf{y}_j\\ \mathbf{u}_j\end{pmatrix} \leq_C \sum_{i=1}^{m}\sum_{j=1}^{n} b_j a_{ij} f\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix}.$$

In addition, from (5.96) we get $\mathbf{a} = \mathbf{b}\mathbf{S}^T$, which amounts to $a_i = \sum\limits_{j=1}^{n} b_j a_{ij}$, $i = 1, 2, \ldots, m$. Consequently, we conclude that

$$\sum_{j=1}^{n} b_j f\begin{pmatrix}\mathbf{y}_j\\ \mathbf{u}_j\end{pmatrix} \leq_C \sum_{i=1}^{m}\left(\sum_{j=1}^{n} b_j a_{ij}\right) f\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix} = \sum_{i=1}^{m} a_i f\begin{pmatrix}\mathbf{x}_i\\ \mathbf{w}_i\end{pmatrix},$$

completing the proof of inequality (5.97). □

We now present the Hardy-Littlewood-Pólya-Karamata type result (5.99) for vector-valued function f and jointly majorized n-tuples of pairs of vectors (cf. [149], see also [25, 31]).

Theorem 5.3.15 *([132]):* *Let $f : J_1 \times J_2 \to W$ be a jointly $\leq_C$ -convex function with convex sets $J_1 \subset V_1$ and $J_2 \subset V_2$. Suppose that*

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in J_1^n \quad \textit{and} \quad \mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \in J_1^n,$$

$$\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n) \in J_2^n \quad \textit{and} \quad \mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \in J_2^n.$$

If

$$\begin{pmatrix}\mathbf{Y}\\ \mathbf{U}\end{pmatrix} \prec \begin{pmatrix}\mathbf{X}\\ \mathbf{W}\end{pmatrix}, \tag{5.98}$$

then

$$\sum_{j=1}^{n} f\begin{pmatrix} \mathbf{y}_j \\ \mathbf{u}_j \end{pmatrix} \leq_C \sum_{i=1}^{n} f\begin{pmatrix} \mathbf{x}_i \\ \mathbf{w}_i \end{pmatrix}. \tag{5.99}$$

If f is jointly $\leq_C$-concave, then the inequality (5.99) is reversed.

Proof: On account of (5.98) we have

$$\left(\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{u}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{y}_2 \\ \mathbf{u}_2 \end{pmatrix}, \dots, \begin{pmatrix} \mathbf{y}_n \\ \mathbf{u}_n \end{pmatrix}\right) = \left(\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{w}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{x}_2 \\ \mathbf{w}_2 \end{pmatrix}, \dots, \begin{pmatrix} \mathbf{x}_n \\ \mathbf{w}_n \end{pmatrix}\right) \mathbf{S}$$

for some $n \times n$ doubly stochastic matrix $\mathbf{S}$.

It is now enough to utilize Theorem 5.3.14 with $\mathbf{a} = \mathbf{b} = \mathbf{e} = (1, \dots, 1) \in \mathbb{R}^n$, since requirement (5.96) is satisfied. So, we infer that (5.97) becomes (5.99). This completes the proof. □

Example 4. *([132])*: *We now illustrate our theory by showing some results connected with [37, Theorem 3]. To this end we apply Theorem 5.3.14 with the following specification*

$$\begin{aligned}
V_1 &= V = \textit{a linear space over } \mathbb{R}, \\
V_2 &= W = \mathbb{R}, \\
J_1 &= K = \textit{a (nonempty) convex cone in } V, \\
J_2 &= (0, \infty) = \textit{the interval of positive numbers in } \mathbb{R}, \\
C &= [0, \infty) = \textit{the convex cone of nonnegative numbers in } \mathbb{R}, \\
\leq_C &\ \textit{is} \ \ \textit{the standard ordering} \leq \textit{on } \mathbb{R}.
\end{aligned}$$

For a convex function g defined on K, we introduce

$$f\begin{pmatrix} \mathbf{x} \\ w \end{pmatrix} = wg\left(\frac{\mathbf{x}}{w}\right) \quad \textit{for } \mathbf{x} \in K, \ w \in (0, \infty)$$

(see [37, p. 1503]). It is easily seen that f is jointly convex on the set $K \times (0, \infty) = \{\binom{\mathbf{x}}{w} : x \in K, \ w \in (0, \infty)\}$, i.e., it holds that

$$f\left(\alpha\begin{pmatrix} \mathbf{x} \\ w \end{pmatrix} + (1-\alpha)\begin{pmatrix} \mathbf{y} \\ u \end{pmatrix}\right) \leq \alpha f\begin{pmatrix} \mathbf{x} \\ w \end{pmatrix} + (1-\alpha) f\begin{pmatrix} \mathbf{y} \\ u \end{pmatrix}$$

for $\alpha \in [0,1]$, $x, y \in K$ *and* $w, u \in (0,\infty)$.

Let

$$\mathbf{x}_i, \mathbf{y}_j \in K, \quad and \quad w_i, u_j \in (0,\infty) \quad and \quad a_i, b_j \in [0,\infty)$$

for $i = 1,2,\ldots,m$ *and* $j = 1,2,\ldots,n$.

Denoting

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) \quad and \quad \mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n),$$

$$\mathbf{W} = (w_1, w_2, \ldots, w_m) \quad and \quad \mathbf{U} = (u_1, u_2, \ldots, u_n),$$

$$\mathbf{a} = (a_1, a_2, \ldots, a_m) \quad and \quad \mathbf{b} = (b_1, b_2, \ldots, b_n),$$

and making use of Theorem 5.3.14, we get the following result.

If

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ \mathbf{W} \end{pmatrix} \mathbf{S} \quad and \quad \mathbf{a} = \mathbf{b}\mathbf{S}^T \tag{5.100}$$

for some $m \times n$ *column stochastic matrix* $S = (s_{ij})$, *then*

$$\sum_{j=1}^{n} b_j u_j g\left(\frac{\mathbf{y}_j}{u_j}\right) \le \sum_{i=1}^{m} a_i w_i g\left(\frac{\mathbf{x}_i}{w_i}\right). \tag{5.101}$$

For instance, if $m = n$, $b_j = 1$ *and* $a_i = 1$ *for* $i, j = 1,2,\ldots,n$, *then it follows from Theorem 5.3.15 that*

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{U} \end{pmatrix} \prec \begin{pmatrix} \mathbf{X} \\ \mathbf{W} \end{pmatrix} \quad \Rightarrow \quad \sum_{i=1}^{n} u_i g\left(\frac{\mathbf{y}_i}{u_i}\right) \le \sum_{i=1}^{n} w_i g\left(\frac{\mathbf{x}_i}{w_i}\right) \tag{5.102}$$

(cf. [37, Theorem 3]).

5.3.3. Applications to operator Csiszár f-divergence

It is now assumed that $\mathcal{A} = \mathcal{B}(H)$ is the C^*-algebra of bounded linear operators on a Hilbert space H, with the following setting

$$\begin{aligned}
\mathcal{A} &= \mathcal{B} = \mathcal{B}(H),\\
V_1 &= \mathcal{L}(\mathcal{A}) = \text{the real linear space of linear maps from } \mathcal{A} \text{ to } \mathcal{A},\\
J_1 &= \mathcal{P}(\mathcal{A}) = \text{the convex cone of positive linear maps from } \mathcal{A} \text{ to } \mathcal{A},\\
V_2 &= \mathcal{A},\\
J_2 &= \text{the convex cone of positive linear operators from } H \text{ to } H,\\
W &= \mathcal{A},\\
C &= \text{the convex cone of positive linear operators from } H \text{ to } H,\\
\leq_C \ &is \ \text{the usual Loewner ordering} \leq \text{on self-adjoint operators in } \mathcal{A}.
\end{aligned}$$

As previously, the symbol $(\cdot)^-$ denotes a generalized inverse operator on positive linear maps such that if $T : \mathcal{A} \to \mathcal{A}$ is a positive linear map then $T^- : \mathcal{A} \to \mathcal{A}$ is positive.

Example 5. *([132]): Let $T = T_A = A^{1/2}(\cdot)A^{1/2}$ for positive operators $0 \leq A \in A$. Then T is a positive linear map. We introduce $T^- = T_A^- = (A^-)^{1/2}(\cdot)(A^-)^{1/2}$, where A^- is a generalized inverse of A. Clearly, T^- is positive. Moreover, T^- is reflexive iff A^- is so.*

Let $T_k : \mathcal{A} \to \mathcal{B}$ be positive linear maps such that $T_k^- : \mathcal{B} \to \mathcal{A}$ are positive, and $0 \leq B_k \in \mathrm{Ran}(T_k)$ be positive operators, $k = 1, 2, \ldots, n$, and let $\left(\sum_{k=1}^n T_k\right)^- : \mathcal{B} \to \mathcal{A}$ be a positive reflexive g.i. of $\sum_{k=1}^n T_k$. We denote $\mathbf{T} = (T_1, \ldots, T_n)$ and $\mathbf{B} = (B_1, \ldots, B_n)$.

Let $f : [0, \infty) \to \mathbb{R}$ be an operator convex function. In light of the operator calculus, the action of f can be extended to the positive self-adjoint operators of the C^*-algebra $\mathcal{A}$.

In this subsection we shall prove that the operator Csiszár f-divergence

$$D_f(\mathbf{T}, \mathbf{B}) = \sum_{k=1}^n T_k f T_k^-(B_k)$$

is a jointly Schur-convex function of two variables $\mathbf{T}$ and $\mathbf{B}$.

To motivate this aim, observe that the Csiszár-Körner type inequality

$$\left(\sum_{k=1}^{n} T_k\right) f\left(\sum_{k=1}^{n} T_k\right)^{-}\left(\sum_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} T_k f T_k^{-}(B_k) \tag{5.103}$$

(see (5.93) in Theorem 5.3.13 can be restated as

$$\sum_{k=1}^{n} \bar{T} f\left(\bar{T}\right)^{-}\left(\bar{B}\right) \leq \sum_{k=1}^{n} T_k f T_k^{-}(B_k), \tag{5.104}$$

where $\bar{T} = \frac{T_1+\ldots+T_n}{n}$ and $\bar{B} = \frac{B_1+\ldots+B_n}{n}$.

In fact, we have

$$\sum_{k=1}^{n} \bar{T} f\left(\bar{T}\right)^{-}\left(\bar{B}\right) = \sum_{k=1}^{n} \left(\frac{\sum\limits_{k=1}^{n} T_k}{n}\right) f\left(\frac{\sum\limits_{k=1}^{n} T_k}{n}\right)^{-}\left(\frac{\sum\limits_{k=1}^{n} B_k}{n}\right)$$

$$= \sum_{k=1}^{n} \frac{1}{n}\left(\sum_{k=1}^{n} T_k\right) f\left(\sum_{k=1}^{n} T_k\right)^{-}\left(\sum_{k=1}^{n} B_k\right) = \left(\sum_{k=1}^{n} T_k\right) f\left(\sum_{k=1}^{n} T_k\right)^{-}\left(\sum_{k=1}^{n} B_k\right).$$

Notice that (5.104) is connected with the *simplest* joint majorization relation

$$\left(\begin{pmatrix}\bar{T}\\ \bar{B}\end{pmatrix}, \ldots, \begin{pmatrix}\bar{T}\\ \bar{B}\end{pmatrix}\right) \prec \left(\begin{pmatrix}T_1\\ B_1\end{pmatrix}, \ldots, \begin{pmatrix}T_n\\ B_n\end{pmatrix}\right),$$

which is induced by the doubly stochastic matrix $\mathbf{S} = \frac{1}{n}\mathbf{E}$, where

$$\mathbf{E} = \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ldots & \vdots \\ 1 & \ldots & 1 \end{pmatrix}$$

is the $n \times n$ matrix of all ones.

The above discussion suggests to try extend the Csiszár-Körner operator type inequality (5.103) to the form related to a *general* joint majorization relation

$$\left(\begin{pmatrix}R_1\\ A_1\end{pmatrix}, \ldots, \begin{pmatrix}R_n\\ A_n\end{pmatrix}\right) \prec \left(\begin{pmatrix}T_1\\ B_1\end{pmatrix}, \ldots, \begin{pmatrix}T_n\\ B_n\end{pmatrix}\right)$$

with positive operators $A_1, \ldots, A_n, B_1, \ldots, B_n$ in $\mathcal{B}$, and with positive linear maps $R_1, \ldots, R_n$, $T_1, \ldots, T_n$ from $\mathcal{A}$ to $\mathcal{B}$.

Lemma 5.3.1. *([132]):* *Let $f : [0, \infty) \to \mathbb{R}$ be an operator convex function. For a positive integer p and $l = 1, 2, \ldots, p$, let $0 < \alpha_l \in \mathbb{R}$ with $\sum_{l=1}^{p} \alpha_l = 1$, and*

$$\mathbf{T}_l = (T_{1l}, T_{2l}, \ldots, T_{nl}) \in J_1^n \quad \textit{and} \quad \mathbf{B}_l = (B_{1l}, B_{2l}, \ldots, B_{nl}) \in J_2^n,$$

where T_{kl} is a positive linear map from $\mathcal{A}$ to $\mathcal{B}$, and B_{kl} is a positive operator in $\mathcal{A}$, $k = 1, 2, \ldots, n$, $l = 1, 2, \ldots, p$.

If

$$B_{kl} \in \operatorname{Ran} T_{kl}\ , \quad k = 1, 2, \ldots, n, l = 1, 2, \ldots, p, \tag{5.105}$$

$$I \in \operatorname{Ran} \left(\sum_{l=1}^{p} \alpha_l T_{kl} \right)^{-} , \quad k = 1, 2, \ldots, n, \tag{5.106}$$

$$\operatorname{Ran} \left(\sum_{l=1}^{p} \alpha_l T_{kl} \right) = \sum_{l=1}^{p} \operatorname{Ran}(\alpha_l T_{kl})\ , \quad k = 1, 2, \ldots, n, \tag{5.107}$$

then

$$D_f \left(\sum_{l=1}^{p} \alpha_l \mathbf{T}_l, \sum_{l=1}^{p} \alpha_l \mathbf{B}_l \right) \leq \sum_{l=1}^{p} \alpha_l D_f(\mathbf{T}_l, \mathbf{B}_l), \tag{5.108}$$

i.e., the function $(\mathbf{T}, \mathbf{B}) \mapsto D_f(\mathbf{T}, \mathbf{B})$ is jointly $\leq$-convex on $(J_1 \times J_2)^n$.

If f is operator concave, then the right-hand side inequality in (5.108) is reversed.

Proof: It is evident that

$$\sum_{l=1}^{p} \alpha_l \mathbf{T}_l = \sum_{l=1}^{p} \alpha_l (T_{1l}, T_{2l}, \ldots, T_{nl}) = (\sum_{l=1}^{p} \alpha_l T_{1l}, \sum_{l=1}^{p} \alpha_l T_{2l}, \ldots, \sum_{l=1}^{p} \alpha_l T_{nl}),$$

$$\sum_{l=1}^{p} \alpha_l \mathbf{B}_l = \sum_{l=1}^{p} \alpha_l (B_{1l}, B_{2l}, \ldots, B_{nl}) = (\sum_{l=1}^{p} \alpha_l B_{1l}, \sum_{l=1}^{p} \alpha_l B_{2l}, \ldots, \sum_{l=1}^{p} \alpha_l B_{nl}).$$

Thus we find that

$$D_f \left(\sum_{l=1}^{p} \alpha_l \mathbf{T}_l, \sum_{l=1}^{p} \alpha_l \mathbf{B}_l \right) = \sum_{k=1}^{n} \left(\sum_{l=1}^{p} \alpha_l T_{kl} \right) f \left(\sum_{l=1}^{p} \alpha_l T_{kl} \right)^{-} \left(\sum_{l=1}^{p} \alpha_l B_{kl} \right).$$

By virtue of Theorem 5.3.13, we infer that

$$\left(\sum_{l=1}^{p}\alpha_l T_{kl}\right) f\left(\sum_{l=1}^{p}\alpha_l T_{kl}\right)^{-}\left(\sum_{l=1}^{p}\alpha_l B_{kl}\right) \leq \sum_{l=1}^{p}(\alpha_l T_{kl})f(\alpha_l T_{kl})^{-}(\alpha_l B_{kl}).$$

In consequence, we can write

$$D_f\left(\sum_{l=1}^{p}\alpha_l \mathbf{T}_l, \sum_{l=1}^{p}\alpha_l \mathbf{B}_l\right) \leq \sum_{k=1}^{n}\sum_{l=1}^{p}\alpha_l T_{kl} f T_{kl}^{-} B_{kl}$$

$$= \sum_{l=1}^{p}\alpha_l \sum_{k=1}^{n} T_{kl} f T_{kl}^{-} B_{kl} = \sum_{l=1}^{p}\alpha_l D_f(\mathbf{T}_l, \mathbf{B}_l),$$

as required. □

Remark 5.3.10.: In Lemma 5.3.1., if the positive linear map $\sum_{l=1}^{p}\alpha_l T_{lk}$ is invertible for any $k = 1, 2, \ldots, n$, then the conditions (5.106) and (5.107) are satisfied automatically. So, they can be dropped off.

Let $p = n!$. In the sequel $\mathcal{P}_n = \{\pi_l : l = 1, 2, \ldots, p\}$ is the permutation group of the index set $\{1, 2, \ldots, n\}$. For $l = 1, 2, \ldots, p$, the symbol $\mathbf{P}_l$ stands for the $n \times n$ permutation matrix generated by π_l. For this reason, if

$$\mathbf{T} = (T_1, T_2, \ldots, T_n) \in J_1^n \quad \text{and} \quad \mathbf{B} = (B_1, B_2, \ldots, B_n) \in J_2^n,$$

then

$$\mathbf{TP}_l = (T_{\pi_l(1)}, T_{\pi_l(2)}, \ldots, T_{\pi_l(n)}) \quad \text{and} \quad \mathbf{BP}_l = (B_{\pi_l(1)}, B_{\pi_l(2)}, \ldots, B_{\pi_l(n)}).$$

Theorem 5.3.16*([132])*: *Let $f : [0, \infty) \to \mathbb{R}$ be an operator convex function. Let*

$$\mathbf{T} = (T_1, T_2, \ldots, T_n) \in J_1^n \quad \textit{and} \quad \mathbf{R} = (R_1, R_2, \ldots, R_n) \in J_1^n,$$

$$\mathbf{B} = (B_1, B_2, \ldots, B_n) \in J_2^n \quad \textit{and} \quad \mathbf{A} = (A_1, A_2, \ldots, A_n) \in J_2^n,$$

where T_k and S_k are positive linear maps from $\mathcal{A}$ to $\mathcal{B}$, and B_k and A_k are positive *operators in* $\mathcal{A}$, $k = 1, 2, \ldots, n$.

If

$$\begin{pmatrix} \mathbf{R} \\ \mathbf{A} \end{pmatrix} \prec \begin{pmatrix} \mathbf{T} \\ \mathbf{B} \end{pmatrix}, \tag{5.109}$$

i.e.,

$$\mathbf{F} = \sum_{l=1}^{p} \alpha_l \mathbf{T}\mathbf{P}_l \quad \textit{and} \quad \mathbf{A} = \sum_{l=1}^{p} \alpha_l \mathbf{B}\mathbf{P}_l$$

for some coefficients $0 \le \alpha_l \in \mathbb{R}$ *with* $\sum\limits_{l=1}^{p} \alpha_l = 1$, $p = n!$, *and if in addition,*

$$B_k \in \operatorname{Ran} T_k\,, \quad k = 1, 2, \ldots, n, \tag{5.110}$$

$$I \in \operatorname{Ran} \left(\sum_{l=1}^{p} \alpha_l T_{\pi_l(k)} \right)^{-}, \quad k = 1, 2, \ldots, n, \tag{5.111}$$

$$\operatorname{Ran} \left(\sum_{l=1}^{p} \alpha_l T_{\pi_l(k)} \right) = \sum_{l=1}^{p} \operatorname{Ran}(\alpha_l T_{\pi_l(k)})\,, \quad k = 1, 2, \ldots, n, \tag{5.112}$$

then

$$D_f(\mathbf{R}, \mathbf{A}) \le D_f(\mathbf{T}, \mathbf{B}). \tag{5.113}$$

If f *is operator concave, then inequality (5.113) is reversed.*

Proof: It follows that the function $(\mathbf{T}, \mathbf{B}) \to D_f(\mathbf{T}, \mathbf{B})$ is jointly permutation-invariant on $(J_1 \times J_2)^n$, i.e., $D_f(\mathbf{TP}, \mathbf{BP}) = D_f(\mathbf{T}, \mathbf{B})$ for any $n \times n$ permutation matrix $\mathbf{P}$. Indeed, we have $\mathbf{TP} = (T_{\pi(1)}, T_{\pi(2)}, \ldots, T_{\pi(n)})$ and $\mathbf{BP} = (B_{\pi(1)}, B_{\pi(2)}, \ldots, B_{\pi(n)})$, where π is the permutation inducing the permutation matrix $\mathbf{P}$, and $(\pi(1), \pi(2), \ldots, \pi(n)) = \pi(1, 2, \ldots, n)$. Therefore,

$$D_f(\mathbf{TP}, \mathbf{BP}) = \sum_{k=1}^{n} T_{\pi(k)} f T_{\pi(k)}^{-}(B_{\pi(k)}) = \sum_{k=1}^{n} T_k f T_k^{-}(B_k) = D_f(\mathbf{T},\ \mathbf{B}),$$

as desired.

For $l = 1, 2, \ldots, p$ with $p = n!$, we define

$$\mathbf{T}_l = \mathbf{T}\mathbf{P}_l \quad \text{and} \quad \mathbf{B}_l = \mathbf{B}\mathbf{P}_l,$$

where $\mathbf{P}_l$ is the $n \times n$ permutation matrix induced by π_l.

By setting $T_{kl} = T_{\pi_l(k)}$ and $B_{kl} = B_{\pi_l(k)}$, from Lemma 5.3.1. we obtain

$$D_f\left(\sum_{l=1}^{p} \alpha_l \mathbf{T}_l, \sum_{l=1}^{p} \alpha_l \mathbf{B}_l\right) \leq \sum_{l=1}^{p} \alpha_l D_f(\mathbf{T}_l, \mathbf{B}_l).$$

Thus we see that

$$D_f(\mathbf{R},\mathbf{A}) \leq \sum_{l=1}^{p} \alpha_l D_f(\mathbf{T}_l,\mathbf{B}_l) = \sum_{l=1}^{p} \alpha_l D_f(\mathbf{T}\mathbf{P}_l,\mathbf{B}\mathbf{P}_l) = \sum_{l=1}^{p} \alpha_l D_f(\mathbf{T},\mathbf{B}) = D_f(\mathbf{T},\mathbf{B}),$$

and the proof of (5.113) is finished. □

Corollary 5.3.21. ***([132]):*** *Let $f : [0,\infty) \to \mathbb{R}$ be an operator convex function. Let*

$$\mathbf{T} = (T_1, T_2, \ldots, T_n) \in J_1^n \quad \textit{and} \quad \mathbf{R} = (R_1, R_2, \ldots, R_n) \in J_1^n,$$

$$\mathbf{B} = (B_1, B_2, \ldots, B_n) \in J_2^n \quad \textit{and} \quad \mathbf{A} = (A_1, A_2, \ldots, A_n) \in J_2^n,$$

where T_k and R_k are positive linear maps from $\mathcal{A}$ to $\mathcal{B}$, and B_k and A_k are positive operators in $\mathcal{A}$, $k = 1, 2, \ldots, n$.

If R_k is invertible and $B_k \in \operatorname{Ran} T_k$ for $k = 1, 2, \ldots, n$, and

$$\begin{pmatrix} \mathbf{R} \\ \mathbf{A} \end{pmatrix} \prec \begin{pmatrix} \mathbf{T} \\ \mathbf{B} \end{pmatrix}, \tag{5.114}$$

then

$$D_f(\mathbf{R},\mathbf{A}) \leq D_f(\mathbf{T},\mathbf{B}). \tag{5.115}$$

If f is operator concave, then inequality (5.115) is reversed.

Proof: From (5.114) we have

$$\mathbf{R} = \sum_{l=1}^{p} \alpha_l \mathbf{T}\mathbf{P}_l \quad \text{and} \quad \mathbf{A} = \sum_{l=1}^{p} \alpha_l \mathbf{B}\mathbf{P}_l$$

for some coefficients $0 \leq \alpha_l \in \mathbb{R}$ with $\sum\limits_{l=1}^{p} \alpha_l = 1$, $p = n!$.

It follows that $R_k = \sum\limits_{l=1}^{p} \alpha_l T_{\pi_l(k)}$, $k = 1, 2, \ldots, n$.

However $R_k : \mathcal{A} \to \mathcal{B}$ is invertible for $k = 1, 2, \ldots, n$, So, $\operatorname{Ran} R_k = \mathcal{B}$. Therefore,

$$\mathcal{B} = \operatorname{Ran} \sum_{l=1}^{p} \alpha_l T_{\pi_l(k)} \subset \sum_{l=1}^{p} \operatorname{Ran}(\alpha_l T_{\pi_l(k)}) \subset \mathcal{B}.$$

Hence we get

$$\operatorname{Ran} \left(\sum_{l=1}^{p} \alpha_l T_{\pi_l(k)} \right) = \sum_{l=1}^{p} \operatorname{Ran}(\alpha_l T_{\pi_l(k)}) \ , \quad k = 1, 2, \ldots, n.$$

In other words, condition (5.112) is met.

Simultaneously, $R_k^- : \mathcal{B} \to \mathcal{A}$ for $k = 1, 2, \ldots, n$. But $R_k : \mathcal{A} \to \mathcal{B}$ is invertible for $k = 1, 2, \ldots, n$. So, we find that $R_k^- = R_k^{-1}$ and $\operatorname{Ran} R_k^- = \mathcal{A}$. In this way, we have

$$I \in \mathcal{A} = \operatorname{Ran} R_k^- = \left(\sum_{l=1}^{p} \alpha_l T_{\pi_l(k)} \right)^- ,$$

which shows that condition (5.111) is fulfilled.

In conclusion, to complete the proof it is now sufficient to use Theorem 5.3.16. $\square$

Bibliography

[1] S. Abramovich, S. Banić and M. Matić, Superquadratic functions in several variables. J. Math. Anal. Appl. 327 (2007) 1444–1460

[2] S. Abramovich, J. Barić and J. Pečarić, Fejer and Hermite-Hadamard type inequalities for superquadratic functions, J. Math. Anal. Appl. 344 (2008) 1048–1056

[3] S. Abramovich, G. Jameson and G. Sinnamon, Inequalities for averages of convex and superquadratic functions, J. Inequal. Pure and Appl. Math. 5 (4) (2004) article 91

[4] S. Abramovich, G. Jameson and G. Sinnamon, Refining Jensen's inequality, Bull. Math. Sc. Math. Roum. 47 (2004) 3-14

[5] A. Abramovich, M. Klaričić Bakula, M. Matić and J. Pečarić, A variant of Jensen - Steffensen's inequality and quasi-arithmetic means, J. Math. Anal. Appl. 307 (2005) 370–386

[6] M. Adil Khan, S. Ivelić Bradanović and J. Pečarić, Generalizations of Sherman's inequality by Hermite's interpolating polynomial, Math. Inequal. Appl. 19 (4) (2016) 1181–1192

[7] M. Adil Khan, S. Ivelić Bradanović and J. Pečarić, On Sherman's type inequalities for n-convex function with applications, Konuralp J. Math. 4 (2) (2016) 255–270

[8] M. Adil Khan, J. Khan and J. Pečarić, Generalizations of Sherman's inequality by Montgomery identity and Green function, Electron. J. Math. Anal. Appl. 5 (1) (2017) 1-16

[9] M. Adil Khan, J. Khan and J. Pečarić, Generalizations of Sherman's theorem by Montgomery identity and new Green functions, Adv. Stud. Contemp. Math. 4 (27) (2017) 495-514

[10] R. P. Agarwal, S. Ivelić Bradanović and J. Pečarić, Generalizations of Sherman's inequality by Lidstone's interpolating polynomial, J. Inequal. Appl. 6 (2016) (2016)

[11] R. P. Agarwal and P. J. Y. Wong, Error Inequalities in Polynomial Interpolation and their Applications, Kluwer Academic Publishers, Dordrecht, 1993.

[12] A. Aglić Aljinović, Asif R. Khan and J. Pečarić, Weighted majorization theorems via generalization of Taylor's formula, Journal of Inequalities and Applications 2015 (2015) 1-22

[13] A. Aglić Aljinović, Asif R. Khan and J. Pečarić, Weighted majorization inequalities for n-convex functions via extension of Montgomery identity using Green function, Arab. J. Math. 7 (2018) 77-90

[14] N. I. Akhiezer, The Classical Moment Problem and Some Related Questions in Analysis, Oliver and Boyd, Edinburgh (1965)

[15] T. W. Anderson, An Introduction to Multivariate Statistical Analysis. Wiley, third ed., 2003.

[16] W. Arveson, A Short Course on Spectral Theory, Springer-Verlag, New York, Heidelberg, Berlin, (2002)

[17] A. Barbir, S. Ivelić Bradanović, Đ. Pečarić and J. Pečarić, Converse to the Sherman inequality with applications, Math. Inequal. Appl. 22 (4) (2019) 1405–1419

[18] N. Bebiano, R. Lamos and J. da Providencia, Inequalities for quantum relative entropy, Linear Algebra Appl. 401 (2005) 159-172

[19] E. F. Beckenbach, Superadditivity Inequalities, Pacific J. Math. 14 (1964), 421-438.

[20] P. R. Beesack, On the Green's function of an N-point boundary value problem, Pacific J. Math. 12 (1962) 801-812

[21] M. Belis, S. Guiaşu, A quantitative-qualitative measure of information in cybernetic systems, IEEE Trans. Inf. Theory. 14 (1968), 593-594

[22] S. N. Bernstein, Sur les fonctions absolument monotones, Acta Math. 52 (1929), 1-66

[23] R. Bhatia, Matrix Analysis. New York: Springer-Verlage, 1997.

[24] R. P. Boas, Representation of functions by Lidstone series. Duke Math. J. 10 (1943) 239-245

[25] J. Borcea, Equilibrium points of logarithmic potentials, Trans. Amer. Math. Soc. 359 (2007) 3209-3237

[26] W. W. Breckner and T. Trif, Convex Functions and Related Functional Equations: Selected Topics, Cluj University Press, Cluj, 2008.

[27] H. D. Brunk, On an inequality for convex functions, Proc. Amer. Math. Soc. 7 (1956) 817–824

[28] P. S. Bullen, Error estimates for some elementary quadrature rules, Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz. 602-633 (1978) 97–103

[29] I. Burbea and C. R. Rao, On the convexity of some divergence measures based on entropy functions, IEEE Trans. Inf. Theory. 28 (1982), 489-495

[30] M. Burgos, A. C. Márquez-Garciá and A. Morales-Campoy, Linear maps strongly preserving Moore-Penrose invertibility, Oper. Matrices 6 (4) (2012) 819-831

[31] A. M. Burtea, Two examples of weight majorization, Ann. Univ. Craiova Math. Comput. Sci. Ser. 32 (2) (2010) 92-99

[32] P. Cerone and S. S. Dragomir, Some new Ostrowski-type bounds for the Čebyšev functional and applications, J. Math. Inequal. 8 (1) (2014) 159-170

[33] I. Csiszár, Information-type measures of difference of probability functions and indirect observations, Studia Sci. Math. Hungar 2 (1967) 299-318

[34] I. Csiszár and J. Körner, Information Theory: Coding Theorem for Discrete Memoryless Systems, Academic Press, New York, 1981.

[35] G. Dahl, Matrix majorization, Linear Algebra Appl., 288 (1999), 53-73

[36] P. J. Davis, Interpolation and Approximation, Blaisdell, Boston, 1961.

[37] S. S. Dragomir, A new refinement of Jensen's inequality in linear spaces with applications, Math. Comput. Modelling 52 (2010) 1497-1505

[38] S. S. Dragomir, A refinement of Jensen's inequality with applications for f-divergence measures, Taiwanese J. Math. 14 (1) (2010)153-164

[39] S. S. Dragomir, Bounds for the normalised Jensen functional, Bull. Austral. Math. Soc. 74 (2006) 471-478

[40] S. S. Dragomir, Other inequalities for Csiszár divergence and applications, Preprint, RGMIA Monographs, Victoria University (2000)

[41] S. S. Dragomir, Upper and lower bounds for Csiszár f-divergence in terms of the Kullback-Leibler distance and applications, Inequalities for Csiszár f-Divergence in Information Theory, RGMIA Monographs, edited by S. S. Dragomir and T. M. Rassias, Victoria University, Australia, 2000

[42] S. S. Dragomir, J. E. Pečarić and L. E. Persson, Properties of some functionals related to Jensen's inequality, Acta Math. Hungar. 70 (1-2) (1996) 129–143

[43] E. G. Effros, A matrix convexity approach to some celebrated quantum inequalities, Proc. Natl. Sci. 106 (2009) 1006–1008

[44] L. Egghe and R. Rousseau, Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science, Elsevier Science Publishers, New York, 1990.

[45] W. Ehm, M. G. Genton and T. Gneiting, Stationary covariance associated with exponentially convex functions, Bernoulli 9 (4) (2003) 607-615

[46] A. El Farissi, Simple proof and refinement of Hermite-Hadamard inequality, J. Math. Inequal. 4 (3) (2010) 365–369

[47] A. M. Fink, Bounds of the deviation of a function from its avereges, Czechoslovak Math. J. 42 (117) (1992) 289-310

[48] J. I. Fujii and E. Kamei, Relative operator entropy in noncommutative information theory, Math. Japon. 34 (1989) 341-348

[49] J. Fujii, M. Fujii, M. Nakamura, J. Pečarić, and Y. Seo, A reverse inequality for the weighted geometric mean due to Lawson-Lim, Linear Algebra Appl. 427 (2007) 272–284.

[50] J. J. Fujii, M. Fujii and Y. Seo, An extension of the Kubo-Ando theory: Solidarities, Math. Japonica 35 (2) (1990) 387–396

[51] T. 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)

[52] S. Furuichi, K. Yanagi and K. Kuriyama, A note on operator inequalities of Tsallis relative operator entropy, Linear Algebra Appl. 407 (2005) 19–31

[53] T. Furuta, Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators, Linear Algebra Appl. 381 (2004) 219–235

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

[55] A. Gelbukh and G. Sidorov, Zipf and Heaps laws' coefficients depend on language, Lect. Notes Comput. Sci. 2001 (2004) 332–335

[56] R. Ger and K. Nikodem, Strongly convex functions of higher order, Advances in Nonlinear Analysis 74 (2011) 661–665

[57] V. L. Gontscharoff, Theory of Interpolation and Approximation of Functions, Gostekhizdat, Moscow (1954)

[58] P. C. Hammer, The midpoint method of numerical integration, Math. Mag. 31 (1957/1958) 193–195

[59] F. Hansen, J. Pečarić and I. Perić, Jensen's operator inequality and its converses, Math. Scan. 100 (2007) 61-73

[60] G. H. Hardy, J. E. Littlewood and G. Pólya, Inequalities, Cambridge University Press, London/New York, 1952

[61] G. H. Hardy, J. E. Littlewood, G. Pólya : Some simple inequalities satisfied by convex functions, Mess. of Math. 58 (1929) 145-152

[62] R. Harte and M. Mbekhta: On generalized inverses in C^* -algebras, Studia Math. 103 (1992) 71-77

[63] R. A. Horn and C. R. Johnson, Matrix Analysis. New York: Cambridge University Press, 1991.

[64] S. Ivelić Bradanović, More Accurate Majorization Inequalities Obtained Via Superquadraticity and Convexity with Application to Entropies, Mediterr. J. Math. 18 (79) (2021) 1-16

[65] S. Ivelić Bradanović, Sherman's inequality and its converse for strongly convex functions with applications to generalized f-divergences, Turk. J. Math. 43 (6) (2019) 2680-2696

[66] S. Ivelić Bradanović, N. Latif, Đ. Pečarić and J. Pečarić, Sherman's and related inequalities with applications in information theory, J. Inequal. Appl. 2018 (2018).

[67] S. Ivelić Bradanović, N. Latif, J. Pečarić, Generalizations of Sherman's inequality by Taylor's formula, J. Inequal. Spec. Funct. 8 (2) (2017) 18-30

[68] S. Ivelić Bradanović, N. Latif, J. Pečarić, Generalizations of Sherman's Inequality Via Fink's Identity and Green's Function, Ukr. Math. J. 70 (8) (2019) 1192-1204

[69] S. Ivelić Bradanović, N. Latif and J. Pečarić, On an upper bound for Sherman's inequality. J. Inequal. Appl. 2016 (2016)

[70] S. Ivelić Bradanović, J. Mićić and J. Pečarić, Sherman's operator inequality, J. Math. Inequal. 15 (2) (2021) 675-699

[71] S. Ivelić Bradanović, Đ. Pečarić and J. Pečarić, n-convexity and weighted majorization with applications to f -divergences and Zipf-Mandelbrot law, Period. Math. Hung. 90 (2025), 57-77

[72] S. Ivelić Bradanović, Đ. Pečarić and J. Pečarić, Sherman's functional, its properties with applications for f-divergence measure, Math. Inequal. Appl. 2 (26) (2023) 351–368

[73] S. Ivelić Bradanović and J. Pečarić, Generalizations of Sherman's inequality, Per. Math. Hung. 74 (2) (2017) 197-219

[74] S. Ivelić Bradanović and J. Pečarić, Extensions and improvements of Sherman's and related inequalities for n -convex functions, Open Math. 15 (1) (2017), 936–947

[75] H. Isa, M. Ito, E. Kamei, H. Tohyama and M. Watanabe: Generalizations of operator Shanon inequality based on Tsallis and Rényi relative entropies, Linear Algebra Appl. 439 (10) (2013) 3148-3155

[76] J. Jakšetić, J. Pečarić, A. Perušić, Steffensen inequality, higher order convexity and exponential convexity , Rend. Circ. Mat. Palermo. 63 (1) (2014) 109-127

[77] J. Jakšetić and J. Pečarić, Exponential convexity method, J. Convex Anal. 20 (1) (2013) 181-197

[78] J. L. W. V. Jensen, Sur les fonctions convexes et les inegalites inégalités entre les valeurs moyennes, Acta Mathematica. 30 (1) (1906) 175-193

[79] B. Jessen, Bemaerkinger om konvekse Funktioner og Uligheder imellem Middelvaerdier. I, Mat. Tidsskrift B (1931) 17-28

[80] J. H. Justice, Maximum Entropy and Bayssian Methods in Applied Statistics, Cambridge University Press, Cambridge, 1986.

[81] J. N. Kapur, On the roles of maximum entropy and minimum discrimination information principles in Statistics, Technical Address of the 38th Annual Conference of the Indian Society of Agricultural Statistics, 1984, 1-44

[82] J. N. Kapur, A comparative assessment of various measures of directed divergence, Advances in Management Studies, 3 (1) (1984) 1-16

[83] J. N. Kapur, Maximum-Entropy Models in Science and Engineering, NY, USA: John Wiley and Sons, 1989.

[84] S. Karlin, Total Positivity, Stanford Univ. Press, Stanford, 1968.

[85] R. Kaur, M. Singh and J. S. Aujla, Generalized matrix version of reverse Hölder inequality, Linear Algebra Appl. 434 (2011) 636-640

[86] M. Khosravi, J. S. Aujla, S. S. Dragomir and M. S. Moslehian, Refinements of Choi-Davis-Jensen's inequality, Bull. Math. Anal. Appl. 3 (2) (2011) 127–133

[87] M. Klaričić Bakula and K. Nikodem, On the converse Jensen inequality for strongly convex functions, J. Math. Anal. Appl. 434 (2016) 516-522

[88] M. Klaričić-Bakula, J. Pečarić, J. Perić, Extensions of the Hermite-Hadamard inequality with applications, Math. Inequal. Appl. 12 (4) (2012) 899–921

[89] M. Klaričić Bakula, J. Pečarić and M. Ribičić, Companion inequalities to Jensen's inequality for m-convex and α, m-convex functions, J. Ineq. Pure & Appl. Math. 7 (5) (2006) Article 194.

[90] P. Kluza and M. Niezgoda, On Csiszár and Tsallis type f-divergences induced by superquadratic and convex functions, Math. Inequal. Appl. 21 (2) (2018) 455-467

[91] A. N. Kolmogorov and S. V. Fomin, Elements of the theory of functions and functional analysis, Graylock, 1957-1961

[92] M. Krnić, N. Lovričević and J. Pečarić, Jessen's functional, its properties and applications, An. St. Univ. Ovidius Constanta 20 (1) (2012) 225-248

[93] M. Krnić, N. Lovričević and J. Pečarić, Jensen's Operator and Applications to Mean Inequalities for Operators in Hilbert Space, Bull. Malays. Math. Sci. Soc. 35 (1) (2012) 1–14

[94] M. Krnić, N. Lovričević, J. Pečarić, Multidimensional Jensen's operator on a Hilbert space and applications, Linear Algebra Appl. 436 (7) (2012) 2583–2596

[95] M. Krnić, N. Lovričević and J. Pečarić, On some properties of Jensen-Mercer's functional, J. Math. Inequal. 6 (1) (2012) 125-139

[96] M. Krnić, N. Lovričević, J. Pečarić and J. Perić, Superadditivity and monotonicity of the Jensen-type functionals (New methods for improving the Jensen-type inequalities in real and in operator cases), Zagreb, Element, 2015.

[97] F. Kubo and T. Ando, Means of positive linear operators, Math. Ann. 246 (1980) 205–224

[98] S. Kullback, Information Theory and Statistics, J. Wiley, New York, 1959

[99] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1) (1951) 79–86

[100] P. Lah and M. Ribarič, Converse of Jensen's inequality for convex functions, Publ. Elektroteh. fak. Univ. Beogr., Ser. Mat. fiz. 412–460 (1973) 201–205

[101] N. Latif, J. Pečarić and I. Perić, On Discrete Favard and Berwald's Inequalities, Commun. Math. Anal. 12 (2) (2012) 34-57

[102] Z. Latreuch, B. Belaïdi, New inequalities for convex sequences with applications, Int. J. Open Problems Comput. Math. 5 (3) (2012) 15–27

[103] G. J. Lidstone, Notes on the extension of Aitken's theorem (for polynomial interpolation) to the Everett types. Proc. Edinb. Math. Soc. 2 (2) (1929) 16-19

[104] G. Maksa and Z. Páles, The equality case in some recent convexity inequalities, Opuscula Math. 31 (2) (2011) 269–277

[105] B. Mandelbrot, Information theory and psycholinguistics: a theory of words frequencies, in: P. Lazafeld, N. Henry (Eds.), Readings in Mathematical Social. Science, (MIT Press, Cambridge, MA, 1966)

[106] A. W. Marshall, I. Olkin and C. Barry Arnold, Inequalities: Theory of Majorization and Its Applications (Second Edition), NY, USA: Springer Series in Statistics, 2011.

[107] A. M. Mercer, A variant of Jensen's inequality, J. Ineq. Pure and Appl. Math. 4 (4) (2003) Article 73

[108] B. Micherda and T. Rajba, On some Hermite-Hadamard-Fejér inequalities for (k, h)-convex functions, Math. Inequal. Appl. 15 (4) (2012) 931–940

[109] J. Mićić, Z. Pavić and J. Pečarić, Some better bounds in converses of the Jensen operator inequality, Oper. Matrices 6 (3) (2012) 589–605

[110] N. Minculete, F.-C. Mitroi, Fejér-type inequalities, Austral. J. Math. Anal. Appl. 9 (1) (2012) 1-8

[111] D. S. Mitrinović, J. E. Pečarić, A. M. Fink, Classical and New Inequalities in Analysis, Kluwer Academic Publishers, Dortrecht-Boston-London, 1993.

[112] D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, Inequalities for functions and their Integrals and Derivatives, Kluwer Academic Publishers, Dordrecht, 1994.

[113] M. Mohan, J. Mitter, On bounds of 'useful' information measure, Indian J. Pure Appl. Math. 9 (9) (1978) 960-964

[114] B. Mond and J. Pečarić, Converses of Jensen inequality for several operators, Rev. Anal. Numér. Théor. Approx. 23 (2) (1994) 179-183

[115] M. A. Montemurro, Beyond the Zipf-Mandelbrot law in quantitative linguistics, Physica A 300 (3-4) (2001) 567–578

[116] M. S. Moslehian and M. Kian, Non-commutative f -divergence functional, Math. Nachr. 286 (14-15) (2013) 1514-1529

[117] D. Mouillot nad A. Lepretre, Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity, Environ. Monit. Assess. 63 (2) (2000) 279-295

[118] C. P. Niculescu and L. E. Persson, Convex Functions and Their Applications. A contemporary Approach, CMS Books in Mathematics 23, Springer-Verlage, New York, 2006.

[119] C. P. Niculescu and L. E. Persson, Old and New on the Hermite-Hadamard inequality, Real Anal. Exch. 29 (2) (2004) 663–685

[120] M. Niezgoda, An extension of results of A. McD. Mercer and I. Gavrea, J. Inequal. Pure Appl. Math. 6 (4) (2005) Art. 107

[121] M. Niezgoda, A generalization of Mercer's result on convex functions, Nonlinear Anal. Theory Methods Appl. 71 (7-8) (2009) 2771-2779

[122] M. Niezgoda, Inequalities for convex sequences and nondecreasing convex functions, Aequat. Math. 91 (1) (2017) 1-20

[123] M. Niezgoda, Inequalities for H-invex Functions with Applications for Uniformly Convex and Superquadratic Functions, Filomat 31 (15) (2017) 4781–4794

[124] M. Niezgoda, Jessen's functional and majorization, Math. Inequal. Appl. 18 (3) (2015) 1003-1011

[125] M. Niezgoda, Jointly subadditive mappings induced by operator convex functions, Math. Inequal. Appl. 18 (1) (2015) 169-184

[126] M. Niezgoda, Nonlinear Sherman-type inequalities, Adv. Nonlinear Anal. 9 (1) (2020), 168–175

[127] M. Niezgoda, On $(k, h; m)$-convex mappings and applications , Math. Inequal. Appl. 17 (2) (2014) 665-678

[128] M. Niezgoda, On Shannon like inequalities for f-connections of positive linear map and positive operators, Linear Algebra Appl. 481 (2015) 186-201

[129] M. Niezgoda, On sub- and superadditive vector-valued maps with applications to group majorization, Linear Algebra Appl. 438 (11) (2013) 4249-4259

[130] M. Niezgoda, Remarks on Sherman like inequalities for (α, β)-convex functions, Math. Ineqal. Appl. 17 (4) (2014) 1579–1590

[131] M. Niezgoda, Sherman, Hermite-Hadamard and Fejér like Inequalities for Convex Sequences and Nondecreasing Convex Functions, Filomat 31 (8) (2017) 2321-2335

[132] M. Niezgoda, Vector joint majorization and generalization of Csiszar-Korner's inequality for f-divergence, Discrete Appl. Math. 198 (2016) 195-205

[133] M. Niezgoda, Vector majorization and Schur-concavity of some sums generated by the Jensen and Jensen-Mercer functionals, Math. Inequal. Appl. 18 (2) (2015) 769-786

[134] K. Nikodem, On Strongly Convex Functions and Related Classes of Functions, Handbook of Functional Equations. NY, USA: Springer (2014) 365-405

[135] K. Nikodem and Z. Páles, Characterizations of inner product spaces by strongly convex functions, Banach J. Math. Anal. 5 (1) (2011) 83–87

[136] C. E. M. Pearce and A. M. Rubinov, P-functions, quasi-convex functions and Hadamard-type inequalities, J. Math. Anal. Appl. 240 (1) (1999) 92–104

[137] J. Pečarić and J. Perić, Improvement of the Giaccardi and the Petrović inequality and related Stolarsky type means, An. Univ. Craiova Ser. Mat. Inform. 39 (1) (2012) 65-75

[138] J. E. Pečarić, F. Proschan and Y. L. Tong, Convex Functions, Partial Orderings, and Statistical Applications, Academic Press, New York, 1992.

[139] R. Pereira and S. Plosker, Dirichlet polynomials, majorization, and trumping, J. Phys. A. 46 (2015) 80–86

[140] B. T. Polyak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, Soviet mathematics-Doklady. 7 (1966) 72–75.

[141] T. Popoviciu, Sur l'approximation des fonctions convexes d'ordre superier, Mathematica 10 (1934) 49-54

[142] H. Poritsky, On certain polynomial and other approximations to analytic functions. Trans. Am. Math. Soc. 34 (1932), 274-331

[143] C. R. Rao and S. M. Mitra, Generalized inverse of a matrix and its applications, Proc. Sixth Berkeley Symp. on Math. Statist. and Prob. 1, Univ. of Calif. Press (1972) 601-620

[144] A. Rényi, On measures of entropy and information, Proc. Fourth Berkeley Symp. Math. Statist. Prob. (1), Univ. of Calif. Press (1961) 547–561

[145] W. Rudin, Real and Complex Analysis, McGraw-Hill, London, 1970.

[146] J. L. Schiff, The Laplace Transform. Theory and Applications, Undergraduate Texts in Mathematics, Springer, New York (1999).

[147] I. Schur, Über eine Klasse von Mittelbildungen mit Anwendungen die Determinanten-Theorie Sitzungsber, Berlin. Math. Gesellschaft 22 (1923) 9-20.

[148] C. E. Shannon and W. Weaver, The Mathemtiatical Theory of Comnunication, Urbana, University of Illinois Press, 1949..

[149] S. Sherman, On a theorem of Hardy, Littlewood, Pólya and Blackwell, Proc. Nat. Acad. Sci. 37 (12) (1957) 826-831

[150] Z. K. Silagadze, Citations and the Zipf-Mandelbrot Law, Complex Systems 11 (1997) 487-499

[151] A. Srivastava, Some new bounds of weighted entropy measures, Cybern. Inf. Technol. 11 (3) 60-65

[152] K. B. Stolarsky, Generalizations of the logarithmic mean, Math. Mag. 48 (1975) 87-92

[153] K. B. Stolarsky, The power and generalized logarithmic means, Amer. Math. Monthly 87 (1980) 545-548

[154] I. J. Taneja, L. Pardo, D. Morales and M. L. Menéndez, On generalized information and divergence measures and their applications: A brief review, Qüstiió 13 (1,2,3) (1989) 47-73.

[155] S. Varošanec, On h-convexity, J. Math. Anal. Appl. 326 (2007) 303–311

[156] J. M. Whittaker, Interpolation Function Theory, Cambridge, 1935.

[157] D. V. Widder, Completely convex function and Lidstone series, Trans. Am. Math. Soc. 51 (1942) 387-398

[158] D. V. Widder, The Laplace Transform, Princeton Univ. Press, New Jersey, 1941.

[159] S. Wu and L. Debnath, Inequalities for convex sequences and their applications, Comput. Math. Appl. 54 (2007) 525–534

[160] K. Yanagi, K. Kuriyama and S. Furuichi, Generalized Shannon inequalities based on Tsallis relative operator entropy, Linear Algebra Appl. 394 (2005) 109-118.

[161] A. Yu. Levin, Some problems bearing on the oscillation of solutions of linear differential equations, Soviet Math. Dokl. 4 (1963), 121-124

[162] G. K. Zipf, Selected Studies of the Principle of Relative Frequency in Language. Cambridge, MA: Harvard University Press. 1932.

[163] L. Zou and Y. Jiang, Improved arithmetic-geometric mean inequality and its application, J. Math. Inequal. 9 (1) (2015) 107–111

Index

Synopsis of the Volume

As a cornerstone of the **Pečarić PSR Series on Mathematical Inequalities**, this treatise offers an exhaustive analytical framework for Sherman's inequality. It bridges profound mathematical theory with sophisticated computational methodologies, addressing complex paradigms in theoretical physics, intelligent modeling, and information theory. This volume is curated to empower researchers and graduate students with advanced insights into optimization algorithms and fractional differential equations.

Profiles of the Authors

Slavica Ivelić Bradanović: An esteemed scholar at the University of Split, Croatia, whose research portfolio encompasses mathematical inequalities, Jensen-type operators, and information theory.

Marek Niezgoda: A prominent mathematician recognized for pioneering contributions to convex analysis, majorization theory, and the structural refinement of functional inequalities.

Đilda Pečarić: A faculty member at University North, Croatia, specializing in entropy-related applications, the Zipf-Mandelbrot law, and advanced bibliometric visualization.

Josip Pečarić: A distinguished Fellow of the Croatian Academy of Sciences and Arts and a global authority in the field of mathematical inequalities and convexity.

ISBN: 978-627-7623-07-4

978-627-7623-07-4

PSR Press
Ptolemy Scientific Research Press
Lahore, Pakistan