UNIVERSITATEA DIN CRAIOVA

ȘCOALA DOCTORALĂ DE ȘTIINȚE

DOMENIUL MATEMATICĂ

TEZĂ DE DOCTORAT

Conducător de doctorat:
Prof. univ. dr. Rovenţa Ionel

Doctorand:
Lăchescu Geanina Maria

CRAIOVA
2024

UNIVERSITATEA DIN CRAIOVA
ȘCOALA DOCTORALĂ DE ȘTIINȚE
DOMENIUL MATEMATICĂ

TEZĂ DE DOCTORAT

Convex analysis and majorization theory
(Analiza convexă și teoria majorizării)

Conducător de doctorat:
Prof. univ. dr. Rovenţa Ionel

Doctorand:
Lăchescu Geanina Maria

CRAIOVA
2024

Contents

Acknowledgments

I would first give my greatest thanks to my supervisor, Professor Ionel Rovenţa, to provide the opportunity of pursuing my PhD throughout 3 years with an unforgettable collaboration. I am profoundly thankful to him for his unwavering support, encouragement, and understanding throughout the ups and downs of this pursuit.

I am equally thankful to the members of my thesis committee of the Doctoral School in Mathematics, for their constructive feedback and essential suggestions that enhanced the quality of my work.

A part of this work in my thesis has been carried out in collaboration with other researchers. I would like to give my gratitude to Professor Emeritus Constantin P. Niculescu for his collaboration and support. I would also sincerely thank to Lector Dr. Maria Mălin for her colaboration and support in my research activity.

My appreciation also goes to the faculty and staff in the Department of Mathematics from University of Craiova, whose resources and assistance have been invaluable.

I would like to thank my family for giving me the opportunity to achieve this milestone and their constant support and push to challenge myself and bring out the best in me. Especially to my Dad thank you... this whole work is dedicated to you. You have raised me up to more than I can be.

Finally I want to mention that the research work within this doctoral thesis has been supported by a grant of the Romanian Ministry of Research, Innovation and Digitalization (MCID), project number 22 - Nonlinear Differential Systems in Applied Sciences, within PNRR-III-C9-2022-I8.

Chapter 1

Introduction, preliminaries and main results

The general aim of the present doctoral thesis is to strengthen the role of Convex Functions Theory (pointed out by majorization theory), as an important link between Mathematics, Engineering and Computer Science. We also emphasize the powerful interdisciplinary role of convexity, having as main tools optimization ideas and methods combined in various ways.

The main topic we are interested in this thesis is given by the link between convex analysis and majorization theory, described in terms of new concepts and refined majorization convex type inequalities. More precisely, we present Jensen's, Hardy-Littlewood-Polya's and Sherman's type inequalities for new type of weakly or strongly convex functions, perturbed by homogeneous symmetric polynomials of even degree. We manage to prove that the behaviour of homogeneous symmetric polynomials is similar to the one of euclidean norms. Moreover, the main novelty here is given by the possibility to extend all the above mentioned inequalities for nonpositive weights in $\mathbb{R}^n$, or even in spaces with curved geometry. These extensions in more general spaces can be done via majorization arguments and based on the extension of the barycenter concept for Steffensen-Popoviciu measures (where the weights are allowed to be nonpositive).

We consider that the research topic of this doctoral thesis becomes over the years an important field of research, due to the necessity to understand and optimize different processes/problems which use convex analysis tools, in order to be applied in different areas of research [28, 54, 68, 116]. In other words, our particular aim is to consolidate a theoretical foundation for studying optimization problems from an applied point of view, such as modeling communication networks and design of communication systems.

The concept of majorization appears in 1905, when Max Lorenz propose a graphical way to model the social differences in a finite population. Later on, Dalton (1920) and Hardy-Littlewood-Polya (1927, 1934), reveal some optimization properties, which led to the notion of Schur-convex function. Applications of majorization in 4G communications networks, are related to data transmission rates with huge dimensions, where the interferences between different links create a strangulations of data transmission rates. An important amelioration was obtained in [11, 67], where the optimal power distribution is studied as a nonlinear optimization problem, non-convex with constraints. The problem was solved by the identification of a Schur-convex structure in the objective function. It can be shown that the optimal power allocation is binary, in a sense that, the data are sent with maximal power or the data transmission is not allowed (which goes back to switching or bang-bang controls strategy). These results obtained in the field of convex analysis are focused on majorization inequalities and can be used to obtain new optimal operating principles for some communications devices as intelligent traffic lights.

The second approach deals with weakly/strongly majorization/convexity concepts in the context of metric spaces with global non-positive curvature (namely global NPC spaces). Besides Hilbert spaces and manifolds, other important examples of global NPC spaces are the Bruhat-Tits buildings [17, 163] (in particular, the trees). It is important to mention that, in [119, 121, 126], Ky-Fan's inequality, Schauder's and Schaeffer's fixed point theorems and Hardy-Littlewood-Polya's majorization theorem have been extended in the context of global NPC spaces [17, 28, 80]. A new type of weak majorization was also discussed in [150].

The subject of majorization in global NPC spaces was successfully studied using some ideas inspired from articles [26, 92, 98, 111]. Applying different kind of majorization concepts (see [150]), for instance, to the trees, we could obtain a feasible model for the optimal distribution in high performances communication networks.

Other significant idea in this area presented in this thesis, is given by the possibility to introduce a new weaker concept of "point of convexity with nonpositive weights" , inspired by the notion of point of convexity introduced in [118]. Early references can be found in [47, 52, 59, 114, 127, 138, 139, 161]. Our aim is to use the notion of "point of convexity with nonpositive weights" in such a way to prove different type of convex inequalities for weaker assumptions, even in the context of global NPC spaces. Finally, note that some weaker or other generalized convexities were successfully used in the study of existence and uniqueness of solutions of partial differential equations. As applications, we mention that in order to establish a sufficient condition for the existence of finite time blow-up solutions for an evolutionary problem, arising naturally in mechanics, biology and population dynamics, in [122, 123], we have successfully used a class of generalized convex functions. See [21, 34, 35, 38, 59].

In the following sentences (of Chapter I) we briefly present the content of each chapter, where we announce the main results of this thesis. Moreover, we are also focused to give some theoretical background, in order to offer a general view and a good understanding of the whole thesis.

The first part of Chapter II is an introductory one and is mainly inspired from [68, 69, 104]. In this part, we present some notions related to convexity, majorization theory and inequalities associated to it. In fact, we recall the main properties of the relation of majorization, which was introduced by G. H . Hardy, J. E. Littlewood and G. Pólya [69] in 1929, and was popularized by their well-known book on inequalities [68]. For other details we also refer to the recent book by A. W. Marshall, I. Olkin and B. Arnold [104].

The second part of Chapter II is based on the paper *G. M. Lăchescu, M. Malin and I. Rovenţa, New Versions of Uniformly Convex Functions via Quadratic Complete Homogeneous Symmetric Polynomials, Mediterranean Journal of Mathematics **20**, 279 (2023).*

In this part we present new versions of uniformly convex functions, namely h_d strongly/weaker convex functions. In other words, we introduce stronger and weaker versions of uniformly convexity for which we recover well-known convex type inequalities such as: Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities. Sherman's and Ingham's type inequalities are also discussed.

More precisely, the topic we address is related to the study of a new family of convex functions which is based on the positivity property of complete homogeneous symmetric polynomials with even degree. The positivity of symmetric polynomial functions was firstly studied in an old paper of Hunter [74]. Later on, in [165] a different way to establish the positivity of such polynomials was considered. In addition, two different ideas are presented in [153], based on a Schur-convexity argument or on a method with divided differences. Note also that, the previous strategies was used to obtain fine estimates on the norms on complex matrices induced by complete homogeneous symmetric polynomials. See [4] and [37].

The family of complete homogeneous symmetric polynomials with n real variables $x_1, \dots, x_n$ and degree $d \in \mathbb{N}$ is given as follows

$$h_0(x_1, \dots, x_n) = 1,$$
$$h_d(x_1, \dots, x_n) := \sum_{1 \le i_1 \le \cdots \le i_d \le n} x_{i_1} \cdots x_{i_d} \qquad (d \ge 1).$$

The main strategy used to prove the positivity of h_d, for all even degrees $d \ge 2$, consists of using Schur-convexity and majorization arguments. Note that, the concept of majorization is a powerful topic of research with several interesting applications: a necessary and sufficient condition for a linear map to preserve group majorizations [131]; properties on superquadratic functions related to Jensen–Steffensen's inequality [1]; other majorization properties [83, 132]. Moreover, we notice that the possibility to define the concept of majorization into the spaces of curved geometry was confirmed in [126]. More results on this topic can be found in [112, 113, 114, 117, 124].

In order to present the current settings we address in this part of the thesis, let us introduce the concepts of stronger and weaker h_d convexity for functions defined on $\mathbb{R}^n$. A positivity result given in [153] asserts that: if $d \ge 2$ is an even natural index, then

$$h_d(x_1, x_2, \dots, x_n) \ge 0 \qquad (x_1, \dots, x_n \in \mathbb{R}). \tag{1.0.1}$$

Using ideas from (1.0.1) we define a new concept of convex function, as a perturbed of convex function with a complete homogeneous symmetric polynomial.

Definition 1.0.1. *Let $C > 0$ and let $d \ge 2$ be an even natural number. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be h_d strongly convex with modulus C if the function $f(\cdot) - C\, h_d(\cdot)$ is convex. Similarly, a function $f : \mathbb{R}^n \to \mathbb{R}$ is called h_d weakly convex with modulus C if the function $f(\cdot) + C\, h_d(\cdot)$ is convex.*

In order to motivate the concept of h_2 strongly/weakly convex function we recall the related notion of uniformly convex function.

Definition 1.0.2. *Let $C > 0$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be uniformly convex with modulus C if $f(\cdot) - C \left\|\cdot\right\|^2$ is convex. Equivalently, the function f is uniformly convex with modulus C if and only if the following inequality holds*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \le (1-\lambda) f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda(1-\lambda) \left\|\mathbf{x} - \mathbf{y}\right\|^2, \tag{1.0.2}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

It is natural to show that (1.0.2) holds similarly, even in the context of h_2 strongly convexity. Moreover, our objective is to study the general and difficult case, i.e. h_d strongly convexity, for any even natural number $d \geq 2$. We also use fine estimates in order to get h_d versions of Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities. Other classical inequalities are also obtained, which certifies that the family of h_d strongly convex functions lead to new ideas of further research. We strongly consider that the new concept and results presented in this part of this chapter can be used to establish connections and further applications related to other important scientific achievements in literature (see [2, 3, 15, 96, 133, 169]).

More precisely, we obtain an inequality related to (1.0.2), in the case of h_2 strongly convex functions.

Proposition 1.0.1. *Let $C > 0$. Then, the function $f : \mathbb{R}^n \to \mathbb{R}$ is h_2 strongly convex with modulus C if and only if*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda(1-\lambda)h_2(\mathbf{x}-\mathbf{y}), \tag{1.0.3}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

In the general case, for any even natural number $d \geq 2$, we get a nice extension of Proposition 1.0.1.

Theorem 1.0.1. *Let $C > 0$ and let $d \geq 2$ be an even natural number. Then, the function $f : \mathbb{R}^n \to \mathbb{R}$ is h_d strongly convex with modulus $C > 0$ if and only if*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}}h_d(\mathbf{x}-\mathbf{y}), \tag{1.0.4}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Moreover, for each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$ we have

$$h_d((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) - (1-\lambda)h_d(\mathbf{x}) - \lambda h_d(\mathbf{y}) \leq -\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}}h_d(\mathbf{x}-\mathbf{y}). \tag{1.0.5}$$

It is worth mentioning that, even if h_d polynomials cannot itself induce a norm (for example, in majorization settings, we have that, for any two vectors satisfying $x \prec y$, $h_2(y) \geq h_2(x) + h_2(y-x)$, see Lemma 2.2.1) we can introduce some polynomial norms. That means to introduce the norms that are the d^{th} root of a h_d polynomials. For more details, see the last chapter of this doctoral thesis, devoted to some perspectives.

For the convenience of the reader, in the following sentences we present some basic theoretical facts about strongly convex functions with modulus $C > 0$. Using our estimates from (1.0.1) and (1.0.3) we recover some well-known classical results within uniform convex functions theory. More details can be found in [97].

Theorem 1.0.2. *Let $C > 0$, $d \geq 2$ an even natural number and let $f : \Omega \to \mathbb{R}$ be a h_d strongly convex with modulus C defined on a convex set $\Omega \subseteq \mathbb{R}^n$. Then the following statements hold true:*

(i) If Ω is an open set, then f is a continuous function on Ω.

(ii) Any local minimizer of f is a global minimum for f.

(iii) Moreover, the global minimizer of f is unique.

In a similar way, we extend the notion of elliptic differentiable functions.

Definition 1.0.3. *We say that a function $J : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is elliptic (α-elliptic) if it is differentiable on Ω and there exists an $\alpha > 0$ such that*

$$\langle \nabla J(\mathbf{x}) - \nabla J(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \alpha \|\mathbf{x} - \mathbf{y}\|^2 \qquad (\mathbf{x}, \mathbf{y} \in \Omega). \tag{1.0.6}$$

Definition 1.0.4. *We say that a function $J : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is h_2-elliptic if it is differentiable on Ω with modulus C if*

$$\langle \nabla J(\mathbf{x}) - \nabla J(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq C h_2(x - y) \qquad (\mathbf{x}, \mathbf{y} \in \Omega). \tag{1.0.7}$$

By considering the convex function $g : \Omega \to \mathbb{R}$, where $g(\mathbf{x}) = J(\mathbf{x}) - C h_2(\mathbf{x})$, and using in addition the well-known convex inequality

$$g(\mathbf{y}) \geq g(\mathbf{x}) + \langle \nabla g(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \qquad (\mathbf{x}, \mathbf{y} \in \Omega),$$

we can get easily the following two results.

Theorem 1.0.3. *Let $J : \Omega \to \mathbb{R}$ be a differentiable function defined on the convex set $\Omega \subseteq \mathbb{R}^n$. Then the following affirmations are equivalent:*

(i) J is h_2 strongly convex with modulus C.

(ii) The following inequality holds true

$$J(\mathbf{x}) - J(\mathbf{y}) \geq \langle \nabla J(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle + C h_2(\mathbf{x} - \mathbf{y}) \qquad (\mathbf{x}, \mathbf{y} \in \Omega). \tag{1.0.8}$$

(iii) J is h_2-elliptic on Ω with modulus $2C$, i.e.

$$\langle \nabla J(\mathbf{x}) - \nabla J(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq 2C h_2(\mathbf{x} - \mathbf{y}) \qquad (\mathbf{x}, \mathbf{y} \in \Omega).$$

Theorem 1.0.4. *If $U \subseteq \mathbb{R}^n$ is a nonempty, closed and convex set, and J is h_2 strongly convex with modulus $C > 0$, then there exists an unique $\mathbf{x} \in U$ such that*

$$J(\mathbf{x}) = \min_{\mathbf{y} \in U} J(\mathbf{y}). \tag{1.0.9}$$

On the other hand, it is natural to present some remarks concerning the possibility of defining a scalar product in terms of h_d symmetric polynomials. Thus, defining the map $\langle \cdot, \cdot \rangle_h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as follows

$$\langle \mathbf{x}, \mathbf{y} \rangle_h = \frac{h_2(\mathbf{x} + \mathbf{y}) - h_2(\mathbf{x} - \mathbf{y})}{4} \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n), \tag{1.0.10}$$

a straightforward computation gives

$$\langle \mathbf{x}, \mathbf{y} \rangle_h = \langle \mathbf{x}, \mathbf{y} \rangle + \frac{1}{2} \left(\sum_{i=1}^n x_i \sum_{i=1}^n y_i - \sum_{i=1}^n x_i y_i \right),$$

where $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ denotes the usual scalar product in $\mathbb{R}^n$.

Notice that $\langle \mathbf{x}, \mathbf{y} \rangle_h$ satisfies the properties needed for a scalar product, i.e.

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle_h &= \langle \mathbf{y}, \mathbf{x} \rangle_h && (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n), \\ \langle \alpha\mathbf{x}, \mathbf{y} \rangle_h &= \alpha\langle \mathbf{x}, \mathbf{y} \rangle_h && (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \alpha \in \mathbb{R}), \\ \langle \mathbf{x}+\mathbf{z}, \mathbf{y} \rangle_h &= \langle \mathbf{x}, \mathbf{y} \rangle_h + \langle \mathbf{z}, \mathbf{y} \rangle_h && (\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n), \\ \langle \mathbf{x}, \mathbf{x} \rangle_h &= h_2(\mathbf{x}) \geq 0 && (\mathbf{x} \in \mathbb{R}^n). \end{aligned}$$

Finally, if $h_2(\mathbf{x}) = 0$ we have

$$h_2(\mathbf{x}) = \frac{1}{2}(x_1 + \cdots + x_n)^2 + \frac{1}{2}(x_1^2 + \cdots + x_n^2) = 0,$$

which gives $\mathbf{x} = 0_n$.

Hence, $\langle \cdot, \cdot \rangle_h$ is a scalar product and a distance (see [97]) can be given by

$$d^2(\mathbf{x}, \mathbf{y}) = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle_h} \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n). \tag{1.0.11}$$

We end the resume of this part of Chapter II by presenting an inequality of Jensen's type in the case of h_d strongly convex functions, for any even natural number $d \geq 2$.

Proposition 1.0.2. *(Jensen's type inequality for h_d strongly convexity) Let $C > 0$ and let $d \geq 2$ be an even natural number. If $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ is a given function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly convex with modulus C on I^n then, for all $x_1, \ldots, x_n \in I$, the following inequality holds*

$$\begin{aligned} f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq{}& \frac{f(x_1) + \cdots + f(x_n)}{n} \\ & - C\frac{1}{n}\binom{n+d-1}{d}\left(\left(\frac{x_1 + \cdots + x_n}{n}\right)^d - \frac{x_1^d + \cdots + x_n^d}{n}\right). \end{aligned} \tag{1.0.12}$$

In order to compare Jensen's type inequalities for h_d strongly convex functions and uniformly convex functions we present the following result (which can be seen as a consequence of the results from [170]).

Proposition 1.0.3. *(Jensen's type inequality for uniform convexity) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be such that $F : I^n \to \mathbb{R}$, defined as $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$, is uniformly convex with modulus C. Then, for all $x_1, \ldots, x_n \in I$ the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - \frac{C}{n}\sum_{1 \leq i < j \leq n}(x_i - x_j)^2. \tag{1.0.13}$$

Remark 1. *Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ such that $F : I^n \to \mathbb{R}$, defined as $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$, is h_2 strongly convex with modulus C. Then, for all $x_1, \ldots, x_n \in I$, the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - C\frac{n+1}{2n^2}\sum_{1 \leq i < j \leq n}(x_i - x_j)^2. \tag{1.0.14}$$

Note that the two constants appearing in front of right hand error term in (1.0.13) *and* (1.0.14) *are different and depend on n. Hence, we cannot move from h_2 strongly convex case to the uniformly convex case, by only changing the modulus.*

In the following we present several majorization type inequalities in the context of h_d strongly convex functions. More precisely, we are dealing with extensions of Hardy-Littlewood-Polya's and Popoviciu's inequalities in the case of our new class of convex functions.

Let us consider $\mathbf{x}^{\downarrow}$ and $\mathbf{y}^{\downarrow}$ two vectors with the same entries as $\mathbf{x}$, respectively $\mathbf{y}$, expressed in decreasing order, as

$$x_1^{\downarrow} \geq \cdots \geq x_n^{\downarrow},\ y_1^{\downarrow} \geq \cdots \geq y_n^{\downarrow}.$$

We say that, the vector $\mathbf{x}$ is *majorized* by $\mathbf{y}$ (abbreviated, $\mathbf{x} \prec \mathbf{y}$) if

$$\begin{aligned} &\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \qquad (1 \leq k \leq n-1), \\ &\sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}. \end{aligned} \tag{1.0.15}$$

More details and applications concerning the majorization theory can be found in [104]. We refer to the monotonicity with respect to the majorization order, the so called Schur-convex property, which has been introduced by I. Schur in 1923.

Definition 1.0.5. *The function $f : A \to \mathbb{R}$, where A is a symmetric subset of $\mathbb{R}^n$, is called Schur-convex if $\mathbf{x} \prec \mathbf{y}$ implies $f(\mathbf{x}) \leq f(\mathbf{y})$.*

A simple computation tool (see, for instance, [104]) which is used to study the Schur-convexity property of a function is given as follows. For any symmetric function $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$ having continuous partial derivatives on $I^n = I \times I \times ... \times I$, the Schur-convexity property is reduced to check the following inequality

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right) \geq 0 \qquad (1 \leq i,\ j \leq n,\ x_i, x_j \in I).$$

We introduce now the notions of h_d strongly Schur convexity and uniformly Schur convexity.

Definition 1.0.6. *Let $C > 0$. A function $f : I^n \to \mathbb{R}$ is said to be h_d strongly Schur-convex with modulus C if the function $f(\cdot) - C\, h_d\,(\cdot)$ is Schur-convex.*

We first remark that a similar Jensen's type inequalities is obtained by using this time majorization arguments obtaining different constants in front of the right hand error term.

Proposition 1.0.4. *(Jensen's type inequality via Hardy-Littlewood-Pólya's inequality) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. Then, for all $x_1, \ldots, x_n \in I$ the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - \frac{C}{2n^2} \sum_{1 \leq i < j \leq n} (x_i - x_j)^2. \tag{1.0.16}$$

We are able now to present Hardy-Littlewood-Pólya's majorization theorem for h_d strongly convexity case.

Theorem 1.0.5. *(Hardy-Littlewood-Pólya's inequality for strongly h_d functions) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. If $\mathbf{x} \prec \mathbf{y}$ on I^n the following inequality holds*

$$\sum_{i=1}^{n} f(y_i) \geq \sum_{i=1}^{n} f(x_i) + Ch_2(\mathbf{y} - \mathbf{x}). \tag{1.0.17}$$

Now, we can present some natural extensions of Popoviciu's inequalities for h_2 strongly convex functions.

Proposition 1.0.5. *(Popoviciu's type inequality for h_2 strongly convexity) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. Then, for all $x, y, z \in I$ the following inequality holds*

$$\begin{aligned} \frac{f(x) + f(y) + f(z)}{3} + f\left(\frac{x+y+x}{3}\right) &\geq \frac{2}{3}\left(f\left(\frac{x+y}{2}\right) + f\left(\frac{x+z}{2}\right) + f\left(\frac{y+z}{2}\right)\right) \\ &+ \frac{C}{36}\left((x-y)^2 + (y-z)^2 + (x-z)^2\right). \end{aligned} \tag{1.0.18}$$

The third part of Chapter II is based on the paper *G. M. Lăchescu, and I. Rovenţa, The Hardy-Littlewood-Pólya inequality of majorization in the context of ω-$\mathbf{m}$-star-convex functions, Aequationes Mathematicae **97** (2023), 523–535.*

In this part, we extend the Hardy-Littlewood-Pólya inequality of majorization for ω-$\mathbf{m}$-star-convex functions in the framework of ordered Banach spaces. Several open problems which seem to be of interest for further extensions of the Hardy-Littlewood-Pólya inequality are also included.

Notice that, in the early 1950s, the Hardy-Littlewood-Pólya inequality was extended by Sherman [160] to the case of continuous convex functions of a vector variable by using a much broader concept of majorization, based on matrices stochastic on lines. The full details can be found in [114], Theorem 4.7.3, p. 219. Over the years, many other generalizations in the same vein have been published. See, for example, [31, 117, 118, 124, 125, 126, 133].

As was noticed in [112] and [113], the Hardy-Littlewood-Pólya inequality of majorization can be extended to the framework of convex functions defined on ordered Banach spaces. Our aim is to prove that the same works for the larger class of ω-m-star-convex functions.

We also present different types of majorization relations in ordered Banach spaces. The corresponding extensions of the Hardy-Littlewood-Pólya inequality constitute another the objective. We end with mentioning several open problems which seem to be of interest for further extensions of the Hardy-Littlewood-Pólya inequality.

Let us consider E a Banach space and C a convex subset of it.

Definition 1.0.7. *Let m be a real parameter belonging to the interval $(0, 1]$. A function $\Phi : C \to \mathbb{R}$ is said to be a perturbed m-star-convex function with modulus $\omega : [0, \infty) \to \mathbb{R}$ (abbreviated as ω-m-star-convex function) if it fulfils an estimate of the form*

$$\Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y}) \leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}) - m\lambda(1-\lambda)\omega\left(\|\mathbf{x} - \mathbf{y}\|\right),$$

for all $\mathbf{x}, \mathbf{y} \in C$ *and* $\lambda \in (0, 1)$.

The ω*-*m*-star-convex functions associated to an identically zero modulus will be called* m*-star-convex. They satisfy the inequality*

$$\Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y}) \leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}),$$

for all $\mathbf{x}, \mathbf{y} \in C$ *and* $\lambda \in (0, 1)$.

Notice that the usual convex functions represent the particular case of m-star-convex functions where $m = 1$. On the other hand every convex function is m-star-convex (for every $m \in (0, 1]$) if $\mathbf{0} \in C$ and $\Phi(\mathbf{0}) \leq 0$. Every ω-m-star-convex function associated to a modulus $\omega \geq 0$ is necessarily m-star-convex. The ω-m-star-convex functions whose moduli ω are strictly positive except at the origin (where $\omega(0) = 0$) are usually called *uniformly* m*-star-convex.* In that case the definitory inequality is strict whenever $\mathbf{x} \neq \mathbf{y}$ and $\lambda \in (0, 1)$.

The theory of m-star-convex functions was initiated by Toader [166], who considered only the case of functions defined on real intervals. For additional results in the same setting see [108] and the references therein. A simple example of a (16/17)-star-convex function which is not convex is

$$f : [0, \infty) \to \mathbb{R}, \quad f(x) = x^4 - 5x^3 + 9x^2 - 5x. \tag{1.0.19}$$

Under the presence of Gâteaux differentiability, ω-m-star-convex functions generate specific gradient inequalities that play a prominent role in our generalization of the Hardy-Littlewood-Pólya inequality of majorization.

Lemma 1.0.1. *Suppose also that* C *is an open convex subset of the Banach space* E *and* $\Phi : C \to \mathbb{R}$ *is a function both Gâteaux differentiable and* ω*-*m*-star-convex. Then*

$$m\Phi(\mathbf{y}) \geq \Phi(\mathbf{x}) + d\Phi(\mathbf{x})(m\mathbf{y} - \mathbf{x}) + m\omega\left(\|\mathbf{x} - \mathbf{y}\|\right), \tag{1.0.20}$$

for all points $\mathbf{x}, \mathbf{y} \in C$.

Remark 2. *Lemma 1.0.1 shows that the critical points* $\mathbf{x}$ *of the differentiable* ω*-*m*-star–convex functions are those for which* $\omega \geq 0$ *fulfill the property*

$$m \inf_{\mathbf{y} \in C} \Phi(\mathbf{y}) \geq \Phi(\mathbf{x}).$$

Unlike the case of convex functions of one real variable, when the isotonicity of the differential is automatic, for several variables, this is not necessarily true in the case of a differentiable convex function of a vector variable. See [112, Remark 4].

In this part of the doctoral thesis we are dealing with functions defined on ordered Banach spaces, that is, on real Banach spaces endowed with order relations $\leq$ that make them ordered vector spaces such that positive cones are closed and

$$\mathbf{0} \leq \mathbf{x} \leq \mathbf{y} \text{ implies } \|\mathbf{x}\| \leq \|\mathbf{y}\|.$$

The Euclidean N-dimensional space $\mathbb{R}^N$ has a natural structure of an ordered Banach space associated to coordinatewise ordering. The usual sequence spaces c_0, c, ℓ^p (for $p \in [1, \infty]$) and the function spaces $C(K)$ (for K a compact Hausdorff space) and $L^p(\mu)$ (for $1 \leq p \leq \infty$ and μ a σ-additive positive measure) are also examples of ordered Banach spaces (with respect to coordinatewise/pointwise ordering and natural norms).

A map $T : E \to F$ between two ordered vector spaces is called *isotone* (or *order preserving*) if

$$\mathbf{x} \leq \mathbf{y} \text{ in } E \text{ implies } T(\mathbf{x}) \leq T(\mathbf{y}) \text{ in } F$$

and *antitone* (or *order reversing*) if $-T$ is isotone. When T is a linear operator, T is isotone if and only if T maps positive elements into positive elements (abbreviated, $T \geq 0$).

For basic informations on ordered Banach spaces see [113]. The interested reader may also consult the classical books of Aliprantis and Tourky [9] and Meyer-Nieberg [106]. As was noticed by Amann [10], Proposition 3.2, p. 184, the Gâteaux differentiability offers a convenient way to recognize the property of isotonicity of functions acting on ordered Banach spaces: the positivity of the differential. We state here his result (following the version given in [112], Lemma 4):

Lemma 1.0.2. *Suppose that E and F are two ordered Banach spaces, C is a convex subset of E with nonempty interior $\operatorname{int} C$ and $\Phi : C \to F$ is a convex function, continuous on C and Gâteaux differentiable on $\operatorname{int} C$. Then Φ is isotone on C if and only if $\Phi'(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \operatorname{int} C$.*

Remark 3. *If the ordered Banach space E has finite dimension, then the statement of Lemma 1.0.2 remains valid when the interior of C is replaced by the relative interior of C. See [114], Exercise 6, p. 81.*

We can now introduce the concept of majorization in the framework of ordered Banach spaces. Since in an ordered Banach space not every string of elements admits a decreasing rearrangement, we will concentrate on the case of pairs of discrete probability measures at least one of which is supported by a monotone string of points. The case where the support of the left measure consists of a decreasing string is defined as follows.

Definition 1.0.8. *Suppose that $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k}$ and $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$ are two discrete Borel probability measures that act on the ordered Banach space E and $m \in (0, 1]$ is a parameter. We say that $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k}$ is weakly $mL^{\downarrow}$-majorized by $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$ (denoted $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k} \prec_{wmL^{\downarrow}} \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$) if the left hand measure is supported by a decreasing string of points*

$$\mathbf{x}_1 \geq \cdots \geq \mathbf{x}_N \tag{1.0.21}$$

and

$$\sum_{k=1}^{n} \lambda_k \mathbf{x}_k \leq \sum_{k=1}^{n} \lambda_k m \mathbf{y}_k \quad \text{for all } n \in \{1, \ldots, N\}. \tag{1.0.22}$$

We say that $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k}$ is $mL^{\downarrow}$-majorized by $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$ (denoted $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k} \prec_{mL^{\downarrow}} \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$) if in addition

$$\sum_{k=1}^{N} \lambda_k \mathbf{x}_k = \sum_{k=1}^{N} \lambda_k m \mathbf{y}_k. \tag{1.0.23}$$

Notice that the context of Definition 1.0.8 makes it necessary that all the weights $\lambda_1, \ldots, \lambda_N$ belong to $(0,1]$ and $\sum_{k=1}^{N} \lambda_k = 1$. The three conditions (1.0.21), (1.0.22) and (1.0.23) imply $m\mathbf{y}_N \leq \mathbf{x}_N \leq \mathbf{x}_1 \leq m\mathbf{y}_1$ but not the ordering $\mathbf{y}_1 \geq \cdots \geq \mathbf{y}_N$. For example, when $N = 3$, one may consider the case where

$$m = 1,\ \lambda_1 = \lambda_2 = \lambda_3 = 1/3,\ \mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = \mathbf{x}$$

and

$$\mathbf{y}_1 = \mathbf{x},\ \mathbf{y}_2 = \mathbf{x} + \mathbf{z},\ \mathbf{y}_3 = \mathbf{x} - \mathbf{z},$$

$\mathbf{z}$ being any positive element.

Our objective is to consider the corresponding extensions of the Hardy-Littlewood-Pólya inequality of majorization for $\prec_{wmL^{\downarrow}}$ and$\prec_{mL^{\downarrow}}$. Moreover, we also present also a Sherman type inequality. The proof of the following theorem is inspired by the techniques successfully used in [101] and [112].

Theorem 1.0.6. *Suppose that $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k}$ and $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$ are two discrete probability measures whose supports are included in an open convex subset C of the ordered Banach space E. If $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k} \prec_{mL^{\downarrow}} \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$, then*

$$m \sum_{k=1}^{N} \lambda_k \Phi(\mathbf{y}_k) \geq \sum_{k=1}^{N} \lambda_k \Phi(\mathbf{x}_k) + \sum_{k=1}^{N} \lambda_k \omega(\|\mathbf{x}_k - \mathbf{y}_k\|), \tag{1.0.24}$$

for every Gâteaux differentiable ω-m-star-convex function $\Phi : C \to F$ whose differential is isotone and satisfies the hypotheses of Lemma 1.0.1.

The conclusion (1.0.24) *still works under the weaker hypothesis $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k} \prec_{wmL^{\downarrow}} \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$, provided that Φ is also an isotone function.*

The last part of Chapter II is based on the paper *G. M. Lăchescu, M. Malin, and I. Rovenţa, Convex type inequalities with nonpositive weights, (2024), submitted for publication.*

The weighted concept of majorization between two vectors $\mathbf{u} = (u_1, \ldots, u_l) \in I^l$, $\mathbf{v} = (v_1, \ldots, v_m) \in I^m$ with nonnegative weights $\mathbf{a} = (a_1, \ldots, a_l) \in [0, \infty)^l$ and $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$, where I is an interval in $\mathbb{R}$ and $m, l \geq 2$, has been defined in S. Sherman [160]. The concept of weighted majorization is defined by assuming the existence of a columns stochastic matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$, i.e. a matrix with nonnegative entries and columns sums equal to 1, such that

$$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \ldots, m), \tag{1.0.25}$$

$$u_i = \sum_{j=1}^{m} v_j \alpha_{ji}, \quad (i = 1, \ldots, l). \tag{1.0.26}$$

Under conditions (1.0.25) – (1.0.26) it is proved that, the following inequality

$$\sum_{i=1}^{l} a_i f(u_i) \leq \sum_{j=1}^{m} b_j f(v_j)$$

holds for every convex function $f : I \to \mathbb{R}$. See [160]. We can write conditions (1.0.25) – (1.0.26) in the matrix form

$$\mathbf{b} = \mathbf{a}\mathbf{A}^{\mathbf{T}} \quad \textit{and} \quad \mathbf{u} = \mathbf{v}\mathbf{A}.$$

We write

$$(\mathbf{u}, \mathbf{a}) \prec (\mathbf{v}, \mathbf{b})$$

and say that a pair $(\mathbf{u}, \mathbf{a})$ is weighted majorized by $(\mathbf{v}, \mathbf{b})$ if (1.0.25) – (1.0.26) are satisfied for some columns stochastic matrix $\mathbf{A}$. Note that, in the case $l = 1$ and $\mathbf{b} = [1]$ we deduce Jensen's inequality. When $m = l$ and all weights a_i and b_j are equal to $1/m$, the condition (1.0.25) assures the *stochasticity on rows,* so in that case we deal with doubly stochastic matrices.

Since all these above inequalities are dealing with positive weights the study of the case of nonpositive weights is very challenging and this is our next important objective in this thesis. In this context we recall one of the first relevant step, the so called Jensen Steffensen inequality. We refer to [115] for the following result.

Theorem 1.0.7. *Let $x_n \leq x_{n-1} \leq \cdots \leq x_1$ be points in $[a, b]$ and let $p_1, \ldots, p_n$ be real numbers such that the partial sums $S_k = \sum_{i=1}^{k} p_i$ verify the relations*

$$0 \leq S_k \leq S_n \quad \textit{and} \quad S_n > 0.$$

Then for every convex functions $f : [a, b] \to \mathbb{R}$ we have the inequality

$$f\left(\frac{1}{S_n}\sum_{k=1}^{n} p_k x_k\right) \leq \frac{1}{S_n}\sum_{k=1}^{n} p_k f(x_k).$$

Our next aim is to present new extensions of the above inequality for the case of nonpositive weights. More precisely, we try to extend Theorem 1.0.7 in the framework of $\mathbb{R}^n$ and then to derive Sherman and Jensen Steffensen's type inequalities for perturbed convex functions with complete homogeneous symmetric polynomials. Our strategy can be also adapted to more general spaces, not only in $\mathbb{R}^n$, but also in spaces with curved geometry.

Inspired from [115] we shall use the following notation related to $\mathbf{z}_1, \ldots, \mathbf{z}_m \in \mathbb{R}^n$ and $p_1, \ldots, p_m \in \mathbb{R}$:

$$\bar{\mathbf{z}} = p_1\mathbf{z}_1 + \cdots + p_m\mathbf{z}_m, \tag{1.0.27}$$

$$P_k = p_1 + \cdots + p_k \qquad (k \in \{1, 2, \ldots, m\}),$$
$$\bar{P}_k = p_k + \cdots + p_m \qquad (k \in \{1, 2, \ldots, m\}).$$

Definition 1.0.9. *We say that a sequence $\mathbf{z}_1, \ldots, \mathbf{z}_m \in \mathbb{R}^n$ is monotonic decreasing with respect to majorization relation iff the following relations hold*

$$\mathbf{z}_m \prec \mathbf{z}_{m-1} \prec \cdots \prec \mathbf{z}_2 \prec \mathbf{z}_1. \tag{1.0.28}$$

We are now in position to present the extension of Jensen-Steffensen's type inequality in $\mathbb{R}^n$.

Theorem 1.0.8. *Let I be an interval in $\mathbb{R}$ and $m, n \geq 1$. If $f : I^n \to \mathbb{R}$ is a convex function invariant under permutation of coordinates, then for every $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$, which is monotonic decreasing with respect to majorization relation, and every real m-tuple $\mathbf{p} = (p_1, \ldots, p_m)$ such that, for every $i \in \{1, 2, \ldots, m\}$ we have*

$$0 \leq P_i \leq P_m = 1,$$

then the following inequality holds

$$f\left(\sum_{i=1}^{m} p_i\mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f(\mathbf{z}_i).$$

For the convenience of the reader we also present the case of increasing sequences with respect to the majorization relation.

Theorem 1.0.9. *If $f : I^n \to \mathbb{R}$ is a convex function invariant under permutation of coordinates, then for every $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$, which is monotonic increasing with respect to majorization relation, and every real m-tuple $\mathbf{p} = (p_1, \ldots, p_m)$ such that, for every $i \in \{1, 2, \ldots, m\}$ we have*

$$0 \leq P_i \leq P_m = 1,$$

then the following inequality holds

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f(\mathbf{z}_i).$$

We can now develop the previous results for the case of nonpositive weights.

The first step is introduce the weighted concept of majorization between two n-tuples $\mathbf{x} = (x_1, \ldots, x_l)$, $\mathbf{y} = (y_1, \ldots, y_m)$, where $\mathbf{z}_1, \ldots, \mathbf{z}_l \in I^n$, $\mathbf{y}_1, \ldots, \mathbf{y}_m \in I^n$, with real weights $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{R}^l$ (which can be nonpositive) and $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$, where I is an interval in $\mathbb{R}$ and $m, l \geq 2$.

We define the concept of weighted majorization $(\mathbf{x}, \mathbf{a}) \prec (\mathbf{y}, \mathbf{b})$ by considering any matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$, verifying

$$0 \leq A_k^i \leq A_k^m = 1, \qquad (1 \leq k, i \leq m) \tag{1.0.29}$$

where

$$A_k^i = \alpha_{1i} + \cdots + \alpha_{ki} \qquad (k \in \{1, 2, \ldots, m\}) \qquad (1 \leq k \leq m), \tag{1.0.30}$$

such that

$$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \ldots, m), \tag{1.0.31}$$

$$\mathbf{x}_i = \sum_{j=1}^{m} \mathbf{y}_j \alpha_{ji}, \quad (i = 1, \ldots, l). \tag{1.0.32}$$

We can present now the extension of Sherman's inequality in $\mathbb{R}^n$, when the weights are allowed to be nonpositive.

Theorem 1.0.10. *If*

$$\mathbf{x}_m \prec \mathbf{x}_{m-1} \prec \cdots \prec \mathbf{x}_2 \prec \mathbf{x}_1. \tag{1.0.33}$$

and let us suppose that conditions (1.0.29)-(1.0.32) *are satisfied. Then, the following inequality*

$$\sum_{i=1}^{l} a_i f(\mathbf{x}_i) \leq \sum_{j=1}^{m} b_j f(\mathbf{y}_j)$$

holds for every convex function $f : I^n \to \mathbb{R}$ which is invariant under permutation of coordinates.

The next topic we address in this chapter is related to implement a similar study of a perturbed family of convex functions by complete homogeneous symmetric polynomials with even degree.

Inspired from the strategy used in [1, 2, 3, 15, 30, 83] we have the following result.

Theorem 1.0.11. *(Jensen-Steffensen's type inequality) Let $C > 0$ and let I be an interval in $\mathbb{R}$. If $f : I^n \to \mathbb{R}$ is h_2 strongly convex with modulus C and invariant under permutation of coordinates, then for every monotonic sequence $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$, as in (1.0.28), and every real n-tuple $\mathbf{p} = (p_1, \ldots, p_m)$ such that, for every $i \in \{1, 2, \ldots, m\}$, $0 \le P_i \le P_m = 1$, the following inequality holds:*

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \le \sum_{i=1}^{m} p_i f(\mathbf{z}_i) - C \sum_{i=1}^{m} p_i h_2(\mathbf{z}_i - \bar{\mathbf{z}}),$$

where $\bar{\mathbf{z}}$ is defined in (1.0.27).

Using our extension of Sherman's results (for nonpositive weights) we can deduce Sherman's inequality for h_2 strongly convex functions with modulus C.

Theorem 1.0.12. *(Sherman's type inequality) Let $C > 0$ and let I be an interval in $\mathbb{R}$. Let $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_l)$, $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$, where $\mathbf{z}_1, \ldots, \mathbf{z}_l \in I^n$, $\mathbf{y}_1, \ldots, \mathbf{y}_m \in I^n$ and let $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{R}^l$ and $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$ be such that $(\mathbf{y}, \mathbf{b}) \prec (\mathbf{z}, \mathbf{a})$. If in addition we assume that*

$$\mathbf{z}_m \prec \mathbf{z}_{m-1} \prec \cdots \prec \mathbf{z}_2 \prec \mathbf{z}_1, \tag{1.0.34}$$

then for every $f : I^n \to \mathbb{R}$ h_2 strongly convex with modulus C and invariant under permutation of coordinates we have

$$\sum_{i=1}^{l} b_i f(\mathbf{y}_i) \le \sum_{j=1}^{m} a_j f(\mathbf{z}_j) - C \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} h_2(\mathbf{z}_j - \mathbf{y}_i).$$

Chapter III of the present doctoral thesis is based on the paper *G.M. Lăchescu, M. Mălin, I. Rovenţa, On the barycenter for discrete Steffensen Popoviciu measures on global NPC spaces, submitted for publication.*

The first part of this chapter is an introductory one, we present theoretically aspects about global NPC spaces (properties and useful results). In the second part of this chapter we put in a new light the concept of barycenter for discrete Steffensen Popoviciu measures supported in some points belonging to a space with curved geometry. More precisely, we ensure the existence of the barycenter if we relax the restrictions imposed to the weights of the measure. As applications, even in the case of nonpositive weights we deduce Jensen-Steffensen's, HLP's and Sherman's type inequalities on global NPC spaces.

Several authors performed an intense research activity to extend majorization theory beyond classical case of probability measures, i.e. Steffensen Popoviciu measures. The main point of interest into this topic of research is to offer a large framework under which Jensen's type inequalities works. Jensen Steffensen's inequality (see [116, Theorem 2.4.4]) reveals an important case when Jensen's inequality works beyond the framework of positive measures. In fact, this is our aim, to relax the concept of barycenter in spaces with curved geometry, in order to provide more insight into the relation between signed measures and Jensen's type inequalities.

In fact, the above result is related to the general concept of Steffensen Popoviciu's measure, as it is presented in [114, 115, 116].

Definition 1.0.10. *Let K be a compact convex subset of a real locally convex Hausdorff space E. A Steffensen Popoviciu measure on K is any real Borel measure μ on K such that $\mu(K) > 0$ and*

$$\int_K f(x)\, d\,\mu(x) \ge 0,$$

for every positive, continuous and convex function $f : K \to \mathbb{R}$.

The characterization of discrete Steffensen Popoviciu's measures is presented in [116, Corollary 9.14].

Proposition 1.0.6. *Suppose that $x_1 \leq \cdots \leq x_n$ are real points and $p_1, \ldots, p_n$ are real weights. Then, the discrete measure $\mu = \sum_{k=1}^{n} p_k \delta_{x_k}$ is a Steffensen Popoviciu measure if*

$$\sum_{k=1}^{n} p_k > 0 \ \text{ and } \ 0 \leq \sum_{k=1}^{m} p_k \leq \sum_{k=1}^{n} p_k \quad (m \in \{1, \ldots, n\}).$$

The concept of barycenter for Steffensen Popoviciu measures was fully discussed in [116, Lemma 9.2.3 and Theorem 9.2.4]. But, our aim is to give a new perspective to the barycenter concept on more general spaces, namely global NPC spaces, via the majorization techniques.

In what follows we shall deal with the relation of weighted majorization $\prec$, for pairs of discrete probability measures. In the context of Euclidean space $\mathbb{R}^n$, the following relation

$$\sum_{i=1}^{l} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{m} \mu_j \delta_{y_j} \tag{1.0.35}$$

means the existence of a $m \times l$-dimensional matrix $A = (a_{ij})_{i,j}$ such that the next four conditions are fulfilled:

$$a_{ij} \geq 0, \ \text{for all } i, j, \tag{1.0.36}$$

$$\sum_{j=1}^{m} a_{ji} = 1, \quad i = 1, \ldots, l, \tag{1.0.37}$$

$$\mu_j = \sum_{i=1}^{l} a_{ji} \lambda_i, \quad j = 1, \ldots, m, \tag{1.0.38}$$

and

$$x_i = \sum_{j=1}^{m} a_{ji} y_j, \quad i = 1, \ldots, l. \tag{1.0.39}$$

Under the above settings, S. Sherman [160] use the concept of weighted majorization and proved that, the following inequality

$$\sum_{i=1}^{l} \lambda_i f(x_i) \leq \sum_{j=1}^{m} \mu_j f(y_j)$$

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

Our next aim is to extend Theorem 1.0.7 in the framework of global NPC spaces and then to derive HLP's, Sherman's and Jensen Steffensen's type inequalities. Hence, we consider the weighted concept of majorization within a class of spaces with curved geometry that verifies ta weaker form of Apollonius' theorem relating the length of a median of a triangle to the lengths of its sides.

Definition 1.0.11. *A global NPC space is a complete metric space $M = (M, d)$ for which the following inequality holds true: for every pair of points $x_0, x_1 \in M$ there exists a point $y \in M$ such that for all points $z \in M$,*

$$d^2(z, y) \leq \frac{1}{2} d^2(z, x_0) + \frac{1}{2} d^2(z, x_1) - \frac{1}{4} d^2(x_0, x_1). \tag{1.0.40}$$

Here "NPC" stands for "nonpositive curvature". Global NPC spaces are also known as CAT(0) spaces or Hadamard spaces. For more details, the interested reader may consult the excellent survey of Sturm [163] (and also the books of Ballman [17], Bridson and Haefliger [32], and Jost [80]).

In a global NPC space, each pair of points $x_0, x_1 \in M$ can be connected by a geodesic (that is, by a rectifiable curve $\gamma : [0,1] \to M$ such that the length of $\gamma|_{[s,t]}$ is $d(\gamma(s), \gamma(t))$ for all $0 \leq s \leq t \leq 1$). Moreover, this geodesic is unique.

The point y that appears in Definition 1.0.11 is the *midpoint* of x_0 and x_1 and has the property

$$d(x_0, y) = d(y, x_1) = \frac{1}{2} d(x_0, x_1).$$

The case of convex combinations for x_0 and x_1 can be introduced as follows:

$$(1-\lambda)x_0 \boxplus \lambda x_1 = \operatorname*{arg\,min}_{z \in M} \left[(1-\lambda)d^2(x_0, z) + \lambda d^2(x_1, z)\right]. \tag{1.0.41}$$

See Bhatia [24], Proposition 6.2.8, for the case $\lambda = 1/2$. Here, an important role is played by the inequality (1.0.40), which assures the uniform convexity of the square distance.

In a global NPC space $M = (M, d)$, the convexity notions are introduced at follows.

Definition 1.0.12. *A set $C \subset M$ is called* convex *if $\gamma([0,1]) \subset C$ for each geodesic $\gamma : [0,1] \to M$ joining the points $\gamma(0), \gamma(1) \in C$.*

A function $f : C \to \mathbb{R}$ is called convex if C is a convex set and for each geodesic $\gamma : [0,1] \to C$ the composition $f \circ \gamma$ is a convex function in the usual sense, that is,

$$f(\gamma(t)) \leq (1-t) f(\gamma(0)) + t f(\gamma(1))$$

for all $t \in [0,1]$.

The distance function d is convex on $M \times M$, while the functions $d^\alpha(\cdot, z)$, with $\alpha \geq 1$, are convex on M. See Sturm [163, Corollary 2.5], for details. Despite the fact that the property of associativity of convex combinations fails it is worth mentioning that Jensen's inequality works in the context of global NPC spaces. Note that, the basic ingredient, the *barycenter* of a discrete probability measures $\lambda = \sum_{i=1}^n \lambda_i \delta_{x_i}$ is defined by the formula

$$\operatorname{bar}(\lambda) = \operatorname*{arg\,min}_{z \in M} \frac{1}{2} \sum_{i=1}^n \lambda_i d^2(z, x_i).$$

In the case of Hilbert spaces, this coincides with the usual definition of barycenter in flat spaces, which is given by $\sum_{i=1}^n \lambda_i x_i$.

The next result is a particular case of the integral form of Jensen's Inequality, which was first noticed by Jost [79] (and later extended by Eells and Fuglede [55]). A probabilistic version can be found in [163].

Theorem 1.0.13. (The discrete form of Jensen's Inequality). *For every continuous convex function $f : M \to \mathbb{R}$ and every discrete probability measure $\lambda = \sum_{i=1}^n \lambda_i \delta_{x_i}$ on M, we have the inequality*

$$f(\operatorname{bar}(\lambda)) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

When $x_1, \ldots, x_m, y_1, \ldots, y_n$ are points in a global NPC space (M, d) and $\lambda_1, \ldots, \lambda_m$ in $[0, 1]$ are weights that sum to 1, we define the relation of majorization

$$\sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{n} \mu_j \delta_{y_j} \tag{1.0.42}$$

by asking the existence of an $m \times n$-dimensional matrix $A = (a_{ij})_{i,j}$ that is stochastic on rows and verifies in addition the following two conditions:

$$\mu_j = \sum_{i=1}^{m} a_{ij} \lambda_i, \quad j = 1, \ldots, n \tag{1.0.43}$$

and

$$x_i = \arg\min_{z \in M} \frac{1}{2} \sum_{j=1}^{n} a_{ij} d^2(z, y_j), \quad i = 1, \ldots, m. \tag{1.0.44}$$

The existence and uniqueness of the optimization problems (1.0.44) is assured by the fact that the objective functions are uniformly convex and positive. See Jost [80], Section 3.1, or Sturm [163, Proposition 1.7, p. 3]. According to our definition, we have

$$\delta_{\mathrm{bar}(\lambda)} \prec \lambda,$$

for every discrete Borel probability measure λ. The following theorem in [126] offers an extension of the Hardy-Littlewood-Pólya Theorem (HLP) in the context of global NPC spaces.

Theorem 1.0.14. *If the relation*

$$\sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{n} \mu_j \delta_{y_j}$$

hold in the global NPC space M, then for every real-valued continuous convex function f defined on a convex subset $U \subset M$ that contains all points x_i and y_j, we have the following inequality

$$\sum_{i=1}^{m} \lambda_i f(x_i) \leq \sum_{j=1}^{n} \mu_j f(y_j).$$

Moreover, using Theorem 1.0.14 we have the following result, in which the properties of convexity and Schur convexity are connected.

Proposition 1.0.7. (*Lim [98], Niculescu and Rovenţa [126]*) *If we have that*

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \prec \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i},$$

in the global NPC space M and $f : M^n \to \mathbb{R}$ is a continuous convex function invariant under the permutation of coordinates, then

$$f(x_1, \ldots, x_n) \leq f(y_1, \ldots, y_n).$$

Inspired from [94], in this thesis we present an extension of barycenter for Steffensen Popoviciu discrete measures, where the most important ingredient in NPC spaces is the *barycenter* of a discrete probability measures $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$. Thus, in what follows we relax the concept of barycenter by considering nonpositive weights for the discrete measures.

Definition 1.0.13. *Let $X := \{x_1, \dots, x_n\}$ be a family of points in a global NPC space M, all these points belonging to the same geodesic $[x_1, x_n]$ and, in addition the following assumptions are verified*

$$x_i \in [x_{i-1}, x_{i+1}] \qquad (i \in \{2, 3, \dots, n-1\}).$$

For any family of real weights $\Lambda := \{\lambda_1, \dots, \lambda_n\}$ which verify

$$0 \leq S_i \leq S_n = 1 \qquad (i \in \{1, 2, \dots, n\}),$$

where

$$S_k = \lambda_1 + \dots + \lambda_k \qquad (k \in \{1, 2, \dots, n\}),$$

we define the notion of weak barycenter of the family of points X with respect to the family of real weights Λ as the unique point $\bar{\mathbf{x}}$ on the geodesic $[x_1, x_n]$ satisfying

$$d(\bar{\mathbf{x}}, x_1) = \bar{S}_2 d(x_2, x_1) + \bar{S}_3 d(x_3, x_2) + \dots + \bar{S}_n d(x_n, x_{n-1}), \tag{1.0.45}$$

or, equivalently,

$$d(x_n, \bar{\mathbf{x}}) = S_1 d(x_2, x_1) + S_2 d(x_3, x_2) + \dots + S_{n-1} d(x_n, x_{n-1}), \tag{1.0.46}$$

where

$$\bar{S}_k = \lambda_k + \dots + \lambda_n \qquad (k \in \{1, 2, \dots, n\}).$$

Remark 4. *Note that, the weak barycenter $\bar{\mathbf{x}}$ from* (1.0.45) *and* (1.0.46) *is well defined and we have that*

$$d(\bar{\mathbf{x}}, x_1 + d(x_n, \bar{\mathbf{x}}) = d(x_2, x_1) + d(x_3, x_2) + \dots + d(x_n, x_{n-1}) = d(x_n, x_1),$$

which confirm the fact that $\bar{\mathbf{x}}$ lies on the geodesic $[x_1, x_n]$. Moreover, using (1.0.45) *or* (1.0.46)*, in flat spaces, so we recover the following classical formula*

$$\bar{\mathbf{x}} = \lambda_1 x_1 + \dots + \lambda_n x_n.$$

We are now in position to present Jensen-Steffensen's inequality in the most relevant case, where we have considered the maximum possible number of nonpositive weights. In fact, in Chapter III a completely new strategy is used to prove the following result.

Theorem 1.0.15. (The discrete form of Jensen-Steffensen's Inequality) *Let X and Λ be given as in Definition 1.0.13, but with nonpositive weights $\lambda_2, \lambda_3, \dots, \lambda_{n-1} \leq 0$.*

Then, for every continuous convex function $f : M \to \mathbb{R}$ we have the inequality

$$f(\bar{\mathbf{x}}) \leq \sum_{i=1}^{n} \lambda_i f(x_i).$$

In order to obtain in thesis, Sherman's type inequalities with nonpositive weights we firstly introduce the relaxed concept of majorization between two n-tuples of points in a global NPC space (M, d).

Definition 1.0.14. *Let $\mathbf{x} = (x_1, \dots, x_n) \in M^n$, $\mathbf{y} = (y_1, \dots, y_n) \in M^n$, $n \geq 2$.*

We define the concept of majorization $\mathbf{x} \prec \mathbf{y}$ by asking the existence of a matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$ such that

- $\alpha_{ji} \leq 0 \qquad (i \neq 1 \text{ or } i \neq l)$,
- $y_i \in [y_{i-1}, y_{i+1}] \qquad (i \in \{2, 3, \ldots, l-1\})$ *verify that all these points belong to the same geodesic* $[y_1, y_m]$,
- x_i *is the weak barycenter of the family of points* $X := \{y_1, \ldots, y_m\}$ *with respect to the family of real weights* Λ^j, *i.e. the unique point* x_i *on the geodesic* $[y_1, y_m]$ *satisfying*

$$d(x_i, y_1) = \bar{S}_2^j d(y_2, y_1) + \bar{S}_3^j d(y_3, y_2) + \cdots + \bar{S}_n^j d(y_n, y_{n-1}), \tag{1.0.47}$$

or, equivalently,

$$d(y_n, x_i) = S_1^j d(y_2, y_1) + S_2^j d(y_3, y_2) + \cdots + S_{n-1}^j d(y_n, y_{n-1}), \tag{1.0.48}$$

where

$$\begin{aligned} \Lambda^j &:= \{\alpha_{1j}, \ldots, \alpha_{nj}\} \qquad (j \in \{1, \ldots, m\}), \\ \bar{S}_k^j &= \alpha_{kj} + \cdots + \alpha_{nj} \qquad (k \in \{1, 2, \ldots, l\}), \\ S_k^j &= \alpha_{1j} + \cdots + \alpha_{kj} \qquad (k \in \{1, 2, \ldots, l\}), \\ 0 \leq S_k^j &\leq S_n^j = 1 \qquad (k \in \{1, 2, \ldots, l\}). \end{aligned}$$

We can present now the extension of HLP's inequality in a global NPC space (M, d), when the weights are allowed to be nonpositive.

Theorem 1.0.16. *In the hypotheses from Definition 1.0.14 let us suppose that conditions* (1.0.47) *are satisfied. Then, the following inequality*

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

holds for every convex function $f : M \to \mathbb{R}$.

We are in position to introduce another result of this thesis, the relaxed weighted concept of majorization between two n-tuples of points in a global NPC space M.

Definition 1.0.15. *Let* $\mathbf{x} = (x_1, \ldots, x_l) \in M^l$, $\mathbf{y} = (y_1, \ldots, y_m) \in M^m$, $m, l \geq 2$. *We consider some real weights* $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{R}^l$ *(which can be nonpositive) and* $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$.

We define the concept of weighted majorization $(\mathbf{x}, \mathbf{a}) \prec (\mathbf{y}, \mathbf{b})$ *by asking the existence of a matrix* $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$ *such that*

- $\alpha_{ji} \leq 0 \qquad (i \neq 1 \text{ or } i \neq l)$,
- $y_i \in [y_{i-1}, y_{i+1}] \qquad (i \in \{2, 3, \ldots, l-1\})$ *verify that all these points belong to the same geodesic* $[y_1, y_m]$,
- x_i *is the weak barycenter of the family of points* $X := \{y_1, \ldots, y_m\}$ *with respect to the family of real weights* Λ^j, *i.e. the unique point* x_i *on the geodesic* $[y_1, y_m]$ *satisfying*

$$d(x_i, y_1) = \bar{S}_2^j d(y_2, y_1) + \bar{S}_3^j d(y_3, y_2) + \cdots + \bar{S}_n^j d(y_n, y_{n-1}), \tag{1.0.49}$$

or, equivalently,

$$d(y_n, x_i) = S_1^j d(y_2, y_1) + S_2^j d(y_3, y_2) + \cdots + S_{n-1}^j d(y_n, y_{n-1}), \tag{1.0.50}$$

where

$$\begin{aligned} \Lambda^j &:= \{\alpha_{1j}, \dots, \alpha_{nj}\} \qquad (j \in \{1, \dots, m\}), \\ \bar{S}_k^j &= \alpha_{kj} + \cdots + \alpha_{nj} \qquad (k \in \{1, 2, \dots, l\}), \\ S_k^j &= \alpha_{1j} + \cdots + \alpha_{kj} \qquad (k \in \{1, 2, \dots, l\}), \\ 0 \le S_k^j &\le S_n^j = 1 \qquad (k \in \{1, 2, \dots, l\}), \end{aligned}$$

- *the following identities hold*

$$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \dots, m). \tag{1.0.51}$$

We can now present the extension of Sherman's inequality in a global NPC space (M, d), when the weights are allowed to be nonpositive.

Theorem 1.0.17. *In the hypotheses from Definition 1.0.15 let us suppose that conditions* (1.0.45) *are satisfied. Then, the following inequality*

$$\sum_{i=1}^{m} a_i f(x_i) \le \sum_{j=1}^{l} b_j f(y_j)$$

holds for every convex function $f : M \to \mathbb{R}$.

The last chapter of this doctoral thesis is devoted to some conclusions, final remarks and further research objectives. We recall some recent results in literature which could be useful to develop the results obtained in this doctoral thesis. We also emphasize several different directions of research as follows: polynomial norms defined in terms of symmetric homogeneous polynomials of even degree and some error estimates in this area; convexity properties of other symmetric polynomials; discrete Korn inequalities, etc.

Chapter 2

New majorization results on h_d strongly convex functions

2.1 Convexity and majorization in $\mathbb{R}^n$

In this section, we present some notions (inspired from [68, 69, 104, 116]) related to convexity, majorization theory and inequalities associated to it. The relation of majorization was introduced by G. H . Hardy, J. E. Littlewood and G. Pólya [69] in 1929, and was popularized by their celebrated book on Inequalities [68]. Part of this research activity is summarized in the 900 pages of the recent book by A. W. Marshall, I. Olkin and B. Arnold [104].

2.1.1 The Hardy-Littlewood-Pólya theory of majorization

The main problem of this subsection is to find necessary and sufficient conditions under which two families of real numbers $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$ accompanied by a family $(p_k)_{k=1}^n$ of positive weights, which verify the inequality

$$\sum_{k=1}^{n} p_k f(x_k) \leq \sum_{k=1}^{n} p_k f(y_k), \tag{2.1.1}$$

for every real-valued continuous convex function f defined on an interval that contains all real numbers of x_k and y_k.

This problem is related to Jensen's inequality, which occurs in particular case where

$$\sum_{k=1}^{n} p_k = 1 \quad \text{and} \quad x_1 = \cdots = x_n = \sum_{k=1}^{n} p_k y_k.$$

Because the identity and its opposite are continuous convex functions, we deduce that the inequality (2.1.1) imposes the equality

$$\sum_{k=1}^{n} p_k x_k = \sum_{k=1}^{n} p_k y_k. \tag{2.1.2}$$

Furthermore, using the convexity of the functions $f = (x - y_k)^+$, we obtain

$$p_1x_1 + \cdots + p_kx_k - (p_1 + \cdots + p_k)y_k \leq \sum_{j=1}^{n} f(x_j)$$

$$\leq \sum_{j=1}^{n} f(y_j) \leq p_1y_1 + \cdots + p_ky_k - (p_1 + \cdots + p_k)y_k,$$

which gives a new set of necessary conditions:

$$\sum_{k=1}^{m} p_kx_k \leq \sum_{k=1}^{m} p_ky_k \quad \text{for all} \quad m \in \{1, \ldots, n-1\}. \tag{2.1.3}$$

Unexpected, the conditions (2.1.2) and (2.1.3) are also sufficient for solving the problem mentioned above, even in the case of real weights. This fact is known as Fuchs' generalization of the Hardy-Littlewood-Pólya inequality of majorization and can be stated as follows:

Theorem 2.1.1. *If $(x_k)_{k=1}^n$ and $(y_k)_{k=1}^n$ are two families of real numbers directed downwards,*

$$x_1 \geq \cdots \geq x_n \text{ and } y_1 \geq \cdots \geq y_n,$$

and $(p_k)_k$ is a family of real weights which satisfy the conditions (2.1.2) *and* (2.1.3)*, then*

$$\sum_{k=1}^{n} p_kf(x_k) \leq \sum_{k=1}^{n} p_kf(y_k), \tag{2.1.4}$$

for every function f whose domain of definition is an interval that contains all numbers x_k and y_k.

Furthermore, in the case when the two families of real numbers $(x_k)_k$ and $(y_k)_k$ are directed upwards, the inequality (2.1.4) *works in the reverse direction.*

Proof. Without loss of generality we may assume that $x_k \neq y_k$ for all indices k. Then, according to Abel's partial summation formula [116, Theorem 2.4.5] we have

$$\sum_{k=1}^{n} p_kf(y_k) - \sum_{k=1}^{n} p_kf(x_k) = \sum_{k=1}^{n} \left[p_k(y_k - x_k)\frac{f(y_k) - f(x_k)}{y_k - x_k} \right]$$

$$= \sum_{k=1}^{n-1} \left(\frac{f(y_k) - f(x_k)}{y_k - x_k} - \frac{f(y_{k+1}) - f(x_{k+1})}{y_{k+1} - x_{k+1}} \right) \left(\sum_{i=1}^{k} p_iy_i - \sum_{i=1}^{k} p_ix_i \right)$$

$$+ \frac{f(y_n) - f(x_n)}{y_n - x_n} \left(\sum_{i=1}^{n} p_iy_i - \sum_{i=1}^{n} p_ix_i \right)$$

$$= \sum_{k=1}^{n-1} \left(\frac{f(y_k) - f(x_k)}{y_k - x_k} - \frac{f(y_{k+1}) - f(x_{k+1})}{y_{k+1} - x_{k+1}} \right) \left(\sum_{i=1}^{k} p_iy_i - \sum_{i=1}^{k} p_ix_i \right) \geq 0$$

due to our hypotheses (2.1.2) and (2.1.3) and the three chords inequality [116, Remark 1.4.1].

Remark that in the case where $x_1 \leq \cdots \leq x_n$ and $y_1 \leq \cdots \leq y_n$, the three chords inequality implies

$$\frac{f(y_k) - f(x_k)}{y_k - x_k} - \frac{f(y_{k+1}) - f(x_{k+1})}{y_{k+1} - x_{k+1}} \leq 0,$$

and the proof is complete. $\square$

An analysis of the argument of Theorem 2.1.1, leads to the next result namely Fuchs' generalization of the Tomić-Weyl inequality of majorization:

Theorem 2.1.2. *If $(x_k)_{k=1}^n$ and $(y_k)_{k=1}^n$ are two families of real numbers directed downwards,*

$$x_1 \geq \cdots \geq x_n \text{ and } y_1 \geq \cdots \geq y_n,$$

and $(p_k)_k$ is a family of real weights which satisfy the inequalities

$$\sum_{k=1}^{m} p_k x_k \leq \sum_{k=1}^{m} p_k y_k \text{ for all } m \in \{1, \ldots, n\},$$

then

$$\sum_{k=1}^{n} p_k f(x_k) \leq \sum_{k=1}^{n} p_k f(y_k), \tag{2.1.5}$$

for every convex and increasing function f whose domain of definition is an interval that contains all numbers x_k and y_k.

Furthermore, in the case when the two families of real numbers $(x_k)_k$ and $(y_k)_k$ are directed upwards, the last inequality works in the reverse direction.

Note that Hardy, Littlewood and Pólya [68], [69] have considered only the unweighted case of their inequalities of majorization, that is, the case where all weights equal to unity. We recall them here by removing the unnecessary assumption on the continuity of the functions involved and noting that the monotonicity of only one of the two families of numbers is sufficient.

Theorem 2.1.3. *(The Hardy-Littlewood-Pólya inequality of majorization) Let f be a convex function defined on an interval I and let $\mathbf{x} = (x_k)_{k=1}^n$ and $\mathbf{y} = (y_k)_{k=1}^n$ be two families of numbers in I such that*

$$\sum_{k=1}^{m} x_k \leq \sum_{k=1}^{m} y_k \text{ for } m = 1, \ldots, n-1 \tag{2.1.6}$$

and

$$\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k. \tag{2.1.7}$$

If $x_1 \geq \cdots \geq x_n$, then

$$\sum_{k=1}^{n} f(x_k) \leq \sum_{k=1}^{n} f(y_k), \tag{2.1.8}$$

while if $y_1 \leq \cdots \leq y_n$, then the last inequality works in the reverse direction.

Proof. When $x_1 \geq \cdots \geq x_n$ and the hypotheses (2.1.6) and (2.1.7) hold, they will continue to work when $\mathbf{y}$ is replaced by the vector obtained from $\mathbf{y}$ by rearranging its components in decreasing order and the conclusion is a direct consequence of Theorem 2.1.1.

Taking into account the property of subdifferentiability of convex functions we can avoid the use of Theorem 2.1.1 and for this, rearrange $\mathbf{y}$ as above and observe that we may assume that $x_k \neq y_k$ for all

k. According to the hypotheses (2.1.6) and (2.1.7), it follows that $x_1 < y_1$ and $x_n > y_n$, so all points x_k are interior to I. Then by using [116, Theorem 2.1.2] we have that

$$\begin{aligned}\sum_{k=1}^{n} f(y_k) - \sum_{k=1}^{n} f(x_k) &= \sum_{k=1}^{n} (f(y_k) - f(x_k)) \geq \sum_{k=1}^{n} f'_+(x_k)(y_k - x_k)\\ &= \sum_{m=1}^{n-1} (f'_+(x_m) - f'_+(x_{m+1}))\left(\sum_{k=1}^{m} (y_k - x_k)\right) + f'_+(x_n)\sum_{k=1}^{n}(x_k)(y_k - x_k)\\ &= \sum_{m=1}^{n-1} (f'_+(x_m) - f'_+(x_{m+1}))\left(\sum_{k=1}^{m} (y_k - x_k)\right),\end{aligned}$$

and the proof ends by noticing that the right derivative f'_+ of the convex function f is increasing on the interior of I.

Remark that the case when $y_1 \leq \cdots \leq y_n$ can be treated in a similar way, by replacing $\mathbf{x}$ by the vector obtained from it by rearranging its components in increasing order. $\square$

An immediate consequence of Theorem 2.1.3 is as follows:

Corollary 1. *(Truncated majorization) Let $f : [0, \infty) \to \mathbb{R}$ be an increasing concave function and let $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m$ $(2 \leq m \leq n)$ be nonnegative numbers such that*

$$\max\{x_1, \ldots, x_n\} \leq \max\{y_1, \ldots, y_m\},$$

$$\max\{x_{i_1} + x_{i_2} : i_1 \neq i_2\} \leq \max\{y_{j_1} + y_{j_2} : j_1 \neq j_2\},$$

$$\ldots$$

$$\max\left\{\sum_{p=1}^{k} x_{i_p} : i_r \neq i_s\right\} \leq \max\left\{\sum_{p=1}^{k} y_{j_p} : j_r \neq j_s\right\},$$

for $k \leq m$ and

$$\sum_{i=1}^{n} x_i \geq \sum_{j=1}^{m} y_j.$$

Then

$$\sum_{i=1}^{n} f(x_i) \geq \sum_{j=1}^{m} y_j + (n - m)f(0).$$

The argument of Theorem 2.1.3 yields the Tomić-Weyl inequality of weak majorization [167, 168]:

Theorem 2.1.4. *Let f be a convex and increasing function defined on an interval I and let $\mathbf{x} = (x_k)_{k=1}^{n}$ and $\mathbf{y} = (y_k)_{k=1}^{n}$ be two families of numbers in I such that*

$$\sum_{k=1}^{m} x_k \leq \sum_{k=1}^{m} y_k \text{ for } m = 1, \ldots, n.$$

If $x_1 \geq \cdots \geq x_n$, then

$$\sum_{k=1}^{n} f(x_k) \leq \sum_{k=1}^{n} f(y_k),$$

while if $y_1 \leq \cdots \leq y_n$, then the last inequality works in the reverse direction.

Now we are in a position to indicate the precise definition of the notion of majorization.

Given a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we denote by $\mathbf{x}^{\downarrow} = (x_1^{\downarrow}, \ldots, x_n^{\downarrow})$ the vector obtained from $\mathbf{x}$ by rearranging its components in decreasing order,

$$x_1^{\downarrow} \geq \cdots \geq x_n^{\downarrow}.$$

In a similar way we may introduce the vector $\mathbf{x}^{\uparrow} = (x_1^{\uparrow}, \ldots, x_n^{\uparrow})$, obtained from $\mathbf{x}$ by rearranging its components in increasing order.

Definition 2.1.1. *Given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, we say that $\mathbf{x}$ is weakly majorized by $\mathbf{y}$ (denoted $\mathbf{x} \prec_{HLPw} \mathbf{y}$) if*

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \text{ for } k = 1, \ldots, n,$$

and that $\mathbf{x}$ is majorized by $\mathbf{y}$ (denoted $\mathbf{x} \prec_{HLP} \mathbf{y}$) if in addition

$$\sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}.$$

When considering the similar relations for vectors rearranged in increasing order, one obtains respectively the corresponding relations denoted $\mathbf{x} \prec_{HLPw}^{*} \mathbf{y}$ and $\mathbf{x} \prec_{HLP}^{*} \mathbf{y}$.

Notice that $\prec_{HLP}$ provides a partial order on $\mathbb{R}^n$ and the initial order of the components in the vectors does not play any role in terms of majorization. As it will be shown in Theorem 2.1.7, the fact that $\mathbf{x} \prec_{HLP} \mathbf{y}$ means geometrically that the components of $\mathbf{x}$ spread out less than those of $\mathbf{y}$.

In information theory, the relation $\mathbf{p} \prec_{HLP} \mathbf{q}$ (for $\mathbf{p}$ and $\mathbf{q}$ probability distributions on N outcomes) implies that $\mathbf{p}$ is more *disordered* than $\mathbf{q}$. Indeed, according to Theorem 2.1.3, we have

$$H(\mathbf{p}) = -\sum_{k=1}^{n} p_k \log_2 p_k \geq H(\mathbf{q}) = -\sum_{k=1}^{n} q_k \log_2 q_k,$$

that is, the Shannon entropy of $\mathbf{q}$ does not exceed the Shannon entropy of $\mathbf{p}$. Since the converse does not work (that is, in general the inequality $H(\mathbf{p}) \geq H(\mathbf{q})$ does not imply $\mathbf{p} \prec_{HLP} \mathbf{q}$), this suggests that the majorization theory could offer stronger criteria for measurement of disorder in a system than the entropic inequalities. Indeed, this is the case, and the details can be found in the survey of M. A. Nielsen [129] and the monograph of M. A. Nielsen and I. L. Chuang [130].

Remark 5. *The fact that the Tomić-Weyl inequality of weak majorization was obtained as a consequence of the Hardy-Littlewood-Pólya inequality of majorization is not an accident. Indeed, as was noticed by G. Pólya [141], if*

$$(x_1, \ldots, x_n) \prec_{HLPw} (y_1, \ldots, y_n),$$

then there exist real numbers x_{n+1} and y_{n+1} such that

$$(x_1, \ldots, x_n, x_{n+1}) \prec_{HLP} (y_1, \ldots, y_n, y_{n+1}).$$

To check this, choose

$$x_{n+1} = \min\{x_1, \ldots, x_n, y_1, \ldots, y_n\} \text{ and } y_{n+1} = \sum_{k=1}^{n+1} x_k - \sum_{k=1}^{n} y_k.$$

W. Arveson and R. Kadison have found another reduction of weak majorization to majorization. Let $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$ be two decreasing sequences of positive real numbers such that

$$x_1 + \cdots + x_k \leq y_1 + \cdots + y_k \text{ for } k = 1, \ldots, n.$$

Then there is a decreasing sequence $\overline{y}_1 \geq \cdots \geq \overline{y}_n$ such that $0 \leq \overline{y}_k \leq y_k$ for all k and

$$(x_1, \ldots, x_n) \prec_{HLP} (\overline{y}_1, \ldots, \overline{y}_n).$$

The aforementioned book [68] of Hardy, Littlewood and Pólya also includes a description of the relation of majorization by averaging means, based on the doubly stochastic matrices. Recall that a matrix $A \in M_n(\mathbb{R})$ is doubly stochastic if A has positive entries and each row and each column sums to unity. A special class of doubly stochastic matrices is that of T-transformations which have the form

$$T = \lambda I + (1 - \lambda) Q,$$

where $0 \leq \lambda \leq 1$ and Q is a permutation mapping which interchanges two coordinates, that is,

$$Tx = (x_1, \ldots, x_{j-1}, \lambda x_j + (1-\lambda) x_k, x_{j+1}, \ldots, x_{k-1}, \lambda x_k + (1-\lambda) x_j, x_{k+1}, \ldots, x_n).$$

Theorem 2.1.5. *If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the following statements are equivalent:*

- (i) $\mathbf{x} \prec_{HLP} \mathbf{y}$*;*
- (ii) $\mathbf{x} = A\mathbf{y}$ *for a suitable doubly stochastic matrix* $A \in M_n(\mathbb{R})$*;*
- (iii) $\mathbf{x}$ *can be obtained from* $\mathbf{y}$ *by successive applications of finitely many* T*-transformations.*

Remark that the implication $(ii) \Rightarrow (i)$ is due to I. Schur [157] and constituted the starting point for this theorem.

Proof. $(iii) \Rightarrow (ii)$ Since T-transformations are doubly stochastic, the product of T-transformations is a doubly stochastic transformation.

$(ii) \Rightarrow (i)$ This implication is a consequence of Theorem 2.1.3. Assume that $A = (a_{jk})_{j,k=1}^n$ and consider an arbitrary continuous convex function f defined on an interval including the components of $\mathbf{x}$ and $\mathbf{y}$. Since $x_k = \sum_j y_i a_{jk}$, and $\sum_j a_{jk} = 1$ for all indices k, it follows from Jensen's inequality that

$$f(x_k) \leq \sum_{j=1}^{n} a_{jk} f(y_j).$$

Then

$$\sum_{k=1}^{n} f(x_k) \leq \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{jk} f(y_i) \right) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{jk} f(y_i) \right) = \sum_{j=1}^{n} f(y_i),$$

and Theorem 2.1.3 applies.

$(i) \Rightarrow (iii)$ Let $\mathbf{x}$ and $\mathbf{y}$ be two distinct vectors in $\mathbb{R}^n$ such that $\mathbf{x} \prec_{HLP} \mathbf{y}$. Since permutations are T-transformations, we may assume that their components verify the conditions

$$x_1 \geq \cdots \geq x_n \text{ and } y_1 \geq \cdots \geq y_n.$$

Let j be the largest index such that $x_j < y_j$ and let k be the smallest index such that $k > j$ and $x_k > y_k$. The existence of such a pair of indices is motivated by the fact that the largest index i with $x_i \neq y_i$ verifies $x_i > y_i$. Then

$$y_j > x_j \geq x_k > y_k.$$

Put

$$\epsilon = \min\{y_j - x_j, x_k - y_k\},\ \lambda = 1 - \frac{\epsilon}{y_j - y_k}$$

and

$$\mathbf{y}^* = (y_1, \ldots, y_{j-1}, y_j - \epsilon, y_{j+1}, \ldots, y_{k-1}, y_k + \epsilon, y_{k+1}, \ldots, y_n).$$

Clearly, $\lambda \in (0,1)$. Denoting by Q the permutation matrix which interchanges the components of order j and k, we see that $\mathbf{y}^* = T\mathbf{y}$ for the representation

$$T = \lambda I + (1-\lambda)Q.$$

According to the implication $(ii) \Rightarrow (i)$, it follows that $\mathbf{y}^* \prec_{HLP} \mathbf{y}$. On the other hand, $\mathbf{x} \prec_{HLP} \mathbf{y}^*$. In fact,

$$\sum_{r=1}^{s} y_r^* = \sum_{r=1}^{s} y_r \geq \sum_{r=1}^{s} x_r \text{ for } s = 1, \ldots, j-1,$$

$$y_j^* \geq x_j \text{ and } y_r^* = y_r \text{ for } r = j+1, \ldots, k-1,$$

$$\sum_{r=1}^{s} y_r^* = \sum_{r=1}^{s} y_r \geq \sum_{r=1}^{s} x_r \text{ for } s = k+1, \ldots, n$$

and

$$\sum_{r=1}^{n} y_r^* = \sum_{r=1}^{n} y_r = \sum_{r=1}^{n} x_r.$$

Letting $d(\mathbf{u}, \mathbf{v})$ be the number of indices r such that $u_r \neq v_r$, it is clear that $d(\mathbf{x}, \mathbf{y}^*) \leq d(\mathbf{x}, \mathbf{y}) - 1$, so by repeating the above algorithm (at most) $n-1$ times, we arrive at $\mathbf{x}$. □

The Hardy-Lttlewood-Pólya inequality of majorization admits a generalization due to S. Sherman [160] to the case where the vectors $\mathbf{x}$ and $\mathbf{y}$ are not necessarily in the same vector space. For functions defined on intervals it reduces to the following result.

Theorem 2.1.6. *Suppose that I is an interval of $\mathbb{R}$ and consider the vectors $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{R}_+^m$, $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}_+^n$, $\mathbf{x} = (x_1, \ldots, x_m) \in I^m$ and $\mathbf{y} = (y_1, \ldots, y_n) \in I^n$. Then*

$$\sum_{i=1}^{n} b_i f(y_i) \leq \sum_{j=1}^{m} a_j f(x_j),$$

for every convex function $f : I \to \mathbb{R}$ if and only if there exists an $n \times m$ dimensional matrix $S = (s_{ij})_{i,j}$ with positive entries such that

$$\mathbf{y} = S\mathbf{x}, \quad \mathbf{a} = S^T\mathbf{b}$$

and

$$\sum_{j=1}^{m} s_{ij} = 1 \text{ for every } i.$$

An elegant proof of this result has been presented by G. Bennett [23], Lemma 6.1, p. 894. The next two results need some preparation concerning the action of the permutation group on the Euclidian space.

The permutation group of order n is the group $\Pi(n)$ of all bijective functions from $\{1, \ldots, n\}$ onto itself. This group acts on $\mathbb{R}^n$ via the map $\Psi : \Pi(n) \times \mathbb{R}^n \to \mathbb{R}^n$, defined by the formula

$$\Psi(\pi, \mathbf{x}) = \pi\mathbf{x} = (x_{\pi(1)}, \ldots, x_{\pi(n)}).$$

The orbits of this action, that is, the sets of the form $\mathcal{O}(x) = \{\pi x : \pi \in \Pi(n)\}$, play an important role in majorization theory.

Definition 2.1.2. *A subset C of $\mathbb{R}^n$ is called invariant under permutations (or $\Pi(n)$-invariant) if $\pi\mathbf{x} \in C$ whenever $\pi \in \Pi(n)$ and $\mathbf{x} \in C$. Therefore, a function F defined on a $\Pi(n)$-invariant subset C is called $\Pi(n)$-invariant (or invariant under permutations) if $F(\pi\mathbf{x}) = F(\mathbf{x})$ whenever $\pi \in \Pi(n)$ and $\mathbf{x} \in C$.*

Note that all elementary symmetric functions (as well as all norms of index $p \in [1, \infty]$) are invariant under permutations. Furthermore, to every convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ one can attach a convex function $\varphi_\Pi : \mathbb{R}^n \to \mathbb{R}$ invariant under permutations via the formula

$$\varphi_\Pi(\mathbf{x}) = \sum_{\pi \in \Pi(n)} \varphi(\pi\mathbf{x}).$$

A geometric insight into majorization was revealed by R. Rado, who noticed that $\mathbf{x} \prec_{HLP} \mathbf{y}$ means that the components of $\mathbf{x}$ spread out less than those of $\mathbf{y}$ in the sense that $\mathbf{x}$ lies in the convex hull of the $n!$ permutations of $\mathbf{y}$.

Theorem 2.1.7. *(R. Rado [146]) $\mathbf{x} \prec_{HLP} \mathbf{y} \in \mathbb{R}^n$ if and olny if $\mathbf{x}$ belongs to the convex hull of the $n!$ permutations of $\mathbf{y}$. Therefore, we have*

$$\{\mathbf{x} : \mathbf{x} \prec_{HLP} \mathbf{y}\} = conv\{\pi\mathbf{y} : \pi \in \Pi(n)\}.$$

Proof. According to Theorem 2.1.5, if $\mathbf{x} \prec_{HLP} \mathbf{y}$, then $\mathbf{x} = A\mathbf{y}$ for some doubly stochastic matrix. Taking into account Birkhoff's theorem, A can be represented as a convex combination $A = \sum_\pi \lambda_\pi A_\pi$ of the $n!$ permutation matrices A_π. Then we have

$$\mathbf{x} = \sum_\pi \lambda_\pi A_\pi(\mathbf{y}) \in \mathrm{conv}\{A_\pi(\mathbf{y}) : \pi \in \Pi(n)\}.$$

Conversely, if $\mathbf{x} \in \mathrm{conv}\{A_\pi(\mathbf{y}) : \pi \in \Pi(n)\}$, then $\mathbf{x}$ admits a convex representation of the form $\mathbf{x} = \sum_\pi \lambda_\pi A_\pi(\mathbf{y})$, whence $\mathbf{x} = (\sum_\pi \lambda_\pi A_\pi)(\mathbf{y})$. □

Remarkably, the relation of majorization gives rise to inequalities of the type (2.1.8) not only for the continuous convex functions of the form

$$F(x_1, \ldots, x_n) = \sum_{k=1}^{n} f(x_k),$$

but also for all quasiconvex functions $F(x_1, \ldots, x_n)$ which are invariant under the action of the permutation group.

Theorem 2.1.8. *(I. Schur [157]) If C is a convex set in $\mathbb{R}^n$ invariant under permutations and $F : C \to \mathbb{R}$ is a quasiconvex function invariant under permutations, then*

$$\boldsymbol{x} \prec_{HLP} \mathbf{y} \text{ implies } F(\mathbf{x}) \leq F(\mathbf{y}).$$

Proof. Indeed, according to Theorem 2.1.7, we have

$$F(\mathbf{x}) \leq \sup\{F(\mathbf{u}) : \mathbf{u} \in \operatorname{conv}\{\pi\mathbf{y} : \pi \in \Pi(n)\}\}$$
$$= \sup\{F(\pi\mathbf{y}) : \pi \in \Pi(n)\} = F(\mathbf{y}).$$

□

Two simple examples of quasiconvex functions invariant under permutations that are not convex are

$$\log\left(\sum_{k=1}^{n} x_k\right) \text{ and } \frac{x_1^\alpha + \cdots + x_n^\alpha}{\sqrt[n]{x_1 \cdots x_n}} \text{ for } \alpha > 0 \text{ and } x_1, \ldots, x_n > 0.$$

For more examples, notice that if f and g are two functions defined on a convex set $C \subset \mathbb{R}^n$ such that f is positive and convex and g is strictly positive and concave, then f/g is quasiconvex.

An illustration of Theorem 2.1.8 is offered by the following result due to R. F. Muirhead [109]:

Theorem 2.1.9. *(Muirhead's inequality) If $\mathbf{x}$ and $\mathbf{y}$ are two vectors in $\mathbb{R}^n$ such that $\mathbf{x} \prec_{HLP} \mathbf{y}$ and $\alpha_1, \ldots, \alpha_n$ are strictly positive numbers, then*

$$\sum_{\pi \in \Pi(n)} \alpha_{\pi(1)}^{x_1} \cdots \alpha_{\pi(n)}^{x_n} \leq \sum_{\pi \in \Pi(n)} \alpha_{\pi(1)}^{y_1} \cdots \alpha_{\pi(n)}^{y_n}, \tag{2.1.9}$$

the sum being taken over all permutations π of the set $\{1, \ldots, n\}$.

Actually, Muirhead has considered only the case where $\mathbf{x}$ and $\mathbf{y}$ have positive integer components. The extension to the case of real components is due to G. H. Hardy, J. E. Littlewood and G. Pólya [68].

Proof. Put $\mathbf{w} = (\log a_1, \ldots, \log a_n)$. Then we have to prove that

$$\sum_{\pi \in \Pi(n)} e^{\langle \mathbf{x}, \pi\mathbf{w} \rangle} \leq \sum_{\pi \in \Pi(n)} e^{\langle \mathbf{y}, \pi\mathbf{w} \rangle}.$$

This follows from Theorem 2.1.8 because the function $\mathbf{u} \to \sum_{\pi \in \Pi(n)} e^{\langle \mathbf{u}, \pi\mathbf{w} \rangle}$ is convex and invariant under permutations. □

The converse of Theorem 2.1.9 also works: If the inequality (2.1.9) is valid for all $\alpha_1, \ldots, \alpha_n > 0$, then $\mathbf{x} \prec_{HLP} \mathbf{y}$. Indeed, the case where $\alpha_1 = \cdots = \alpha_n > 0$ gives us

$$\alpha_1^{\sum_{k=1}^n x_k} \leq \alpha_1^{\sum_{k=1}^n y_k},$$

so that

$$\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k,$$

since $\alpha_1 > 0$ is arbitrary. Denote by $\mathcal{P}$ the set of all subsets of $\{1, \ldots, n\}$ of size k and take $\alpha_1 = \cdots = \alpha_k > 1$ and $\alpha_{k+1} = \cdots = \alpha_n = 1$. By our hypotheses,

$$\sum_{S \in \mathcal{P}} \alpha_1^{\sum_{i \in S} x_i} \leq \sum_{S \in \mathcal{P}} \alpha_1^{\sum_{i \in S} y_i}.$$

If $\sum_{j=1}^{k} x_j^{\downarrow} > \sum_{j=1}^{k} y_j^{\downarrow}$, this leads to a contradiction for α_1 large enough. Thus $\mathbf{x} \prec_{HLP} \mathbf{y}$.

2.1.2 Several applications to linear algebra

A well known result in linear algebra states that the trace of a matrix equals the sum of its eigenvalues. What more can be said about the possible diagonal entries of the real symmetric matrices having a fixed set of eigenvalues? The answer to this problem is given by the Schur-Horn theorem:

Theorem 2.1.10. *Suppose that*

$$\mathbf{d} = (d_1, \ldots, d_n) \text{ and } \lambda = (\lambda_1, \ldots, \lambda_n)$$

are two vectors in $\mathbb{R}^n$. Then there is a real symmetric matrix with diagonal entries $d_1, \ldots, d_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if $\mathbf{d} \prec_{HLP} \lambda$.

Schur's contribution was the striking remark concerning the implication of majorization to this matter.

Lemma 2.1.1. *(I. Schur [157]) Let $B \in M_n(\mathbb{R})$ be a self-adjoint matrix with diagonal elements $b_{11}, \ldots, b_{nn}$ and eigenvalues $\lambda_1, \ldots, \lambda_n$. Then*

$$(b_{11}, \ldots, b_{nn}) \prec_{HLP} (\lambda_1, \ldots, \lambda_n).$$

Proof. Using the spectral decomposition theorem, $B = UDU^*$, where $U = (u_{kj})_{k,j}$ is orthogonal and D is diagonal, with diagonal entries $\lambda_1, \ldots, \lambda_n$. The diagonal elements of B are

$$b_{kk} = \langle B\mathbf{e}_k, \mathbf{e}_k \rangle = \sum_{j=1}^{n} \lambda_i u_{kj}^2 = \sum_{j=1}^{n} a_{kj} \lambda_j,$$

where $a_{kj} = u_{kj}^2$; as usually, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denote the natural basis of $\mathbb{R}^n$. Since U is orthogonal, the matrix $A = (a_{kj})_{k,j}$ is doubly stochastic and Theorem 2.1.5 applies. □

Since the function log is concave, from Theorem 2.1.3 and Lemma 2.1.1 we infer the following inequality:

Corollary 2. *(Hadamard's determinant inequality) If B is an $n \times n$-dimensional positive matrix with diagonal elements $b_{11}, \ldots, b_{nn}$ and eigenvalues $\lambda_1, \ldots, \lambda_n$, then*

$$\prod_{k=1}^{n} b_{kk} \geq \prod_{k=1}^{n} \lambda_k.$$

A. Horn [71] has proved a converse to Lemma 2.1.1, which led to the statement of Theorem 2.1.10:

Lemma 2.1.2. *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are two vectors in* $\mathbb{R}^n$ *such that* $\mathbf{x} \prec_{HLP} \mathbf{y}$, *then there exists a symmetric matrix* B *such that the entries of* $\mathbf{x}$ *are the diagonal elements of* B *and the entries of* $\mathbf{y}$ *are the eigenvalues of* B.

Proof. We follow here the argument of W. Arveson and R. Kadison [13].

Step 1: If $B = (b_{ij})_{i,j} \in M_n(\mathbb{R})$ is a symmetric matrix with diagonal $\mathbf{d}$, then for every T-transform T there exists a unitary matrix U such that UBU^* has diagonal $T\mathbf{d}$.

Indeed, suppose that

$$T = (1 - \cos^2\theta)I + (\sin^2\theta)P_\pi,$$

where π is a permutation that interchanges i_0 and j_0. Then the matrix $U = (u_{ij})_{i,j}$, obtained by modifying four entries of the identity matrix as follows,

$$u_{i_0 i_0} = i\sin\theta, \quad u_{i_0 j_0} = -\cos\theta,$$

$$u_{j_0 i_0} = i\cos\theta, \quad u_{j_0 j_0} = \sin\theta,$$

is unitary and a straightforward computation shows that the diagonal of UBU^* is equal to $(1 - \cos^2\theta)d + (\sin^2\theta)P_\pi(d)$.

Step 2: Let $\Lambda = \mathrm{Diag}(\mathbf{y})$. By Theorem 2.1.5, $\mathbf{x}$ can be obtained from $\mathbf{y}$ by successive applications of finitely many T-transformations,

$$\mathbf{x} = T_m T_{m-1} \cdots T_1 \mathbf{y}.$$

By Step 1, there is a unitary matrix U_1 such that $U_1 \Lambda U_1^*$ has diagonal $T_1 y$. Similarly, there is a unitary matrix U_2 such that $U_2(U_1 \Lambda U_1^*)U_2^*$ has diagonal $T_2(T_1 y)$. Iterating this argument, we obtain a self-adjoint matrix

$$B = (U_m U_{m-1} \cdots U_1)\Lambda(U_m U_{m-1} \cdots U_1)^*$$

whose diagonal elements are the entries of $\mathbf{x}$ and the eigenvalues are the entries of $\mathbf{y}$. $\square$

2.1.3 The Schur-convexity property

Taking into account the Theorem 2.1.8, the quasiconvex functions $f : \mathbb{R}^n \to \mathbb{R}$ invariant under permutations are isotonic with respect to the relation of majorization, that is,

$$\mathbf{x} \prec_{HLP} \mathbf{y} \text{ implies } f(\mathbf{x}) \leq f(\mathbf{y}). \tag{2.1.10}$$

The implication (2.1.10) holds true beyond the framework of quasiconvex functions invariant under permutations. Simple examples such as $f(x_1, x_2) = -x_1 x_2$ on $\mathbb{R}^2$ can justify this implication. This led I. Schur [157] to initiate a systematic study of the functions that verify the property (2.1.10).

Definition 2.1.3. *A function* $f : C \to \mathbb{R}$ *defined on a set invariant under permutations is called Schur-convex if*

$$\mathbf{x} \prec_{HLP} \mathbf{y} \textit{ implies } f(\mathbf{x}) \leq f(\mathbf{y}).$$

If in addition $f(\mathbf{x}) < f(\mathbf{y})$ *whenever* $\mathbf{x} \prec_{HLP} \mathbf{y}$ *but* $\mathbf{x}$ *is not a permutation of* $\mathbf{y}$, *then* f *is said to be strictly Schur-convex.*

We call the function f *(strictly) Schur-concave if* $-f$ *is (strictly) Schur-convex.*

Every Schur-convex function defined on a set C invariant under permutations is a function invariant under permutations. This is a consequence of the fact that $\mathbf{x} \prec_{HLP} P_\pi(\mathbf{x})$ and $P_\pi(\mathbf{x}) \prec_{HLP} \mathbf{x}$ for every vector $\mathbf{x} \in \mathbb{R}^n$ and every permutation matrix $P_\pi \in M_n(\mathbb{R})$.

The Schur-convex functions contain a large variety of examples such as:

Example 1. *If $J : \mathbb{R}^n \to \mathbb{R}$ is a function (strictly) increasing in each variable and $f_1, \ldots, f_n$ are (strictly) Schur-convex functions on $\mathbb{R}^n$, then the function $j(\mathbf{x}) = J(f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ is (strictly) Schur-convex on $\mathbb{R}^n$. Some particular examples of strictly Schur-convex functions are:*

$$\max\{x_1, \ldots, x_n\} \text{ and } \log\left(\sum_{k=1}^{n} x_k^2\right) \text{ on } \mathbb{R}^n;$$

$$-\prod_{k=1}^{n} x_k \text{ on } (0, \infty)^n.$$

Theorem 2.1.11. *(The Schur-Ostrowski criterion of Schur-convexity) Let I be a nonempty open interval. A differentiable function $f : I^n \to \mathbb{R}$ is Schur-convex if and only if it fulfils the following two conditions:*

(i) f is invariant under permutations;

(ii) for every $\mathbf{x} \in I^n$ and $i, j \in \{1, \ldots, n\}$ we have

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i}(\mathbf{x}) - \frac{\partial f}{\partial x_j}(\mathbf{x})\right) \geq 0.$$

Proof. Necessity. For (i), see the comments after Definition 2.1.3. This reduces the verification of (ii) to the case where $i = 1$ and $j = 2$. Fix arbitrarily $\mathbf{x} \in I^n$ and choose $\epsilon > 0$ sufficiently small such that

$$\mathbf{x}(t) = ((1-t)x_1 + tx_2, tx_1 + (1-t)x_2, x_3, \ldots, x_n) \in D, \tag{2.1.11}$$

for $t \in (0, \epsilon]$. Then $\mathbf{x}(t) \prec_{HLP} \mathbf{x}$, which yields $f(\mathbf{x}(t)) \leq f(\mathbf{x})$. Therefore

$$0 \geq \lim_{t \to 0} \frac{f(\mathbf{x}(t)) - f(\mathbf{x})}{t} = \left.\frac{df(\mathbf{x}(t))}{dt}\right|_{t=0} = -(x_1 - x_2)\left(\frac{\partial f}{\partial x_1}(\mathbf{x}) - \frac{\partial f}{\partial x_2}(\mathbf{x})\right).$$

Sufficiency. We have to prove that $\mathbf{y} \prec_{HLP} \mathbf{x}$ implies $f(\mathbf{y}) \leq f(\mathbf{x})$. According to Theorem 2.1.5, it suffices to consider the case where

$$\mathbf{y} = ((1-s)x_1 + sx_2, sx_1 + (1-s)x_2, x_3, \ldots, x_n),$$

for some $s \in \left[0, \frac{1}{2}\right]$. Consider $x(t)$ as in formula (2.1.11). Then

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= \int_0^s \frac{d}{dt} f(\mathbf{x}(t))dt \\ &= -\int_0^s (x_1 - x_2)\left(\frac{\partial f}{\partial x_1}(\mathbf{x}(t)) - \frac{\partial f}{\partial x_2}(\mathbf{x}(t))\right)dt \\ &= -\int_0^s \frac{pr_1\mathbf{x}(t) - pr_2\mathbf{x}(t)}{1-2t}\left(\frac{\partial f}{\partial x_1}(\mathbf{x}(t)) - \frac{\partial f}{\partial x_2}(\mathbf{x}(t))\right)dt, \end{aligned}$$

where pr_k denotes the projection to the k-th coordinate. According to condition (ii), $f(\mathbf{y}) - f(\mathbf{x}) \leq 0$, and the proof is complete. $\square$

The elementary symmetric functions of n variables are defined by the formulas

$$\begin{aligned}
e_0(x_1, x_2, \ldots, x_n) &= 1,\\
e_1(x_1, x_2, \ldots, x_n) &= x_1 + x_2 + \cdots + x_n,\\
&\vdots\\
e_k(x_1, x_2, \ldots, x_n) &= \sum_{1\le i_1<\cdots<i_k\le n} x_{i_1}\ldots x_{i_k},\\
&\vdots\\
e_n(x_1, x_2, \ldots, x_n) &= x_1x_2\cdots x_n.
\end{aligned}$$

Clearly, they are invariant under permutations. A small computation shows that

$$\frac{\partial}{\partial x_i} e_k(x_1, \ldots, x_n) = e_{k-1}(x_1, \ldots, \widehat{x_i}, \ldots, x_n)$$

and

$$\frac{\partial}{\partial x_i} e_k(x_1, \ldots, x_n) - \frac{\partial}{\partial x_j} e_k(x_1, \ldots, x_n) = -(x_i - x_j)e_{k-2}(x_1, \ldots, \widehat{x_i}, \ldots, \widehat{x_j}, \ldots, x_n),$$

where the cup indicates the omission of the coordinate underneath. This leads us to the following consequence of Theorem 2.1.11:

Corollary 3. *(I. Schur) All elementary symmetric*

$$e_k(x_1, x_2, \ldots, x_n)$$

of n variables are Schur-concave on $\mathbb{R}^n_+$.

Since

$$\left(\frac{x_1 + \cdots + x_n}{n}, \ldots, \frac{x_1 + \cdots + x_n}{n}\right) \prec_{HLP} (x_1, \ldots, x_n),$$

for every $(x_1, \ldots, x_n) \in \mathbb{R}^n_+$, we infer from Corollary 3 that

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{1\le i_1<\cdots<i_k\le n} x_{i_1}\ldots x_{i_k} \le \binom{n}{k}\left(\frac{x_1 + \cdots + x_n}{n}\right)^k,$$

for $k = n$ we retrieve the AM-GM inequality.

Remark 6. *The relation of majorization is closely related to the duality of cones, more precisely, to the fact that the dual cone of the monotone cone*

$$\mathbb{R}^n_{\geq} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \geq \cdots \geq x_n\}$$

is the cone

$$(\mathbb{R}^n_{\geq})^* = \left\{\mathbf{y} \in \mathbb{R}^n : \sum_{k=1}^{m} y_k \geq 0 \text{ for } m = 1, \ldots, n-1, \text{ and } \sum_{k=1}^{n} y_k = 0\right\}.$$

Indeed, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{\geq}$, then

$$\mathbf{x} \prec_{HLP} \mathbf{y} \text{ if and only if } \mathbf{y} - \mathbf{x} \in (\mathbb{R}^n_{\geq})^*.$$

This leads to the more general concept of majorization with respect to a convex cone C in a real vector space V,

$$\mathbf{x} \prec_C \mathbf{y} \text{ if and only if } \mathbf{y} - \mathbf{x} \in C^*,$$

and, implicitly, to a generalization of Schur convexity. In the case of self-dual cones C (like $\mathbb{R}^n_+$ and $Sym^+(n, \mathbb{R})$), the corresponding concept of Schur convexity coincides with that of a function which is monotone increasing on C.

2.1.4 On vector majorization in $\mathbb{R}^n$

The usual relation of majorization, described in Section 2.1.1 as a relation between strings of real numbers can be easily generalized as a relation between strings of weighted vectors in $\mathbb{R}^n$. This was done by S. Sherman [160], inspired by the equivalence of conditions (i) and (ii) in Theorem 2.1.5.

Definition 2.1.4. *The relation of majorization*

$$(\mathbf{x}_1, \ldots, \mathbf{x}_m; \lambda_1, \ldots, \lambda_m) \prec_{Sh} (\mathbf{y}_1, \ldots, \mathbf{y}_n; \mu_1, \ldots, \mu_n) \tag{2.1.12}$$

between two strings of weighted points in $\mathbb{R}^n$ is defined by asking the existence of an $m \times n$-dimensional matrix $A = (a_{ij})_{i,j}$ such that

$$a_{ij} \geq 0 \text{ for } (i, j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}, \tag{2.1.13}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \text{ for } i = 1, \ldots, m, \tag{2.1.14}$$

$$\mu_j = \sum_{i=1}^{m} a_{ij}\lambda_i \text{ for } j = 1, \ldots n, \tag{2.1.15}$$

and

$$\mathbf{x}_i = \sum_{j=1}^{n} a_{ij}\mathbf{y}_j \text{ for } i = 1, \ldots, m. \tag{2.1.16}$$

We assume that all weights λ_i and μ_j belong to $[0, 1]$ and

$$\sum_{i=1}^{m} \lambda_i = \sum_{j=1}^{n} \mu_j = 1.$$

The matrices verifying the conditions (2.1.13) *and* (2.1.14) *are called stochastic on rows. When $m = n$ and all weights λ_i and μ_j are equal to each other, then the condition* (2.1.15) *assures the stochasticity on columns, so in that case we deal with doubly stochastic matrices.*

Remark 7. *The relation of majorization introduced by Definition 2.1.4 can be restated as a relation between probability measures, letting*

$$\sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec_{Sh} \sum_{j=1}^{n} \mu_j \delta_{y_j},$$

if the conditions (2.1.13) - (2.1.16) *hold. In this context, the condition* (2.1.16) *means that*

$$x_i = bar\left(\sum_{j=1}^{n} a_{ij}\delta_{y_j}\right) \text{ for } i = 1, \dots, m,$$

a fact that allows easily to define the relation of majorization between probability measures not only in $\mathbb{R}^n$*, but also in any space where the notion of barycenter of a probability measure makes sense. Notice also that the conditions* (2.1.14) *and* (2.1.16) *imply that*

$$\mathbf{x}_1, \dots, \mathbf{x}_m \in conv\{\mathbf{y}_1, \dots, \mathbf{y}_n\}.$$

The following theorem provides a large extension of the Hardy-Littlewood-Pólya inequality of majorization:

Theorem 2.1.12. *(S. Sherman [160]) Suppose that* $\sum_{i=1}^{m} \lambda_i \delta_{x_i}$ *and* $\sum_{j=1}^{n} \mu_j \delta_{y_j}$ *are two Borel probability measures on* $\mathbb{R}^n$*. Then the following assertions are equivalent:*

(i) $\sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec_{Sh} \sum_{j=1}^{n} \mu_j \delta_{y_j}$*;*

(ii) $\mathbf{x}_1, \dots, \mathbf{x}_m \in conv\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ *and every continuous convex functions* f *defined on* $conv\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ *(or to a larger convex subset of* $\mathbb{R}^n$*) verifies the inequality*

$$\sum_{i=1}^{m} \lambda_i f(\mathbf{x}_i) \le \sum_{j=1}^{n} \mu_j f(\mathbf{y}_j).$$

2.1.5 Convex type inequalities

We present a first example describes how convex functions relate the arithmetic means of the subfamilies of a given triplet of numbers.

Theorem 2.1.13. *(Popoviciu's inequality [144]) If* $f : I \to \mathbb{R}$ *is a continuous function, then* f *is convex if and only if*

$$\frac{f(x) + f(y) + f(z)}{3} + f\left(\frac{x+y+z}{3}\right) \ge \frac{2}{3}\left[f\left(\frac{x+y}{2}\right) + f\left(\frac{y+z}{2}\right) + f\left(\frac{z+x}{2}\right)\right],$$

for all $x, y, z \in I$*. In the variant of strictly convex functions, the above inequality is strict except for* $x = y = z$*.*

Proof. Necessity. This part does not make use of the continuity of f. It suffices to consider the case where f is the absolute value function, that is, to show that

$$|x| + |y| + |z| + |x+y+z| \ge |x+y| + |y+z| + |z+x|, \tag{2.1.17}$$

for all $x, y, z \in I$. This is an immediate consequence of the order properties of real numbers.

A second approach, based on the polynomial identity,

$$x^2 + y^2 + z^2 + (x+y+z)^2 = (x+y)^2 + (y+z)^2 + (z+x)^2,$$

has the advantage to extend to the framework of Euclidian spaces. Indeed,

$$(|x|+|y|+|z|+|x+y+z|-|x+y|-|y+z|-|z+x|)\times(|x|+|y|+|z|+|x+y+z|)$$
$$\begin{aligned}&=(|x|+|y|-|x+y|)(|z|+|x+y+z|-|x+y|)\\&+(|y|+|z|-|y+z|)(|x|+|x+y+z|-|y+z|)\\&+(|z|+|x|-|z+x|)(|y|+|x+y+z|-|z+x|)\geq 0,\end{aligned}$$

and the necessity part is done.

Sufficiency. Popoviciu's inequality when applied for $y=z$, yields the following substitute for the condition of midpoint convexity:

$$\frac{1}{4}f(x)+\frac{3}{4}f\left(\frac{x+2y}{3}\right)\geq f\left(\frac{x+y}{2}\right)\quad\text{for all}\quad x,y\in I.$$

□

Theorem 2.1.14. *(Abel's partial summation formula) If $(a_k)_{k=1}^n$ and $(b_k)_{k=1}^n$ are two families of complex numbers, then*

$$\sum_{k=1}^n a_kb_k=\sum_{k=1}^{n-1}\left[(a_k-a_{k+1})\left(\sum_{j=1}^k b_j\right)\right]+a_n\left(\sum_{j=1}^n b_j\right).$$

Corollary 4. *(The Abel-Steffensen inequality [162]) If $x_1,\dots,x_n$ and $y_1,\dots,y_n$ are two families of real numbers that verify one of the following two conditions*

(i) $x_1\geq\cdots\geq x_n\geq 0$ *and* $\sum_{k=1}^j y_k\geq 0$ *for all* $j\in\{1,2,\dots,n\}$,

(ii) $0\leq x_1\leq\cdots\leq x_n$ *and* $\sum_{k=j}^n y_k\leq 0$ *for all* $j\in\{1,2,\dots,n\}$,

then

$$\sum_{k=1}^n x_ky_k\geq 0.$$

Therefore, if $x_1,\dots,x_n$ is a monotonic family and $y_1,\dots,y_n$ is a family of real numbers such that

$$0\leq\sum_{k=1}^j y_k\leq\sum_{k=1}^n y_k,$$

for $j=1,\dots,n$, then we have

$$\left(\min_{1\leq k\leq n}x_k\right)\sum_{k=1}^n y_k\leq\sum_{k=1}^n x_ky_k\leq\left(\max_{1\leq k\leq n}x_k\right)\sum_{k=1}^n y_k.$$

Another result whose proof can be considerably simplified by the piecewise linear approximation of convex functions in the following generalization of Jensen's inequality, due to J. F. Steffensen [162]. Unlike Jensen's inequality, it allows the use of negative weights.

Theorem 2.1.15. *(The Jensen-Steffensen inequality) Suppose that $x_1, \ldots, x_n$ is a monotonic family of points in an interval $[a, b]$ and $w_1, \ldots, w_n$ are real weights such that*

$$\sum_{k=1}^{n} w_k = 1 \text{ and } 0 \leq \sum_{k=1}^{m} w_k \leq \sum_{k=1}^{n} w_k,$$

for every $m \in \{1, \ldots, n\}$.
Then every convex function f defined on $[a, b]$ verifies the inequality

$$f\left(\sum_{k=1}^{n} w_k x_k\right) \leq \sum_{k=1}^{n} w_k f(x_k).$$

Proof. We may reduce ourselves to the case of absolute value function. Assuming the ordering $x_1 \leq \cdots \leq x_n$, we infer that

$$0 \leq x_1^+ \leq \cdots \leq x_n^+ \text{ and } x_1^- \geq \cdots \geq x_n^- \geq 0.$$

According to Corollary 4,

$$\sum_{k=1}^{n} w_k x_k^+ \geq 0 \text{ and } \sum_{k=1}^{n} w_k x_k^- \geq 0$$

which yields

$$\left|\sum_{k=1}^{n} w_k x_k\right| \leq \sum_{k=1}^{n} w_k |x_k|$$

and the proof is complete. □

The integral version of Jensen-Steffensen inequality can be established in the same manner, using integration by part instead of Abel's partial summation formula.

Theorem 2.1.16. *(The integral version of Jensen-Steffensen inequality) Suppose that $g : [a, b] \to \mathbb{R}$ is a monotone function and $w : [a, b] \to \mathbb{R}$ is an integrable function such that*

$$0 \leq \int_a^x w(t)dt \leq \int_a^b w(t)dt = 1 \text{ for every } x \in [a, b].$$

Then every convex function f defined on an interval I that includes the range of g verifies the inequality

$$f\left(\int_a^b g(t)w(t)dt\right) \leq \int_a^b f(g(t))w(t)dt.$$

Another application of Abel's partial summation formula is as follows.

Theorem 2.1.17. *(The discrete form of Hardy-Littlewood rearrangement inequality [68]) Let $x_1, \ldots, x_n, y_1, \ldots, y_n$ be real numbers. Then*

$$\sum_{k=1}^{n} x_k^{\downarrow} y_{n-k+1}^{\downarrow} \leq \sum_{k=1}^{n} x_k y_k \leq \sum_{k=1}^{n} x_k^{\downarrow} y_k^{\downarrow},$$

where $x_1^{\downarrow} \geq \cdots \geq x_n^{\downarrow}$ denotes the decreasing rearrangement of a family $x_1, \ldots, x_n$ of real numbers.

Proof. Notice first that we may assume that $x_1 \geq \cdots \geq x_n \geq 0$. Then, apply Abel's partial summation formula. □

2.2 New versions of uniformly convex functions via quadratic complete homogeneous symmetric polynomials

In this section, we introduce new versions of uniformly convex functions, namely h_d strongly (weaker) convex functions. Based on the positivity of complete homogeneous symmetric polynomials with even degree, recently studied in [153, 165], we introduce stronger and weaker versions of uniformly convexity. In this context, we recover well-known type inequalities such as: Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities. Some final remarks related to Sherman's and Ingham's type inequalities are also discussed.

The topic we address in this section of the thesis is related to the study of a new family of convex functions which is based on the positivity property of complete homogeneous symmetric polynomials with even degree. The study of the positivity of symmetric polynomial functions goes back to an old paper of Hunter [74], where a tricky argument is proposed. Afterwards, in [165] was considered a genuinely different way to establish the positivity of such polynomials. Moreover, two different ideas are presented in [153], one of them being based on a Schur-convexity argument and the other one following a method with divided differences. Fine estimates on the norms on complex matrices induced by complete homogeneous symmetric polynomials are obtained in [4] and [37].

The family of complete homogeneous symmetric polynomials with n variables $x_1, \ldots, x_n$ and degree $d \in \mathbb{N}$ is defined as follows

$$h_0(x_1, \ldots, x_n) = 1,$$
$$h_d(x_1, \ldots, x_n) := \sum_{1 \leq i_1 \leq \cdots \leq i_d \leq n} x_{i_1} \cdots x_{i_d} \qquad (d \geq 1).$$

A key strategy to prove the positivity of h_d, for all even degrees $d \geq 2$, consists of using Schur-convexity and majorization techniques. The concept of majorization lies in the core of a powerful topic of research with interesting recent results. In this regard, we just enumerate few of them: a necessary and sufficient condition for a linear map to preserve group majorizations can be found in [131]; new majorization results are studied in [83, 132]; interesting properties on superquadratic functions related to Jensen–Steffensen's inequality are obtained in [1]. All these ideas are also based on the theory of uniformly convex functions, which in addition gives the possibility to define the concept of majorization into the spaces of curved geometry (see [126]). More results on this topic can be found in [112, 113, 114, 117, 124].

In order to present the current settings we address in this section, let us introduce the above mentioned concepts of stronger and weaker h_d convexity for functions defined on $\mathbb{R}^n$.

The key point to introduce this new versions of uniformly convexity is based on a positivity result given in [153], which asserts that: if $d \geq 2$ is an even natural index, then

$$h_d(x_1, x_2, \ldots, x_n) \geq 0 \qquad (x_1, \ldots, x_n \in \mathbb{R}). \tag{2.2.1}$$

Based on (2.2.1) we define a new class of convex functions by considering a perturbation of convex functions with complete homogeneous symmetric polynomials.

Definition 2.2.1. *Let $C > 0$ and let $d \geq 2$ be an even natural number. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be h_d strongly convex with modulus C if the function $f(\cdot) - C\, h_d(\cdot)$ is convex. Similarly, a function $f : \mathbb{R}^n \to \mathbb{R}$ is called h_d weakly convex with modulus C if the function $f(\cdot) + C\, h_d(\cdot)$ is convex.*

The above definition is inspired from the notion ω-m-star convex function (see, for instance, [96]). In order to motivate the concept of h_2 strongly/weakly convex function we first recall the notion of uniformly convex function.

Definition 2.2.2. *Let $C > 0$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be uniformly convex with modulus C if $f(\cdot) - C\, \|\cdot\|^2$ is convex. Equivalently, the function f is uniformly convex with modulus C if and only if the following inequality holds*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda(1-\lambda)\,\|\mathbf{x}-\mathbf{y}\|^2\,, \tag{2.2.2}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

A first objective of this section consists in showing that (2.2.2) holds similarly, even in the context of h_2 strongly convexity (see Proposition 2.2.1).

We simply remark that, based on the following estimate

$$\frac{1}{2}\,\|x\|^2 \leq h_2(x) \leq \frac{n+1}{2}\|x\|^2 \qquad (x \in \mathbb{R}^n,\, n \in \mathbb{N}^*),$$

we have, in general, the equivalence between the concepts of uniformly convexity and h_d strongly convexity. But, in particular, going deeply to the modulus, we cannot prove the existence of two positive constants C_1 and C_2 such that: a function is h_2 strongly convex with modulus C_1 if and only if it is uniformly convex with modulus C_2. In this sense, other related remarks and examples are given in Proposition 2.2.2. On the other hand, we prove that Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities for h_2 strongly convex functions produce different constants in the error right hand term (comparing with the ones obtained in the case of uniformly convex functions).

A second objective of this section consists in studying the general and difficult case, i.e. h_d strongly convexity, for any even natural number $d \geq 4$. We show that (2.2.2) can be also extended in this very general case (see Theorem 2.2.1). We use fine estimates and computations in order to get h_d versions of Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities. Other classical inequalities are also obtained in this section (see Theorem 2.2.2, but also Proposition 2.2.3 - Proposition 2.2.10). This confirm that the family of h_d strongly convex functions lead to new ideas of further research.

We strongly consider that the new concept and results presented in this section can be used to establish connections and further applications related to other important scientific achievements in literature (see [2, 3, 15, 96, 133, 169]), as we have shortly explained in the final subsection of this section. The motivation of studying such functions is successfully accomplished by Jensen's, Hardy-Littlewood-Polya's and Popoviciu's inequalities in a very general case.

It is worth mentioning that even if h_d polynomials cannot induce a norm (for example, in majorization settings, we have that for any two vectors satisfying $x \prec y$, $h_2(y) \geq h_2(x) + h_2(y-x)$, see Lemma 2.2.1) similar results (as in the uniform convexity settings) can be obtained. An interesting approach related to this idea was given in [135], for the case of strongly convexity and hence, further research can be now done in our settings. This is why we compare our results, all along the section, with the ones obtained for classical strongly convexity.

The rest of the section is organised as follows: In Subsection 2.2.1 we present our main results, starting with the case of h_2 strongly convex functions, where we deduce similar estimates as the ones for uniformly convex functions (Jensen's type inequalities). The rest of the section is devoted to the h_d strongly convex functions: Theorem 2.2.1 represents a fundamental result which confirm that the family of h_d strongly convex functions have a nice behaviour, even in the general case. Majorization properties gluing together with h_d strongly Schur convex functions are revealed in Subsection 2.2.2, where Hardy-Littlewood-Polya's and Popoviciu's type inequalities are obtained. The last subsection is devoted to some final remarks related to Sherman's inequalities for ω-m star convex functions. Further consequences on Ingham's type inequalities in control theory are also discussed.

2.2.1 New results on h_d strongly convex functions

In this subsection we present important results concerning h_d strongly convex functions, where $d \geq 2$ is any even natural number. Despite the main result of this subsection presented in Theorem 2.2.1 we deduce other interesting consequences which confirm the relevance of h_d strongly convexity.

Firstly, we present a surprisingly property of h_2 strongly convex functions which consists of an inequality similar to (2.2.2).

Proposition 2.2.1. *Let $C > 0$. Then, the function $f : \mathbb{R}^n \to \mathbb{R}$ is h_2 strongly convex with modulus C if and only if*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda(1-\lambda)h_2(\mathbf{x}-\mathbf{y}), \tag{2.2.3}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$. Taking into account the identity

$$h_2(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2 + \sum_{1\leq i<j\leq n} x_i x_j,$$

the property that $f(\cdot) - C\, h_2(\cdot)$ is convex can be expressed as

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) - C\sum_{i=1}^{n}((1-\lambda)x_i + \lambda y_i)^2 - C\sum_{1\leq i<j\leq n}((1-\lambda)x_i + \lambda y_i)((1-\lambda)x_j + \lambda y_j)$$

$$\leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C(1-\lambda)\left(\sum_{i=1}^{n} x_i^2 + \sum_{1\leq i<j\leq n} x_i x_j\right) - C\lambda\left(\sum_{i=1}^{n} y_i^2 + \sum_{1\leq i<j\leq n} y_i y_j\right).$$

Based on the fact that

$$\sum_{i=1}^{n}((1-\lambda)x_i + \lambda y_i)^2 + \sum_{1\leq i<j\leq n}((1-\lambda)x_i + \lambda y_i)((1-\lambda)x_j + \lambda y_j)$$

$$-(1-\lambda)\left(\sum_{i=1}^{n} x_i^2 + \sum_{1\leq i<j\leq n} x_i x_j\right) - \lambda\left(\sum_{i=1}^{n} y_i^2 + \sum_{1\leq i<j\leq n} y_i y_j\right)$$

$$= (\lambda^2 - \lambda)\left(\sum_{i=1}^{n}(x_i - y_i)^2 + \sum_{1\le i<j\le n}(x_i - y_i)(x_j - y_j)\right)$$
$$= (\lambda^2 - \lambda)h_2(\mathbf{x} - \mathbf{y}),$$

we obtain

$$h_2((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) - (1-\lambda)h_2(\mathbf{x}) - \lambda h_2(\mathbf{y}) = h_2(\mathbf{x} - \mathbf{y})(\lambda^2 - \lambda).$$

Hence, the proof of (2.2.3) is complete. □

In order to emphasize the difference between the notions of uniformly convexity and strongly h_d convexity we present the case of very particular family of polynomial functions.

Proposition 2.2.2. *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a function defined as*

$$f(x, y, z) = \frac{a}{2}x^2 + ay^2 + \frac{a}{2}z^2 + (a - \alpha)xz + b \qquad (a, b, \alpha \in \mathbb{R}).$$

Then, for any $a \in (0, \infty)$ and $b \in \mathbb{R}$ there exists $C > 0$ and $\alpha > 0$ such that f is h_2 strongly convex with modulus C. Futhermore, for any $\varepsilon > 0$ there exist a, b, α such that f is not uniformly convex with modulus ε.

Proof. The Hessian matrix for the function f is given by

$$H_f = \begin{pmatrix} a & 0 & a-\alpha \\ 0 & 2a & 0 \\ a-\alpha & 0 & a \end{pmatrix}.$$

Notice that, for any $a > 0$ there exist $C > 0$ and $\alpha \in \mathbb{R}$ such that the Hessian matrix corresponding to the function $g(\cdot) = f(\cdot) - Ch_2(\cdot)$ is positive definite, and therefore g is convex. More precisely, we have

$$H_g = \begin{pmatrix} a-2C & -C & a-\alpha-C \\ -C & 2a-2C & -C \\ a-\alpha-C & -C & a-2C \end{pmatrix},$$

and thus, for $n \ge 4$, let $C = \frac{n}{2n+1}a$ and let $\alpha = \frac{n-1}{n}C = \frac{n-1}{2n+1}a$ in order to fulfill the desired property for H_g, i.e. $\det(H_g) = \frac{2C^3}{n}\left(n^2 - 3n - 3\right) > 0$ (other two diagonal determinants of order 1 and 2 are also strictly positive).

On the other hand, the Hessian matrix of the function $g_{unif}(\cdot) = f(\cdot) - \varepsilon\|\cdot\|^2$ is given by

$$H_{g_{unif}} = \begin{pmatrix} a-2\varepsilon & 0 & a-\alpha \\ 0 & 2a-2\varepsilon & 0 \\ a-\alpha & 0 & a-2\varepsilon \end{pmatrix}.$$

In this case, we get

$$\det(H_{g_{unif}}) = 2(a-\varepsilon)\left((a-2\varepsilon)^2 - (a-\alpha)^2\right),$$

which is strictly negative for all $\epsilon > 0$, $a \ge \alpha$ and $\alpha = \frac{n-1}{n}\varepsilon$ (or $\alpha = 0$). □

In the general case, for any even natural number $d \ge 2$, we get a natural but powerful extension of Proposition 2.2.1.

Theorem 2.2.1. *Let $C > 0$ and let $d \geq 2$ be an even natural number. Then, the function $f : \mathbb{R}^n \to \mathbb{R}$ is h_d strongly convex with modulus $C > 0$ if and only if*

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y})) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - C\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}}h_d(\mathbf{x}-\mathbf{y}), \tag{2.2.4}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$.

Moreover, for each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$ we have

$$h_d((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) - (1-\lambda)h_d(\mathbf{x}) - \lambda h_d(\mathbf{y}) \leq -\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}}h_d(\mathbf{x}-\mathbf{y}). \tag{2.2.5}$$

Proof. Notice that the case $d = 2$ and $n \geq 1$ is already proved in Proposition 2.2.1. Our proof strategy is based on an induction argument with respect to the even natural number $d \geq 2$, but also with respect to the number of variables $n \geq 1$. For the convenience of the reader, we take into account firstly the case $d = 4$. We consider this case in order to be confident on the mathematical induction argument.

Thus, by using Definition 2.2.1 for $d = 4$ we have to prove that for all $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$ we have

$$h_4((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) - (1-\lambda)h_4(\mathbf{x}) - \lambda h_4(\mathbf{y}) \leq -\lambda^2(1-\lambda)^2 h_4(\mathbf{x}-\mathbf{y}). \tag{2.2.6}$$

Firstly, we consider the case $n = 1$. In this particular case, for each $x, y \in \mathbb{R}$ and $\lambda \in [0,1]$, (2.2.6) becomes

$$((1-\lambda)x + \lambda y)^4 \leq (1-\lambda)x^4 + \lambda y^4 - \lambda^2(1-\lambda)^2(x-y)^4. \tag{2.2.7}$$

The above inequality is a consequence of the following computations

$$\begin{gathered}
((1-\lambda)x + \lambda y)^4 - (1-\lambda)x^4 - \lambda y^4 = \left((1-\lambda)^4 - (1-\lambda)\right)x^4 + 4(1-\lambda)^3\lambda x^3 y \\
+6(1-\lambda)^2\lambda^2x^2y^2 + 4(1-\lambda)\lambda^3xy^3 + \left(\lambda^4 - \lambda\right)y^4 \\
= 4\lambda(1-\lambda)^3x^3(y-x) + 4\lambda^3(1-\lambda)y^3(x-y) \\
+\lambda(1-\lambda)\left(3(1-\lambda)^2 - (1-\lambda) - 1\right)x^4 + 6\lambda^2(1-\lambda)^2x^2y^2 + \lambda(1-\lambda)\left(3\lambda^2 - \lambda - 1\right)y^4 \\
= \lambda(1-\lambda)(x-y)^2\left((-\lambda^2 + 3\lambda - 3)x^2 + (2\lambda^2 - 2\lambda - 2)xy - (\lambda^2 + \lambda + 1)y^2\right) \\
= -\lambda^2(1-\lambda)^2(x-y)^4 \\
-\lambda(1-\lambda)(x-y)^2\left((2\lambda^2 - 4\lambda + 3)x^2 + 2(-2\lambda^2 + 2\lambda + 1)xy + (2\lambda^2 + 1)y^2\right) \\
\leq -\lambda^2(1-\lambda)^2(x-y)^4,
\end{gathered}$$

where for the last inequality we have used that

$$(2\lambda^2 - 4\lambda + 3)x^2 + 2(-2\lambda^2 + 2\lambda + 1)xy + (2\lambda^2 + 1)y^2 \geq 0 \quad (x, y \in \mathbb{R},\ \lambda \in [0,1]).$$

Secondly, for $n \geq 2$, let us consider $\mathbf{x} = (x_1, \ldots, x_n)\,,\ \mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ and $\lambda \in [0,1]$. Thanks to the following expansion formulae

$$h_4(x_1, \ldots, x_n) = x_1\left(\sum_{1 \leq i \leq n} x_i h_2(x_i, x_{i+1} \ldots, x_n)\right) + h_4(x_2, \ldots, x_n)\,, \tag{2.2.8}$$

$$h_4\left(x_1,\dots,x_n\right)=\sum_{j=1}^{n-1}x_j\left(\sum_{j\le i\le n}x_ih_2(x_j,x_{i+1}\dots,x_n)\right)+h_4\left(x_n\right), \tag{2.2.9}$$

by using an induction argument, our aim is to reduce the number of variables of inequalities which are to be proved. More precisely, starting with $\mathbf{x}=(x_1,\dots,x_n)$, $\mathbf{y}=(x_1,\dots,x_n)$ in $\mathbb{R}^n$ we show that (2.2.6) holds for vectors $\widetilde{\mathbf{x}}=(x_2,\dots,x_n)$, $\widetilde{\mathbf{y}}=(y_2,\dots,y_n)$ in $\mathbb{R}^{n-1}$, and then we continue the process in a similar way until we reach the end in $\mathbb{R}$. Finally, we use the fact (2.2.7) holds for $n=1$ and the proof of (2.2.6) is then complete. For each $j\ge 1$, let us consider

$$T_j^2(x)=x_j\sum_{j\le i\le n}x_ih_2(x_i,x_{i+1}\dots,x_n). \tag{2.2.10}$$

Then, by using (2.2.8) and (2.2.9), we can reduce the proof of (2.2.6) to the following inequality

$$\sum_{j=1}^{n-1}\left(T_j^2((1-\lambda)\mathbf{x}+\lambda\mathbf{y})-(1-\lambda)T_j^2(\mathbf{x})-\lambda T_j^2(\mathbf{y})\right)\le-\lambda^2(1-\lambda)^2\sum_{j=1}^{n-1}T_j^2(\mathbf{x}-\mathbf{y}). \tag{2.2.11}$$

Therefore, once we prove (2.2.11) we can then use (2.2.8) in a repeated way and based on (2.2.7) and (2.2.9) at the end of process, we finally get

$$\begin{aligned}h_4((1-\lambda)\mathbf{x}+\lambda\mathbf{y})-(1-\lambda)h_4(\mathbf{x})-\lambda h_4(\mathbf{y})&=T_1^2((1-\lambda)\mathbf{x}+\lambda\mathbf{y})-(1-\lambda)T_1^2(\mathbf{x})-\lambda T_1^2(\mathbf{y})\\&\quad+h_4((1-\lambda)\widetilde{\mathbf{x}}+\lambda\widetilde{\mathbf{y}})-(1-\lambda)h_4(\widetilde{\mathbf{x}})-\lambda h_4(\widetilde{\mathbf{y}})\\&\quad\cdots\cdots\cdots\\&\le-\lambda^2(1-\lambda)^2\sum_{i=1}^{n-1}\left(T_j^2((1-\lambda)\mathbf{x}+\lambda\mathbf{y})-(1-\lambda)T_j^2(\mathbf{x})-\lambda T_j^2(\mathbf{y})\right)\\&\quad+h_4((1-\lambda)x_n+\lambda y_n)-(1-\lambda)h_4(x_n)-\lambda h_4(y_n)\\&\le-\lambda^2(1-\lambda)^2\left(\sum_{i=1}^{n-1}T_i^2(\mathbf{x}-\mathbf{y})+h_4(x_n-y_n)\right)=-\lambda^2(1-\lambda)^2h_4(\mathbf{x}-\mathbf{y}),\end{aligned}$$

and hence, (2.2.6) holds.

We prove now (2.2.11) by replacing $\widetilde{x}_i=(x_i,x_{i+1}\dots,x_n)$, $\widetilde{y}_i=(y_i,y_{i+1}\dots,y_n)$ in relation (2.2.10), i. e.

$$\begin{aligned}&\sum_{j=1}^{n-1}\left(T_j^2((1-\lambda)\mathbf{x}+\lambda\mathbf{y})-(1-\lambda)T_j^2(\mathbf{x})-\lambda T_j^2(\mathbf{y})\right)\\&=\sum_{j=1}^{n-1}((1-\lambda)x_j+\lambda y_j)\sum_{i=j}^{n}((1-\lambda)x_i+\lambda y_i)\,h_2\left((1-\lambda)\widetilde{x}_i+\lambda\widetilde{y}_i\right)\\&\quad-(1-\lambda)\sum_{j=1}^{n-1}x_j\sum_{i=j}^{n}x_ih_2\left(\widetilde{x}_i\right)-\lambda\sum_{j=1}^{n-1}y_j\sum_{i=j}^{n}y_ih_2\left(\widetilde{y}_i\right)\\&=-\lambda(1-\lambda)\sum_{j=1}^{n-1}\sum_{i=j}^{n}p_{ij}h_2\left((1-\lambda)\widetilde{x}_i+\lambda\widetilde{y}_i\right)+\lambda(1-\lambda)\sum_{j=1}^{n-1}\sum_{i=j}^{n}p_{ij}\end{aligned}$$

$$\times \left(\frac{(1-\lambda)x_j x_i + \lambda y_j y_i}{\lambda(1-\lambda)p_{ij}} h_2\left((1-\lambda)\widetilde{x}_i + \lambda \widetilde{y}_i\right) - \frac{(1-\lambda)x_j x_i}{\lambda(1-\lambda)p_{ij}} h_2\left(\widetilde{x}_i\right) - \frac{\lambda y_j y_i}{\lambda(1-\lambda)p_{ij}} h_2\left(\widetilde{y}_i\right) \right)$$

$$\leq -\lambda(1-\lambda) \sum_{j=1}^{n-1} \sum_{i=j}^{n} p_{ij} \lambda(1-\lambda) h_2\left(\widetilde{x}_i - \widetilde{y}_i\right) = -\lambda^2(1-\lambda)^2 \sum_{j=1}^{n-1} \sum_{i=j}^{n} T_j^2(\mathbf{x}-\mathbf{y}),$$

where the last estimates are due to the fact that

$$((1-\lambda)x_j + \lambda y_j)\left((1-\lambda)x_i + \lambda y_i\right) = -\lambda(1-\lambda)p_{ij} + (1-\lambda)x_j x_i + \lambda y_j y_i,$$

$$p_{ij} = (x_j - y_j)(x_i - y_i).$$

Hence, the proof of (2.2.6) is now complete.

Note that, at this stage, we have proved (2.2.4) and (2.2.5) only for the case for $d = 4$. In addition, we have also obtained the following generalized inequality in the case $n = 2$

$$\sum_{j=1}^{n-1} \left(T_j^d((1-\lambda)\mathbf{x} + \lambda \mathbf{y}) - (1-\lambda)T_j^d(\mathbf{x}) - \lambda T_j^d(\mathbf{y}) \right) \leq -\lambda^{\frac{d}{2}+1}(1-\lambda)^{\frac{d}{2}+1} \sum_{j=1}^{n-1} T_j^d(\mathbf{x}-\mathbf{y}), \qquad (2.2.12)$$

where

$$T_j^d(x) = x_j \sum_{j \leq i \leq n} x_i h_d(x_i, x_{i+1} \ldots, x_n). \qquad (2.2.13)$$

In what follows we prove (2.2.4) and (2.2.5) in the general case, by using (2.2.11), for any even natural number $d \geq 2$, and by using again the induction argument, this time with respect to the even parameter d. In order to do this, we suppose that for a fixed natural even number $d \geq 2$ the following inequality holds

$$h_d((1-\lambda)\mathbf{x} + \lambda \mathbf{y}) - (1-\lambda)h_d(\mathbf{x}) - \lambda h_d(\mathbf{y}) \leq -\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}} h_d(\mathbf{x}-\mathbf{y}). \qquad (2.2.14)$$

By using (2.2.13) we get

$$\sum_{j=1}^{n-1} T_j^d((1-\lambda)\mathbf{x} + \lambda \mathbf{y}) - (1-\lambda)T_j^d(\mathbf{x}) - \lambda T_j^d(\mathbf{y})$$

$$= \sum_{j=1}^{n-1} ((1-\lambda)x_j + \lambda y_j) \sum_{i=j}^{n} ((1-\lambda)x_i + \lambda y_i)\, h_d\left((1-\lambda)(x_i, \ldots, x_n) + \lambda(y_i, \ldots, y_n)\right)$$

$$-(1-\lambda) \sum_{j=1}^{n-1} x_j \sum_{i=j}^{n} x_i h_d(x_i, x_{i+1} \ldots, x_n) - \lambda \sum_{j=1}^{n-1} y_j \sum_{i=j}^{n} y_i h_d(y_i, y_{i+1} \ldots, y_n)$$

$$= -\lambda(1-\lambda) \sum_{j=1}^{n-1} \sum_{i=j}^{n} p_{ij} h_d((x_i - y_i, x_{i+1} - y_{i+1} \ldots, x_n - y_n)) + \lambda(1-\lambda) \sum_{j=1}^{n-1} \sum_{i=j}^{n} p_{ij}$$

$$\times \left(\frac{(1-\lambda)x_j x_i + \lambda y_j y_i}{\lambda(1-\lambda)p_{ij}} h_d\left((1-\lambda)(x_i, x_{i+1} \ldots, x_n) + \lambda(y_i, y_{i+1} \ldots, y_n)\right) \right.$$

$$\left. - \frac{(1-\lambda)x_j x_i}{\lambda(1-\lambda)p_{ij}} h_d\left((x_i, x_{i+1} \ldots, x_n)\right) - \frac{\lambda y_j y_i}{\lambda(1-\lambda)p_{ij}} h_d\left((y_i, y_{i+1} \ldots, y_n)\right) \right)$$

$$\leq -\lambda(1-\lambda)\sum_{j=1}^{n-1}\sum_{i=j}^{n} p_{ij}\lambda^{\frac{d}{2}}(1-\lambda)^{\frac{d}{2}} h_d((x_i-y_i, x_{i+1}-y_{i+1}\ldots, x_n-y_n))$$

$$= -\lambda^{\frac{d}{2}+1}(1-\lambda)^{\frac{d}{2}+1}\sum_{j=1}^{n-1}\sum_{i=j}^{n} T_j^d(\mathbf{x}-\mathbf{y}),$$

where we have used the following estimates

$$p_{ij} = (x_j - y_j)(x_i - y_i)$$

and

$$((1-\lambda)x_1 + \lambda y_1)((1-\lambda)x_i + \lambda y_i) = -\lambda(1-\lambda)p_{ij} + (1-\lambda)x_1 x_i + \lambda y_1 y_i.$$

We can now continue inductively the sequence of inequalities in order to get

$$h_{d+2}((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) - (1-\lambda)h_{d+2}(\mathbf{x}) - \lambda h_{d+2}(\mathbf{y}) \leq -\lambda^{\frac{d}{2}+1}(1-\lambda)^{\frac{d}{2}+1}h_{d+2}(\mathbf{x}-\mathbf{y}), \tag{2.2.15}$$

by using as the main tool the following generalised expansion formulas

$$h_{d+2}(x_1,\ldots,x_n) = x_1\left(\sum_{1\leq i\leq n} x_i h_d(x_i, x_{i+1}\ldots, x_n)\right) + h_{d+2}(x_2,\ldots,x_n), \tag{2.2.16}$$

$$h_{d+2}(x_1,\ldots,x_n) = \sum_{i=1}^{n-1} x_i\left(\sum_{i\leq j\leq n} x_j h_d(x_j, x_{j+1}\ldots, x_n)\right) + h_{d+2}(x_n). \tag{2.2.17}$$

More precisely, we obtain

$$\begin{aligned}
&h_{d+2}((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) - (1-\lambda)h_{d+2}(\mathbf{x}) - \lambda h_{d+2}(\mathbf{y})\\
&= T_1^d((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) - (1-\lambda)T_1^d(\mathbf{x}) - \lambda T_1^d(\mathbf{y})\\
&+h_d((1-\lambda)\widetilde{\mathbf{x}}+\lambda\widetilde{\mathbf{y}}) - (1-\lambda)h_d(\widetilde{\mathbf{x}}) - \lambda h_d(\widetilde{\mathbf{y}})\\
&\cdots\cdots\cdots\\
&\leq -\lambda^2(1-\lambda)^2\sum_{i=1}^{n-1}\Big(T_j^d((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) - (1-\lambda)T_j^d(\mathbf{x}) - \lambda T_j^2(\mathbf{y})\Big)\\
&+h_d((1-\lambda)x_n+\lambda y_n) - (1-\lambda)h_d(x_n) - \lambda h_d(y_n)
\end{aligned}$$

$$\leq -\lambda^{\frac{d}{2}+1}(1-\lambda)^{\frac{d}{2}+1}\left(\sum_{i=1}^{n-1} T_i^d(\mathbf{x}-\mathbf{y}) + h_d(x_n-y_n)\right) = -\lambda^{\frac{d}{2}+1}(1-\lambda)^{\frac{d}{2}+1}h_d(\mathbf{x}-\mathbf{y}),$$

and hence, (2.2.6) holds in the general case $d \geq 2$. Therefore, the proof of (2.2.4) and (2.2.5) follows easily. □

We end this subsection by presenting an inequality of Jensen's type in the case of h_d strongly convex functions, for any even natural number $d \geq 2$.

Proposition 2.2.3. *(Jensen's type inequality for h_d strongly convexity) Let $C > 0$ and let $d \geq 2$ be an even natural number. If $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ is a given function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly convex with modulus C on I^n then, for all $x_1, \ldots, x_n \in I$, the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - C\frac{1}{n}\binom{n+d-1}{d}\left(\left(\frac{x_1 + \cdots + x_n}{n}\right)^d - \frac{x_1^d + \cdots + x_n^d}{n}\right). \tag{2.2.18}$$

Proof. Based on convexity properties of $F(\cdot) - Ch_d(\cdot)$ we have that

$$F\left(\frac{\mathbf{x}^1 + \cdots + \mathbf{x}^n}{n}\right) - Ch_d\left(\frac{\mathbf{x}^1 + \cdots + \mathbf{x}^n}{n}\right) \leq \frac{F(\mathbf{x}^1) - Ch_d(\mathbf{x}^1) + \cdots + F(\mathbf{x}^n) - Ch_d(\mathbf{x}^n)}{n},$$

where $\mathbf{x}^1 = (x_1, \ldots, x_1)$, $\mathbf{x}^2 = (x_2, \ldots, x_2)$, $\ldots$, $\mathbf{x}^n = (x_n, \ldots, x_n)$ belong to I^n.

Consequently, we get

$$nf\left(\frac{x_1 + \cdots + x_n}{n}\right) - C\binom{n+d-1}{d}\left(\frac{x_1 + \cdots + x_n}{n}\right)^d \leq f(x_1) + \cdots + f(x_n) - \frac{C}{n}\binom{n+d-1}{d}\left(x_1^d + \cdots + x_n^d\right).$$

Finally, we easily get (2.2.18) and the proof is complete. □

2.2.2 Open problems on h_d strongly convexity

In this subsection we are dealing with majorization results concerning h_d strongly convex functions. More precisely, we obtain Jensen's type inequalities in two different contexts and we notice that even the results are of the same type, the constants appearing in front of error term are different. We also succeed to develop majorization results, such as Hardy-Littlewood-Polya's and Popoviciu's inequalities.

In order to compare Jensen's type inequalities for h_d strongly convex functions and uniformly convex functions we present the following result (which can be seen as an easily consequence of the results from [170]).

Proposition 2.2.4. *(Jensen's type inequality for uniform convexity) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be such that $F : I^n \to \mathbb{R}$, defined as $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$, is uniformly convex with modulus C. Then, for all $x_1, \ldots, x_n \in I$ the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - \frac{C}{n}\sum_{1 \leq i < j \leq n}(x_i - x_j)^2. \tag{2.2.19}$$

Remark 8. *Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ such that $F : I^n \to \mathbb{R}$, defined as $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$, is h_2 strongly convex with modulus C. Then, for all $x_1, \ldots, x_n \in I$, by using* (2.2.18) *the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - C\frac{n+1}{2n^2} \sum_{1 \leq i < j \leq n} (x_i - x_j)^2. \tag{2.2.20}$$

Moreover, we can also take in Proposition 2.2.3 other type of functions, which are related to the construction of h_d polynomials. In order to do this, let us introduce

$$F_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} f(x_{i_1}) \cdots f(x_{i_k}) \qquad (k = 2, \ldots, n). \tag{2.2.21}$$

Based on a similar strategy we can deduce the following result.

Proposition 2.2.5. *Let $C > 0, k \geq 2$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be such that $F_k : I^n \to \mathbb{R}$, which is defined in* (2.2.21)*, is h_2 strongly convex with modulus C. Then, for all $x_1, \ldots, x_n \in I$ the following inequality holds*

$$\binom{n}{k}\left(f^k\left(\frac{x_1 + \cdots + x_n}{n}\right) - \frac{f^k(x_1) + \cdots + f^k(x_n)}{n}\right) \leq -C\frac{n+1}{2n} \sum_{1 \leq i < j \leq n} (x_i - x_j)^2. \tag{2.2.22}$$

Proof. The proof is very similar with the one presented in Proposition 2.2.3. □

In the second part of this subsection we present several majorization type inequalities in the context of h_d strongly convex functions. More precisely, we are dealing with extensions of Hardy-Littlewood-Polya's and Popoviciu's inequalities in the case of our new class of convex functions.

Let us consider $\mathbf{x}^{\downarrow}$ and $\mathbf{y}^{\downarrow}$ two vectors with the same entries as $\mathbf{x}$, respectively $\mathbf{y}$, expressed in decreasing order, as

$$x_1^{\downarrow} \geq \cdots \geq x_n^{\downarrow},\ y_1^{\downarrow} \geq \cdots \geq y_n^{\downarrow}.$$

We say that, the vector $\mathbf{x}$ is *majorized* by $\mathbf{y}$ (abbreviated, $\mathbf{x} \prec \mathbf{y}$) if

$$\begin{aligned} &\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow} \qquad (1 \leq k \leq n-1), \\ &\sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}. \end{aligned} \tag{2.2.23}$$

More details and applications concerning the majorization theory can be found in [104]. We refer to the monotonicity with respect to the majorization order, the so called Schur-convex property, which has been introduced by I. Schur in 1923.

Definition 2.2.3. *The function $f : A \to \mathbb{R}$, where A is a symmetric subset of $\mathbb{R}^n$, is called Schur-convex if $\mathbf{x} \prec \mathbf{y}$ implies $f(\mathbf{x}) \leq f(\mathbf{y})$.*

A simple computation tool (see, for instance, [104]) which is used to study the Schur-convexity property of a function is given as follows. For any symmetric function $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$ having continuous partial derivatives on $I^n = I \times I \times \cdots \times I$, the Schur-convexity property is reduced to check the following inequality

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right) \geq 0 \qquad (1 \leq i,\ j \leq n,\ x_i, x_j \in I).$$

We introduce now the notions of h_d strongly Schur convexity and uniformly Schur convexity.

Definition 2.2.4. *Let $C > 0$. A function $f : I^n \to \mathbb{R}$ is said to be h_d strongly Schur-convex with modulus C if the function $f(\cdot) - C\, h_d(\cdot)$ is Schur-convex.*

Definition 2.2.5. *Let $C > 0$. A function $f : I^n \to \mathbb{R}$ is said to be uniformly Schur-convex with modulus C if the function $f(\cdot) - C\, \|\cdot\|^2$ is Schur-convex.*

Proposition 2.2.6. *Let A be a symmetric subset of $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be a symmetric function with continuous partial derivatives. Then, f is strongly Schur-convex if and only if*

$$\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right)(x_i - x_j) \geq C(x_i - x_j)^2 \qquad (1 \leq i,\ j \leq n,\ x_i, x_j \in \mathbb{R}).$$

Proof. The proof easily follows from

$$\frac{\partial h_2(x_1, x_2, \ldots, x_n)}{\partial x_1} - \frac{\partial h_2(x_1, x_2, \ldots, x_n)}{\partial x_2} = (x_1 - x_2).$$

□

We first remark that a similar Jensen's type inequalities is obtained by using this time majorization arguments and obtaining another constants in front of the right hand error term.

Proposition 2.2.7. *(Jensen's type inequality via Hardy-Littlewood-Polya's inequality) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. Then, for all $x_1, \ldots, x_n \in I$ the following inequality holds*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} - \frac{C}{2n^2} \sum_{1 \leq i < j \leq n} (x_i - x_j)^2. \tag{2.2.24}$$

Proof. Based on the following well-known majorization result

$$\left(\frac{x_1 + ... + x_n}{n}, ..., \frac{x_1 + \cdots + x_n}{n}\right) \prec (x_1, \ldots, x_n)$$

and by using Schur convexity property of $F(\cdot) - Ch_2(\cdot)$ we get

$$nf\left(\frac{x_1 + \cdots + x_n}{n}\right) - Ch_2\left(\frac{x_1 + \cdots + x_n}{n}, \ldots, \frac{x_1 + \cdots + x_n}{n}\right)$$
$$\leq f(x_1) + \cdots + f(x_n) - Ch_2(x_1, \ldots, x_n),$$

which yields (2.2.24). □

We are able now to deduce and to compare two different versions of Hardy-Littlewood-Polya's majorization theorem, one for uniform convexity case and another one for h_d strongly convexity case.

We start by proving an useful lemma, which represents the triangle inequality in the reverse way, in majorization settings. In this context, connections with the subdifferential concept can be established. For more details, see [135].

Lemma 2.2.1. *For each $\mathbf{x} \prec \mathbf{y}$ on $\mathbb{R}^n$ the following inequality holds*

$$h_2(\mathbf{y}) \geq h_2(\mathbf{x}) + h_2(\mathbf{y} - \mathbf{x}). \tag{2.2.25}$$

Proof. By using (2.2.23) we get

$$h_2(\mathbf{y}) - h_2\mathbf{x}) = \frac{1}{2}\left(\left(\sum_{i=1}^n y_i^\downarrow\right)^2 - \left(\sum_{i=1}^n x_i^\downarrow\right)^2 + \sum_{i=1}^n (y_i^\downarrow)^2 - \sum_{i=1}^n (x_i^\downarrow)^2\right)$$

$$= h_2(\mathbf{x} - \mathbf{y}) + \sum_{i=1}^n (y_i^\downarrow - x_i^\downarrow)(S + x_i^\downarrow),$$

where $S = \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. Thus, we have to prove that the sum from the right hand side of the above inequality is positive, and this can be done by using a classical telescopic sums trick as follows

$$\sum_{i=1}^n (y_i^\downarrow - x_i^\downarrow)(S + x_i^\downarrow) = \sum_{i=1}^n (y_i^\downarrow - x_i^\downarrow)x_i^\downarrow$$

$$= \sum_{i=1}^{n-1} \left(x_i^\downarrow - x_{i+1}^\downarrow\right)\left(y_1^\downarrow - x_1^\downarrow + \cdots + y_i^\downarrow - x_i^\downarrow\right) + x_n^\downarrow \sum_{i=1}^n \left(y_i^\downarrow - x_i^\downarrow\right) \geq 0,$$

where for the last estimates we have essentially used (2.2.23). ☐

Proposition 2.2.8. *(Hardy-Littlewood-Polya's inequality for uniformly convexity) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is uniformly Schur convex with modulus C on I^n. If $\mathbf{x} \prec \mathbf{y}$ on I^n the following inequality holds*

$$\sum_{i=1}^n f(y_i) \geq \sum_{i=1}^n f(x_i) + C\sum_{i=1}^n (y_i - x_i)^2. \tag{2.2.26}$$

Proof. Since the function $f(\cdot) - C|\cdot|^2$ is convex, by using the properties of the subdifferential we get the existence of $\lambda_i \in \partial g(x_i)$, i. e.

$$f(y) - Cy^2 \geq f(x_i) - Cx_i^2 + (\lambda_i - 2Cx_i)(y - x_i) \qquad (i = 1, \ldots, n,\ y \in I).$$

Hence, summing all the above inequalities for each $y = y_i$ we obtain

$$\sum_{i=1}^n f(y_i) \geq \sum_{i=1}^n f(x_i) + C\sum_{i=1}^n \left(y_i^2 - x_i^2\right) + \sum_{i=1}^n (\lambda_i - 2Cx_i)(y_i - x_i),$$

which leads to

$$\sum_{i=1}^{n} f(y_i) \geq \sum_{i=1}^{n} f(x_i) + C\sum_{i=1}^{n}(y_i - x_i)^2 + \sum_{i=1}^{n}\lambda_i(y_i - x_i).$$

In order to complete the proof of (2.2.26), it is enough to show that $\sum_{i=1}^{n}\lambda_i(y_i - x_i) \geq 0$. But, this can be done by using a similar argument as in the proof of Lemma 2.2.1. $\square$

Theorem 2.2.2. *(Hardy-Littlewood-Polya's inequality for strongly h_d functions) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. If $\mathbf{x} \prec \mathbf{y}$ on I^n the following inequality holds*

$$\sum_{i=1}^{n} f(y_i) \geq \sum_{i=1}^{n} f(x_i) + Ch_2(\mathbf{y} - \mathbf{x}). \tag{2.2.27}$$

Proof. By using the definition of h_d strongly Schur convexity with modulus C we get

$$\sum_{i=1}^{n} f(y_i) \geq \sum_{i=1}^{n} f(x_i) + C\left(h_2(\mathbf{y}) - h_2(\mathbf{x})\right),$$

which finally gives (2.2.27), as a direct consequence of Lemma 2.2.1. $\square$

In the end of this subsection we give some natural extensions of Popoviciu's inequalities for h_2 strongly convex functions, but also for uniformly convex functions.

Proposition 2.2.9. *(Popoviciu's type inequality for h_2 strongly convexity) Let $C > 0$ and let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a function such that $F(x_1, \ldots, x_n) = f(x_1) + \cdots + f(x_n)$ is h_d strongly Schur convex with modulus C on I^n. Then, for all $x, y, z \in I$ the following inequality holds*

$$\begin{aligned} \frac{f(x) + f(y) + f(z)}{3} + f\left(\frac{x+y+x}{3}\right) \geq & \frac{2}{3}\left(f\left(\frac{x+y}{2}\right) + f\left(\frac{x+z}{2}\right) + f\left(\frac{y+z}{2}\right)\right) \\ & + \frac{C}{36}\left((x-y)^2 + (y-z)^2 + (x-z)^2\right). \end{aligned} \tag{2.2.28}$$

Proof. We begin by recalling the following majorization relation (see, for instance, [114]), i. e. for all $x, y, z \in \mathbb{R}$ we have that $u \prec v$, where

$$u = \left(\frac{x+y}{2}, \frac{x+y}{2}, \frac{x+z}{2}, \frac{x+z}{2}, \frac{y+z}{2}, \frac{y+z}{2}\right),$$
$$v = \left(\frac{x+y+z}{3}, \frac{x+y+z}{3}, \frac{x+y+z}{3}, x, y, z\right).$$

By Schur-convexity properties of $F(\cdot) - Ch_2(\cdot)$ we obtain

$$\begin{aligned} f(x) + f(y) + f(z) + 3f\left(\frac{x+y+x}{3}\right) \geq 2\left(f\left(\frac{x+y}{2}\right) + f\left(\frac{x+z}{2}\right) + f\left(\frac{y+z}{2}\right)\right) \\ + C\left(h_2(v) - h_2(u)\right). \end{aligned}$$

By using the following tricky identity

$$h_2(v) - h_2(u) = \frac{1}{2}h_2(x, y, z) - \frac{1}{3}(x+y+z)^2,$$

we deduce that (2.2.28) holds and the proof is complete. $\square$

Proposition 2.2.10. *(Popoviciu's type inequality for uniform convexity) Let $f : I \to \mathbb{R}$ and let $C > 0$ be such that $g(\cdot) = f(\cdot) - C|\cdot|^2$ is convex on I. Then, for all $x, y, z \in I$the following inequality holds*

$$\frac{f(x)+f(y)+f(z)}{3} + f\left(\frac{x+y+x}{3}\right) \geq \frac{2}{3}\left(f\left(\frac{x+y}{2}\right) + f\left(\frac{x+z}{2}\right) + f\left(\frac{y+z}{2}\right)\right)$$
$$+\frac{C}{18}\left((x-y)^2 + (y-z)^2 + (x-z)^2\right). \tag{2.2.29}$$

Proof. For each x, y, $z \in I$, by applying Popoviciu's classical inequality for the convex function $f(\cdot) - C|\cdot|^2$ we have that

$$f(x)+f(y)+f(z)+3f\left(\frac{x+y+x}{3}\right) \geq 2\left(f(\frac{x+y}{2}) + f(\frac{x+z}{2}) + f(\frac{y+z}{2})\right)$$
$$+C\left(x^2+y^2+z^2+3\left(\frac{x+y+x}{3}\right)^2 - 2\left(\frac{x+y}{2}\right)^2 - 2\left(\frac{x+z}{2}\right)^2 - 2\left(\frac{y+z}{2}\right)^2\right).$$

Finally, several computations in the last term of the above inequality gives (2.2.29) and the proof is complete. □

Remark 9. *Notice that, all inequalities in the above two propositions differ only by the constants appearing in the right hand error term, which can be related to the variance/dispersion in probability theory. We can also remark that different constants cannot appear here by only moving from one concept to the other one, since we have no equivalence between h_2 strongly convex functions with modulus C_1 and uniformly convex functions with modulus C_2. See also the case of Jensen's type inequalities. We end this remark be mentioning that, in all the above results, we can also consider the case of functions defined in* (2.2.21)*.*

Finally, we emphasize that the notion of h_d strongly convexity and the results presented in this subsection can also be seen in connection with other results existing in literature. We are strongly confident that the present section gives the possibility to develop other interesting results on this topic, such as in other relevant papers (see, for instance [2, 3, 15, 96, 133, 169]).

More precisely, our first proposal of new research is related to the class of ω-m-star convex functions, for which modulus function ω can be replaced with the polynomial function h_d. This is motivated by [96], where some similar properties are presented and for which we are able to express ω-m-star convex property, for some particular function ω, in terms of convexity of a suitable perturbed function. The second aim is to get Sherman's type inequalities for h_d strongly convex functions. Also, our ideas can also be extended on spaces related to other weaker notions (relative convexity, spaces with global nonpositive curvature, see [124, 126]).

On the other hand, new Ingham's type weighted inequalities are recently proved in [154, 155] by using the positivity of quadratic polynomials. The proofs are essentially based on an Ingham's proof technique inspired from [78]. As applications, the authors consider families of frequencies with relevance in the approximation of controls theory, for which the uniform (with respect to the mesh-size) controllability property of the semi-discrete model is proved, when the spurious frequencies (the gap between them tends to zero when the mesh size goes to zero) are eliminated.

The third future aim is to use the theory developed in this section to prove the positivity of a very general class of weighted symmetric polynomials. Then, as a direct consequence we may provide Ingham's type inequalities, when we eliminate frequencies which are very close to each other. We can also study the uniform boundedness of a sequence of discrete controls, related to such inequalities. This is needed in order to study the approximations of the controls of the continuous wave equation. See, for instance, [45, 77, 78, 81, 99, 100, 172]. We also expect that, a deep development of the results from this section to give us the possibility to prove Ingham's type inequalities in a very general framework, by considering h_d strongly convexity assumptions and eliminating the frequencies in the area where the gap is lost.

Finally, based on the fact that different kind of strongly and weakly convexity notions are studied in the literature, some clarifying remarks are needed. Starting with strongly convex and strongly quasy-convex functions introduced by Polyak [142] other notion of uniformly convex function at a point are studied in [169]. More precisely, in order to define this pointwise notion of convexity it was introduced a positive function δ depending on the term $\|x - y\|$, with have a similar role as the error term in the right hand side of (2.2.2). Note that, in our case, the error term of the form $h_d(x - y)$ is a function of several variables and cannot be seen as a function of the form $\delta(\|x - y\|)$. Moreover, in [82, 170, 171] the notion of convexifiable function (in the sense of Definition 2.2.1) is studied. In this context, similar inequalities, even of integral type, are obtained. See also [46], for other interesting results. Hence, the idea to consider positive symmetric polynomials instead of functions depending on the norm and the possibility to obtain similar results offer a new and fresh perspective within the topic of convexity.

2.3 The Hardy-Littlewood-Pólya inequality of majorization in the context of ω-m-star-convex functions

In this section, the Hardy-Littlewood-Pólya inequality of majorization is extended for ω-$\mathbf{m}$-star-convex functions to the framework of ordered Banach spaces. Several open problems which seem to be of interest for further extensions of the Hardy-Littlewood-Pólya inequality are also included. The Hardy-Littlewood-Pólya theorem of majorization is an important result in convex analysis that lies at the core of majorization theory, a subject that has attracted a great deal of attention due to its numerous applications in mathematics, statistics, economics, quantum information etc. See [103, 104, 114, 128, 139, 140] and [156] to cite just a few books treating this topic.

The relation of majorization was initially formulated as a relation between pairs of vectors with real entries rearranged downward, but nowadays its formulation as a preordering of probability measures.

For the reader's convenience we briefly recall here the most basic facts concerning the theory of majorization.

Given two discrete probability measures $\mu = \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k}$ and $\nu = \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$, supported by a compact interval $[a, b]$, we say that μ is *majorized* by ν (denoted $\mu \prec \nu$) if the following three conditions are fulfilled:

$$\begin{array}{ll} (M1) & \mathbf{x}_1 \geq \mathbf{x}_2 \geq \cdots \geq \mathbf{x}_N \\ (M2) & \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \leq \sum_{i=1}^{k} \lambda_i \mathbf{y}_i \quad \text{for } k = 1, \ldots, N; \text{ and} \\ (M3) & \sum_{i=1}^{N} \lambda_i \mathbf{x}_i = \sum_{i=1}^{N} \lambda_i \mathbf{y}_i. \end{array}$$

When only conditions $(M1)$ and $(M2)$ occur, we say that μ is *weakly majorized* by ν (denoted $\mu \prec_w \nu$).

Hardy, Littlewood and Pólya [68] used a stronger formulation of $(M1)$, by requiring also that $\mathbf{y}_1 \geq \mathbf{y}_2 \geq \cdots \geq \mathbf{y}_N$. Later, their result was improved by Maligranda, Pečarić and Persson [101] who were able to prove that

$$\mu \prec \nu \text{ implies } \int_a^b f \, \mathrm{d}\mu = \sum_{k=1}^{N} f(\mathbf{x}_k) \leq \int_a^b f \, \mathrm{d}\nu = \sum_{k=1}^{N} f(\mathbf{y}_k), \tag{HLP}$$

for all continuous convex functions $f : [a, b] \to \mathbb{R}$. Moreover, the same conclusion holds in the case of weak majorization and convex and nondecreasing functions.

Nowadays the inequality HLP is known as the Hardy-Littlewood-Pólya inequality of majorization.

In the early 1950s, the Hardy-Littlewood-Pólya inequality was extended by Sherman [160] to the case of continuous convex functions of a vector variable by using a much broader concept of majorization, based on matrices stochastic on lines. The full details can be found in [114], Theorem 4.7.3, p. 219. Over the years, many other generalizations in the same vein have been published. See, for example, [31, 117, 118, 124, 125, 126] and [133].

As was noticed in [112] and [113], the Hardy-Littlewood-Pólya inequality of majorization can be extended to the framework of convex functions defined on ordered Banach spaces alongside the conditions $(M1) - (M3)$. The aim of this section is to prove that the same works for the larger class of ω-m-star-convex functions.

The main features of these functions are discussed in subsection 2.3.1. In subsection 2.3.2 we present different types of majorization relations in ordered Banach spaces. The corresponding extensions of the Hardy-Littlewood-Pólya inequality constitute the objective of subsection 2.3.3. The section ends with mentioning several open problems which seem to be of interest for further extensions of the Hardy-Littlewood-Pólya inequality.

2.3.1 Preliminaries on ω-m-star-convex functions

Throughout this subsection E is a Banach space and C is a convex subset of it.

Definition 2.3.1. *Let m be a real parameter belonging to the interval $(0, 1]$. A function $\Phi : C \to \mathbb{R}$ is said to be a perturbed m-star-convex function with modulus $\omega : [0, \infty) \to \mathbb{R}$ (abbreviated as ω-m-star-convex function) if it fulfils an estimate of the form*

$$\Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y}) \leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}) - m\lambda(1-\lambda)\omega\left(\|\mathbf{x} - \mathbf{y}\|\right),$$

for all $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in (0, 1)$.

The ω-m-star-convex functions associated to an identically zero modulus will be called m-star-convex. They satisfy the inequality

$$\Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y}) \leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}),$$

for all $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in (0, 1)$.

Notice that the usual convex functions represent the particular case of m-star-convex functions where $m = 1$. On the other hand every convex function is m-star-convex (for every $m \in (0, 1]$) if $\mathbf{0} \in C$ and $\Phi(\mathbf{0}) \leq 0$. Indeed, we have

$$\begin{aligned}\Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y}) &= \Phi((1-\lambda)\mathbf{x} + \lambda m\mathbf{y} + (\lambda - \lambda m)\mathbf{0}) \\ &\leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}) + (\lambda - \lambda m)\boldsymbol{\Phi}(\mathbf{0}) \\ &= (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}).\end{aligned}$$

Every ω-m-star-convex function associated to a modulus $\omega \geq 0$ is necessarily m-star-convex. The ω-m-star-convex functions whose moduli ω are strictly positive except at the origin (where $\omega(0) = 0$) are usually called *uniformly m-star-convex.* In their case the definitory inequality is strict whenever $\mathbf{x} \neq \mathbf{y}$ and $\lambda \in (0, 1)$.

By reversing the inequalities, one obtains the notions of *ω-m-star-concave function* and *uniformly m-star-concave function.*

The theory of m-star-convex functions was initiated by Toader [166], who considered only the case of functions defined on real intervals. For additional results in the same setting see [108] and the references therein.

A simple example of a (16/17)-star-convex function which is not convex is

$$f : [0, \infty) \to \mathbb{R}, \quad f(x) = x^4 - 5x^3 + 9x^2 - 5x. \tag{2.3.1}$$

See [108], Example 2. Note that if $\Phi : C \to \mathbb{R}$ and $\Psi : C \to \mathbb{R}$ are ω-m-star-convex functions and $\alpha, \beta \in \mathbb{R}_+$, then

$$\alpha\Phi + \beta\Psi \text{ and } \sup\{\Phi, \Psi\}$$

are functions of the same nature. So is

$$\Phi \times \Psi : C \times C \to \mathbb{R}, \quad (\Phi \times \Psi)(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x}) + \Psi(\mathbf{y}).$$

The class of ω-m-star-convex functions is also stable under pointwise convergence (when it exists).

Assuming $C \subset E$ is a convex cone with vertex at the origin, the *perspective* of a function $f : C \to \mathbb{R}$ is the positively homogeneous function

$$\widetilde{f} : C \times (0, \infty) \to \mathbb{R}, \quad \widetilde{f}(\mathbf{x}, t) = tf\left(\frac{\mathbf{x}}{t}\right).$$

Lemma 2.3.1. *The perspective of every m-star-convex/concave function is a function of the same nature.*

Proof. Indeed, assuming (to make a choice) that f is ω-m-star-convex, then for all $(\mathbf{x}, s), (\mathbf{y}, t) \in C \times (0, \infty)$ and $\lambda \in [0, 1]$ we have

$$\begin{aligned} f\left(\frac{(1-\lambda)\mathbf{x} + \lambda m\mathbf{y}}{(1-\lambda)s + \lambda mt}\right) &= f\left(\frac{(1-\lambda)s}{(1-\lambda)s + \lambda mt} \cdot \frac{\mathbf{x}}{s} + \frac{\lambda mt}{(1-\lambda)s + \lambda mt} \cdot \frac{\mathbf{y}}{t}\right) \\ &\leq \frac{(1-\lambda)s}{(1-\lambda)s + \lambda mt} f\left(\frac{\mathbf{x}}{s}\right) + \frac{\lambda mt}{(1-\lambda)s + \lambda mt} f\left(\frac{\mathbf{y}}{t}\right)\end{aligned}$$

that is,

$$\widetilde{f}((1-\lambda)\mathbf{x}+\lambda m\mathbf{y},(1-\lambda)s+\lambda mt)\leq(1-\lambda)\widetilde{f}(\mathbf{x},s)+\lambda m\widetilde{f}(\mathbf{y},t).$$

$\square$

Lemma 2.3.1, allows us to easily produce nontrivial examples of m-star-convex functions of several variables with some nice properties. For example, starting from (2.3.1), we conclude that

$$\Phi(x,t)=\frac{x^4-5x^3t+9x^2t^2-5xt^3}{t^3}$$

is a (16/17)-star-convex function on $[0,\infty)\times(0,\infty)$.

Under the presence of Gâteaux differentiability, ω-m-star-convex functions generate specific gradient inequalities that play a prominent role in our generalization of the Hardy-Littlewood-Pólya inequality of majorization.

Lemma 2.3.2. *Suppose also that C is an open convex subset of the Banach space E and $\Phi : C\to\mathbb{R}$ is a function both Gâteaux differentiable and ω-m-star-convex. Then*

$$m\Phi(\mathbf{y})\geq\Phi(\mathbf{x})+d\Phi(\mathbf{x})(m\mathbf{y}-\mathbf{x})+m\omega\left(\|\mathbf{x}-\mathbf{y}\|\right), \tag{2.3.2}$$

for all points $\mathbf{x},\mathbf{y}\in C$.

Proof. Indeed, we have

$$\frac{\Phi((1-\lambda)\mathbf{x}+m\lambda\mathbf{y})-\Phi(\mathbf{x})}{\lambda}\leq-\Phi(\mathbf{x})+m\Phi(\mathbf{y})-m(1-\lambda)\omega\left(\|\mathbf{x}-\mathbf{y}\|\right)$$

and the proof ends with passing to the limit as $\lambda\to 0+$. $\square$

Remark 10. *Lemma 2.3.2 shows that the critical points $\mathbf{x}$ of the differentiable ω-m-star–convex functions are those for which $\omega\geq 0$ fulfil the property*

$$m\inf_{\mathbf{y}\in C}\Phi(\mathbf{y})\geq\Phi(\mathbf{x}).$$

Unlike the case of convex functions of one real variable, when the isotonicity of the differential is automatic, for several variables, this is not necessarily true in the case of a differentiable convex function of a vector variable. See [112], Remark 4.

In this section we deal with functions defined on ordered Banach spaces, that is, on real Banach spaces endowed with order relations $\leq$ that make them ordered vector spaces such that positive cones are closed and

$$\mathbf{0}\leq\mathbf{x}\leq\mathbf{y}\text{ implies }\|\mathbf{x}\|\leq\|\mathbf{y}\|.$$

The Euclidean N-dimensional space $\mathbb{R}^N$ has a natural structure of an ordered Banach space associated to coordinatewise ordering. The usual sequence spaces c_0, c, ℓ^p (for $p\in[1,\infty]$) and the function spaces $C(K)$ (for K a compact Hausdorff space) and $L^p(\mu)$ (for $1\leq p\leq\infty$ and μ a σ-additive positive measure) are also examples of ordered Banach spaces (with respect to coordinatewise/pointwise ordering and natural norms).

A map $T : E \to F$ between two ordered vector spaces is called *isotone* (or *order preserving*) if

$$\mathbf{x} \leq \mathbf{y} \text{ in } E \text{ implies } T(\mathbf{x}) \leq T(\mathbf{y}) \text{ in } F$$

and *antitone* (or *order reversing*) if $-T$ is isotone. When T is a linear operator, T is isotone if and only if T maps positive elements into positive elements (abbreviated, $T \geq 0$).

For basic information on ordered Banach spaces see [113]. The interested reader may also consult the classical books of Aliprantis and Tourky [9] and Meyer-Nieberg [106].

As was noticed by Amann [10], Proposition 3.2, p. 184, the Gâteaux differentiability offers a convenient way to recognize the property of isotonicity of functions acting on ordered Banach spaces: the positivity of the differential. We state here his result (following the version given in [112], Lemma 4):

Lemma 2.3.3. *Suppose that E and F are two ordered Banach spaces, C is a convex subset of E with nonempty interior* $\operatorname{int} C$ *and $\Phi : C \to F$ is a convex function, continuous on C and Gâteaux differentiable on* $\operatorname{int} C$. *Then Φ is isotone on C if and only if $\Phi'(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in \operatorname{int} C$.*

Remark 11. *If the ordered Banach space E has finite dimension, then the statement of Lemma* 2.3.3 *remains valid when the interior of C is replaced by the relative interior of C. See* [114]*, Exercise* 6*, p.* 81.

As was noticed in [108], Example 7, the function

$$\gamma : (-\infty, 1] \to \mathbb{R}, \quad \gamma(x) = -2x^3 + 5x^2 + 6x$$

is convex on $(-\infty, 5/6]$, concave on $[5/6, 1]$, and m-star-convex on $(-\infty, 1]$, with $m = 27/28$. The last assertion follows from a formula due to Mocanu,

$$m = \inf \left\{ \frac{x\gamma'(x) - \gamma(x)}{y\gamma'(x) - \gamma(y)} : y\gamma'(x) - \gamma(y),\ x, y \in I \right\},$$

mentioned at the bottom of page 72 in [108].

Proceeding like in Lemma 2.3.1, one can prove that the function associated to γ,

$$\Upsilon : (-\infty, 1] \times [1, \infty) \to \mathbb{R}, \quad \Upsilon(x, y) = -\frac{2x^3}{y^2} + \frac{5x^2}{y} + 6x,$$

is 27/28-star-convex. The function Υ is also Gateaux differentiable, with

$$(x, y) = \left(\frac{1}{y^2} \left(-6x^2 + 10xy + 6y^2\right), \frac{x^2}{y^3} (4x - 5y) \right).$$

According to Lemma 2.3.3, the map

$$d\Upsilon : (-\infty, 1] \times [1, \infty) \subset \mathbb{R}^2 \to \mathbb{R}^2$$

is isotone on the domain where $d^2\Upsilon = d(d\Upsilon)$ is positive, that is, where the Hessian of Υ,

$$\begin{pmatrix} -\frac{2}{y^2} (6x - 5y) & 2\frac{x}{y^3} (6x - 5y) \\ 2\frac{x}{y^3} (6x - 5y) & -2\frac{x^2}{y^4} (6x - 5y) \end{pmatrix},$$

has nonnegative entries only. Therefore $d\Upsilon$ is isotone on $(-\infty, 1] \times [1, \infty)$.

2.3.2 The majorization relation on ordered Banach spaces

In this subsection we discuss the concept of majorization in the framework of ordered Banach spaces. Since in an ordered Banach space not every string of elements admits a decreasing rearrangement, in this section will concentrate on the case of pairs of discrete probability measures at least one of which is supported by a monotone string of points. The case where the support of the left measure consists of a decreasing string is defined as follows.

Definition 2.3.2. *Suppose that $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}$ and $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ are two discrete Borel probability measures that act on the ordered Banach space E and $m\in(0,1]$ is a parameter. We say that $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}$ is weakly $mL^{\downarrow}$-majorized by $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ (denoted $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{wmL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$) if the left hand measure is supported by a decreasing string of points*

$$\mathbf{x}_1\geq\cdots\geq\mathbf{x}_N \tag{2.3.3}$$

and

$$\sum_{k=1}^{n}\lambda_k\mathbf{x}_k\leq\sum_{k=1}^{n}\lambda_k m\mathbf{y}_k\quad \textit{for all } n\in\{1,\ldots,N\}. \tag{2.3.4}$$

We say that $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}$ is $mL^{\downarrow}$-majorized by $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ (denoted $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{mL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$) if in addition

$$\sum_{k=1}^{N}\lambda_k\mathbf{x}_k=\sum_{k=1}^{N}\lambda_k m\mathbf{y}_k. \tag{2.3.5}$$

Notice that the context of Definition 2.3.2 makes it necessary that all the weights $\lambda_1,\ldots,\lambda_N$ belong to $(0,1]$ and $\sum_{k=1}^{N}\lambda_k=1$.

The three conditions (2.3.3), (2.3.4) and (2.3.5) imply $m\mathbf{y}_N\leq\mathbf{x}_N\leq\mathbf{x}_1\leq m\mathbf{y}_1$ but not the ordering $\mathbf{y}_1\geq\cdots\geq\mathbf{y}_N$. For example, when $N=3$, one may consider the case where

$$m=1,\ \lambda_1=\lambda_2=\lambda_3=1/3,\ \mathbf{x}_1=\mathbf{x}_2=\mathbf{x}_3=\mathbf{x}$$

and

$$\mathbf{y}_1=\mathbf{x},\ \mathbf{y}_2=\mathbf{x}+\mathbf{z},\ \mathbf{y}_3=\mathbf{x}-\mathbf{z},$$

$\mathbf{z}$ being any positive element.

Under these circumstances it is natural to introduce the following companion to Definition 2.3.2, involving the ascending strings of elements as support for the right hand measure.

Definition 2.3.3. *The relation of* weak $mR^{\uparrow}$-majorization*,*

$$\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{wmR^{\uparrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k},$$

between two discrete Borel probability measures means the fulfillment of the condition (2.3.4) *under the presence of the ordering*

$$\mathbf{y}_1\leq\cdots\leq\mathbf{y}_N; \tag{2.3.6}$$

assuming in addition the condition (2.3.5)*, we say that $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}$ is $mR^{\uparrow}$-majorized by $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ (denoted $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{mR^{\uparrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$).*

When every element of E is the difference of two positive elements, the weak majorization relations $\prec_{mL^{\downarrow}}$and $\prec_{mR^{\uparrow}}$ can be augmented so as to obtain majorization relations.

2.3.3 The extension of the Hardy-Littlewood-Polya inequality of majorization

The objective of this subsection is to consider the corresponding extensions of the Hardy-Littlewood-Pólya inequality of majorization for $\prec_{wmL^{\downarrow}}$, $\prec_{mL^{\downarrow}}$, $\prec_{wmR^{\uparrow}}$and $\prec_{mR^{\uparrow}}$. Moreover, we also present also a Sherman type inequality.

The proof of the following theorem is inspired by the techniques succesfully used in [101] and [112].

Theorem 2.3.1. *Suppose that $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}$ and $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ are two discrete probability measures whose supports are included in an open convex subset C of the ordered Banach space E. If $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{mL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$, then*

$$m\sum_{k=1}^{N}\lambda_k\Phi(\mathbf{y}_k)\geq\sum_{k=1}^{N}\lambda_k\Phi(\mathbf{x}_k)+\sum_{k=1}^{N}\lambda_k\omega(\|\mathbf{x}_k-\mathbf{y}_k\|),\tag{2.3.7}$$

for every Gâteaux differentiable ω-m-star-convex function $\Phi:C\to F$ whose differential is isotone and satisfies the hypotheses of Lemma 2.3.2.

The conclusion (2.3.7) *still works under the weaker hypothesis $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{wmL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$, provided that Φ is also an isotone function.*

Proof. According to the gradient inequality (2.3.2), we have

$$\begin{aligned}m\sum_{k=1}^{N}\lambda_k\Phi(\mathbf{y}_k)-\sum_{k=1}^{N}\lambda_k\Phi(\mathbf{x}_k)&=\sum_{k=1}^{N}\lambda_k\left(m\Phi(\mathbf{y}_k)-\Phi(\mathbf{x}_k)\right)\\&\geq\sum_{k=1}^{N}\Phi'(\mathbf{x}_k)(\lambda_k m\mathbf{y}_k-\lambda_k\mathbf{x}_k)+\sum_{k=1}^{N}\lambda_k\omega\left(\|\mathbf{x}_k-\mathbf{y}_k\|\right),\end{aligned}$$

whence, by using Abel's trick of interchanging the order of summation ([114], Theorem 1.9.5, p. 57), one obtains

$$\begin{aligned}&\sum_{k=1}^{N}\lambda_k m\Phi(\mathbf{y}_k)-\sum_{k=1}^{N}\lambda_k\Phi(\mathbf{x}_k)-\sum_{k=1}^{N}\lambda_k\omega\left(\|\mathbf{x}_k-\mathbf{y}_k\|\right)\\&\quad\geq\Phi'(\mathbf{x}_1)(\lambda_1 m\mathbf{y}_1-\lambda_1\mathbf{x}_1)+\sum_{m=2}^{N}\Phi'(\mathbf{x}_m)\Big[\sum_{k=1}^{m}(\lambda_k\mathbf{y}_k-\lambda_k\mathbf{x}_k)-\sum_{k=1}^{m-1}(\lambda_k\mathbf{y}_k-\lambda_k\mathbf{x}_k)\Big]\\&\quad=\sum_{m=1}^{N-1}\Big[(\Phi'(\mathbf{x}_m)-\Phi'(\mathbf{x}_{m+1}))\sum_{k=1}^{m}(\lambda_k m\mathbf{y}_k-\lambda_k\mathbf{x}_k)\Big]+\Phi'(\mathbf{x}_N)\left(\sum_{k=1}^{N}(\lambda_k m\mathbf{y}_k-\lambda_k\mathbf{x}_k)\right).\end{aligned}$$

When $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{mL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$, the last term vanishes and the fact that $D\geq0$ is a consequence of the isotonicity of Φ'. When $\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{x}_k}\prec_{wmL^{\downarrow}}\sum_{k=1}^{N}\lambda_k\delta_{\mathbf{y}_k}$ and Φ is isotone, one applies Lemma 2.3.3 to infer that

$$\Phi'(\mathbf{x}_N)\left(\sum_{k=1}^{N}(\lambda_k m\mathbf{y}_k-\lambda_k\mathbf{x}_k)\right)\geq0.$$

The other cases can be treated in a similar way. □

Remark 12. *Even in the context of usual convex functions, the isotonicity of the differential is not only sufficient but also* necessary *for the validity of Theorem 2.3.1. See [112], Remark 5.*

We leave it to the reader as an exercise to formulate the variant of Theorem 2.3.1 in the case of relations $\prec_{wmR^{\uparrow}}$and $\prec_{mR^{\uparrow}}$.

2.3.4 Further results and open problems

In the following we mention some open problems which might be of interest for further research on.

Notice first that any perturbation of an ω-m-star-convex function Φ satisfying the hypotheses of Theorem 2.3.1 by a bounded function Π verify an inequality of majorization very close to (2.3.7). Precisely, if $|\Pi| \leq \delta$ and $\sum_{k=1}^{N} \lambda_k \delta_{\mathbf{x}_k} \prec_{mL^{\downarrow}} \sum_{k=1}^{N} \lambda_k \delta_{\mathbf{y}_k}$, then $\Psi = \Phi + \Pi$ will verify the relation

$$m \sum_{k=1}^{N} \lambda_k \Phi(\mathbf{y}_k) \geq \sum_{k=1}^{N} \lambda_k \Phi(\mathbf{x}_k) + \sum_{k=1}^{N} \lambda_k \omega(\|\mathbf{x}_k - \mathbf{y}_k\|) - (1+m)\delta.$$

This call the attention to the following class of *approximately* ω-m-star-convex functions:

Definition 2.3.4. *A function $\Phi : C \to \mathbb{R}$ is said to be δ-ω-m-star-convex function if it verifies an estimate of the form*

$$\Phi((1-\lambda)\mathbf{x} + \lambda m \mathbf{y}) \leq (1-\lambda)\Phi(\mathbf{x}) + m\lambda\Phi(\mathbf{y}) - m\lambda(1-\lambda)\omega\left(\|\mathbf{x} - \mathbf{y}\|\right) + \delta,$$

for some $\delta \geq 0$ and all $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in (0,1)$.

The above definition extends (for $\omega = 0$ and $m = 1$) the concept of δ-convex function, first considered by Hyers and Ulam [75] in a paper dedicated to the stability of convex functions. It is natural to rise the problem wheather their result extends to the framework of δ-ω-m-star-convex functions:

Problem. *Suppose that C is a convex subset of $\mathbb{R}^N$. Is that true that every δ-ω-m-star-convex function $\Phi : C \to \mathbb{R}$ can be written as $\Phi = \Psi + \Pi$, where Ψ is an ω-m-star-convex function and Π is a bounded function whose supremum norm is not larger than $k_N\delta$, where the positive constant k_N depends only on the dimension N of the underlying space?*

Of some interest seems to be the concept of local approximate m-star-convexity suggested by [48], Definition 1, which clearly yields new extensions of the majorization inequality:

Definition 2.3.5. *A function $\Phi : C \to \mathbb{R}$ is called locally approximately m-star-convex if for every $x_0 \in C$, and every $\varepsilon > 0$ there exists $\delta > 0$ such that for all x, y in the open ball of center x_0 and radius δ and all $\lambda \in (0,1)$,*

$$\Phi((1-\lambda)x + m\lambda y) \leq (1-\lambda)\Phi(x) + m\lambda\Phi(y) + t(1-t)\left\|x - y\right\|.$$

The whole discussion above can be placed in the more general context of M_p-convexity.

Recall that the weighted M_p-mean is defined for every pair of positive numbers a, b by the formula

$$M_p(a, b; 1-\lambda, \lambda) = \begin{cases} ((1-\lambda)a^p + \lambda b^p)^{1/p}, & \text{if } p \in \mathbb{R}\backslash\{0\} \\ a^{1-\lambda}b^\lambda, & \text{if } p = 0 \\ \max\{a, b\}, & \text{if } p = \infty, \end{cases}$$

where $\lambda \in [0,1]$. If $p > 0$, then it is usual to extend M_p to all pairs of nonnegative numbers.

Definition 2.3.6. *A function $\Phi : C \to \mathbb{R}$ is called ω-m-M_p-star-convex if there exist a number $p \in \mathbb{R}$ and a modulus $\omega : [0, \infty) \to \mathbb{R}$ such that*

$$\Phi\left((1-\lambda)\mathbf{x} + \lambda\mathbf{y}\right) \leq ((1-\lambda)\Phi(\mathbf{x})^p + m\lambda\Phi(\mathbf{y})^p)^{1/p} - m\lambda(1-\lambda)\omega\left(\|\mathbf{x}-\mathbf{y}\|\right),$$

for all $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in (0,1)$.

Reversing the inequality one obtain the concept of ω-m-M_p-star-concave functions.

The usual M_p-convex/M_p-concave functions represent the particular case where $m = 1$ and $\omega = 0$.

It is worth noticing that the M_p-convex (M_p-concave) functions for $p \neq 0$ are precisely the functions Φ such that Φ^p is convex (concave), while the M_0-convex (M_0-concave) functions are nothing but the log-convex (log-concave) functions. Notice also that the M_∞-convex ($M_{-\infty}$-concave) functions are precisely the quasi-convex (quasi-concave) functions.

The next result represents the extension of Lemma 2.3.2 to the case of ω-m-M_p-star-convex functions.

Lemma 2.3.4. *Suppose that C is an open convex subset of the Banach space E and $\Phi : C \to \mathbb{R}_+$ is a function both Gâteaux differentiable and ω-m-M_p-star-convex. If $p \neq 0$, then Φ verifies the inequality*

$$\Phi^p(\mathbf{y}) \geq \Phi^p(\mathbf{x}) + p\Phi(\mathbf{x})^{p-1}d\Phi(\mathbf{x})(\mathbf{y}-\mathbf{x}) + m\omega\left(\|\mathbf{x}-\mathbf{y}\|\right),$$

for all $\mathbf{x}, \mathbf{y} \in C$.

The analogue of this result for $p = 0$ and $\omega = 0$ requires the strict positivity of the function Φ and can be stated as

$$\log\Phi(\mathbf{y}) - \log\Phi(\mathbf{x}) \geq \frac{d\Phi(\mathbf{x})(\mathbf{y}-\mathbf{x})}{\Phi(\mathbf{x})},$$

for all $\mathbf{x}, \mathbf{y} \in C$. The last two inequalities work in the reverse direction in the case of ω-m-M_p-star-concave functions.

While it is clear that Lemma 2.3.4 allows us to prove Hardy-Littlewood-Pólya type inequalities more general than those provided by Theorem 2.3.1, the exploration of the world of ω-m-M_p-star-convex/concave functions for $\omega \neq 0$ and $m \in (0,1)$ is just at the beginning.

2.4 Convex type inequalities with nonpositive weights

In this section we extend Jensen-Steffensen's inequality via majorization arguments, into the framework of $\mathbb{R}^n$, for any $n \geq 2$. Generally speaking, our aim is to prove convex type inequalities relaxing the weights which are allowed to be nonpositive. We are dealing with monotonic increasing or decreasing sequences with respect to majorization relation in $\mathbb{R}^n$ and the well known behaviour under convex functions invariant on permutation of variables. More precisely, Jensen-Steffensen's and Sherman's type inequalities are obtained, even in the context of strongly convex functions. Moreover, applications concerning relative convexity aspects and extensions on spaces with curved geometry could be also derived.

In the twentieth century, an intense research activity and many significant results were obtained in geometric functional analysis, mathematical economics, convex analysis, and nonlinear optimization. The classical books [68, 104] played a prominent role related to the subject of convex functions, in which one of the most relevant topic is devoted to the concept of majorization.

For any two vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ let us consider $\mathbf{u}^{\downarrow}$ and $\mathbf{v}^{\downarrow}$ two vectors with the same entries as $\mathbf{u}$ and $\mathbf{v}$, expressed in decreasing order, as

$$u_1^{\downarrow} \geq ... \geq u_n^{\downarrow},\ v_1^{\downarrow} \geq ... \geq v_n^{\downarrow}.$$

We say that, the vector $\mathbf{u}$ is *majorized* by $\mathbf{v}$ (abbreviated, $\mathbf{u} \prec \mathbf{v}$) if

$$\begin{aligned} &\sum_{i=1}^{k} u_i^{\downarrow} \leq \sum_{i=1}^{k} v_i^{\downarrow} \qquad (1 \leq k \leq n-1), \\ &\sum_{i=1}^{n} u_i^{\downarrow} = \sum_{i=1}^{n} v_i^{\downarrow}. \end{aligned} \tag{2.4.1}$$

For other relevant details and various applications concerning the majorization theory we refer to [104]. In this context, the monotonicity with respect to the majorization order is called Schur-convex property and has been introduced by I. Schur in 1923. It is well known that

$$\mathbf{u} \prec \mathbf{v} \quad \text{iff} \quad \mathbf{u} = \mathbf{v}\mathbf{A},$$

for some doubly stochastic matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{n,n}(\mathbb{R})$, i.e. a matrix with nonnegative entries and rows and columns sums equal to 1.

The concept of majorization is a powerful topic of research with relevant results in different areas. In this regard, we just enumerate few of them: a necessary and sufficient condition for a linear map to preserve group majorizations can be found in [131]; new majorization results are studied in [83, 132]; interesting properties on superquadratic functions related to Jensen–Steffensen's inequality are obtained in [1]. All these ideas are also based on the theory of uniformly convex functions, which in addition gives the possibility to define the concept of majorization into the spaces of curved geometry (see [126]). For other results see [114, 124, 131, 133, 134, 135].

The weighted concept of majorization between two vectors $\mathbf{u} = (u_1, \ldots, u_l) \in I^l$, $\mathbf{v} = (v_1, \ldots, v_m) \in I^m$ with nonnegative weights $\mathbf{a} = (a_1, \ldots, a_l) \in [0, \infty)^l$ and $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$, where I is an interval in $\mathbb{R}$ and $m, l \geq 2$, has been defined in S. Sherman [160]. The concept of weighted majorization is defined by assuming the existence of a columns stochastic matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$, i.e. a matrix with nonnegative entries and columns sums equal to 1, such that

$$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \ldots, m), \tag{2.4.2}$$

$$u_i = \sum_{j=1}^{m} v_j \alpha_{ji}, \quad (i = 1, \ldots, l). \tag{2.4.3}$$

Under conditions (2.4.2) – (2.4.3) it is proved that, the following inequality

$$\sum_{i=1}^{l} a_i f(u_i) \leq \sum_{j=1}^{m} b_j f(v_j)$$

holds for every convex function $f : I \to \mathbb{R}$. See [160]. We can write conditions (2.4.2) – (2.4.3) in the matrix form

$$\mathbf{b} = \mathbf{a}\mathbf{A}^{\mathbf{T}} \quad \textit{and} \quad \mathbf{u} = \mathbf{v}\mathbf{A}.$$

In the rest of the section we write

$$(\mathbf{u}, \mathbf{a}) \prec (\mathbf{v}, \mathbf{b})$$

and say that a pair $(\mathbf{u}, \mathbf{a})$ is weighted majorized by $(\mathbf{v}, \mathbf{b})$ if (2.4.2) – (2.4.3) are satisfied for some columns stochastic matrix $\mathbf{A}$. Note that, in the case $l = 1$ and $\mathbf{b} = [1]$ we deduce Jensen's inequality. When $m = l$ and all weights a_i and b_j are equal to $1/m$, the condition (2.4.2) assures the *stochasticity on rows,* so in that case we deal with doubly stochastic matrices.

Since all these above inequalities are dealing with positive weights the study of the case of nonpositive weights is very challenging. In this context we recall one of the first relevant step, the so called Jensen Steffensen inequality. We refer to [115] for the following result.

Theorem 2.4.1. *Let $x_n \leq x_{n-1} \leq \cdots \leq x_1$ be points in $[a, b]$ and let $p_1, \ldots, p_n$ be real numbers such that the partial sums $S_k = \sum_{i=1}^{k} p_i$ verify the relations*

$$0 \leq S_k \leq S_n \quad \textit{and} \quad S_n > 0.$$

Then for every convex functions $f : [a, b] \to \mathbb{R}$ we have the inequality

$$f\left(\frac{1}{S_n} \sum_{k=1}^{n} p_k x_k\right) \leq \frac{1}{S_n} \sum_{k=1}^{n} p_k f(x_k).$$

The aim of this section is to present new extensions of the above inequality for the case of finite dimensional spaces. More precisely, our first aim is to extend Theorem 2.4.1 in the framework of $\mathbb{R}^n$ and then to derive Sherman and Jensen Steffensen's type inequalities for perturbed convex functions with complete homogeneous symmetric polynomials. We are very confident that our strategy ca be also adapted to more general spaces, not only in $\mathbb{R}^n$, but also in spaces with curved geometry.

The structure of the section is at follows: in Subsection 2.4.1 we briefly present and introduction with the motivation and some preliminaries concerning historical aspects of the main problem we study; Subsection 2.4.2 is devoted to Jensen Steffensen's inequalities in the framework of $\mathbb{R}^n$ via majorization concept; in Subsection 2.4.3 we present some application related to Sherman's inequalities when the weights can be chosen to be nonpositive and also to Jensen Steffensen's type inequalities for perturbed convex functions with complete homogeneous symmetric polynomials; Subsection 2.4.4 presents some conclusions and further applications related to relative convexity aspects and the possibility to transfer similar results into the spaces with curved geometry.

2.4.1 An extension of Jensen-Steffensen's inequality context in $\mathbb{R}^n$ via majorization ordering

In this subsection we introduce majorization concept in order to present a general strategy which allow us to extend Jensen-Steffensen inequality from $\mathbb{R}$ to multidimensional space $\mathbb{R}^n$. More precisely, we prove our first result, given by an extension of the Jensen-Steffensen inequality.

The following lemma is used to prove the main result of this section.

Lemma 2.4.1. *For any U_1, U_2, $\cdots$, U_m doubly stochastic matrices in $\mathcal{M}_{n,n}(\mathbb{R})$ we have*

$$U_1\mathbf{x}_1 + U_2\mathbf{x}_2 + \cdots + U_m\mathbf{x}_m \prec \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_m,$$

for any $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^n$.

Proof. If we denote by $\mathbf{u}_1 = U_1\mathbf{x}_1, \ldots, \mathbf{u}_m = U_m\mathbf{x}_m$ we need to prove that

$$\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_m \prec \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_m.$$

Since $\mathbf{u}_i \prec \mathbf{x}_i$, using (2.4.1), for any $i = 1, \ldots, m$, and summing up all the inequalities we get the conclusion.

□

In the following sentences we recall some basic facts relevant in our context.

Remark 13. *Note that, every convex function defined on $\mathbb{R}^n$ admits one sided directional derivatives at any point and, moreover, $\partial f(a)$ is singleton precisely when f has directional derivative $f'(a;v)$ and, in that case we have that $\partial f(a)$ consists of the mapping $v \to f'(a;v)$.*

Based on [115, Remark 3.6.1.], we have that

$$f'_+(a;v) \geq \langle d, v\rangle \geq f'_-(a;v) \qquad (a, v \in \mathbb{R}^n, d \in \partial f(a)),$$

where

$$f'_\pm(a;v) = \lim_{t\to 0_+} \frac{f(a+tv) - f(a)}{t} \qquad (a, v \in \mathbb{R}^n, d \in \partial f(a)),$$

and z belongs to the subdifferential of f at the point a, namely $\partial f(a)$, means that

$$f(x) \geq f(a) + \langle x - a, z\rangle \qquad (x \in \mathbb{R}^n).$$

Hence, taking into account the above relations we get

$$f(\mathbf{z}) \geq f(\mathbf{y}) + \langle \mathbf{d}, \mathbf{z} - \mathbf{y}\rangle \qquad (\mathbf{x} \prec \mathbf{y} \prec \mathbf{z},\ \mathbf{d} \in \partial f(\mathbf{x})), \tag{2.4.4}$$

$$f(\mathbf{z}) \leq f(\mathbf{y}) + \langle \mathbf{d}, \mathbf{z} - \mathbf{y} \rangle \qquad (\mathbf{y} \prec \mathbf{z} \prec \mathbf{x},\ \mathbf{d} \in \partial f(\mathbf{x}))\,. \tag{2.4.5}$$

Moreover, using the linearity of the scalar product we can also have

$$\langle \mathbf{v}, \mathbf{z} - \mathbf{y} \rangle \geq \langle \mathbf{u}, \mathbf{z} - \mathbf{y} \rangle \geq 0 \qquad (\mathbf{a} \prec \mathbf{b},\ \mathbf{y} \prec \mathbf{z},\ \mathbf{u} \in \partial f(\mathbf{a}),\ \mathbf{v} \in \partial f(\mathbf{b}))\,.$$

Inspired from [115] we shall use the following notation related to $\mathbf{z}_1, \ldots, \mathbf{z}_m \in \mathbb{R}^n$ and $p_1, \ldots, p_m \in \mathbb{R}$:

$$\begin{aligned}
\bar{\mathbf{z}} &= p_1\mathbf{z}_1 + \cdots + p_m\mathbf{z}_m, \\
P_k &= p_1 + \cdots + p_k \qquad (k \in \{1, 2, \ldots, m\})\,, \\
\bar{P}_k &= p_k + \cdots + p_m \qquad (k \in \{1, 2, \ldots, m\})\,.
\end{aligned}$$

Definition 2.4.1. *We say that a sequence $\mathbf{z}_1, \ldots, \mathbf{z}_m \in \mathbb{R}^n$ is monotonic decreasing with respect to majorization relation iff the following relations hold*

$$\mathbf{z}_m \prec \mathbf{z}_{m-1} \prec \cdots \prec \mathbf{z}_2 \prec \mathbf{z}_1. \tag{2.4.6}$$

We are now in position to present the extension of Jensen-Steffensen's type inequality in $\mathbb{R}^n$.

Theorem 2.4.2. *Let I be an interval in $\mathbb{R}$ and m, $n \geq 1$. If $f : I^n \to \mathbb{R}$ is a convex function invariant under permutation of coordinates, then for every $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$, which is monotonic decreasing with respect to majorization relation, and every real m-tuple $\mathbf{p} = (p_1, \ldots, p_m)$ such that, for every $i \in \{1, 2, \ldots, m\}$ we have*

$$0 \leq P_i \leq P_m = 1,$$

then the following inequality holds

$$f\left(\sum_{i=1}^{m} p_i\mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f\left(\mathbf{z}_i\right).$$

Proof. From (2.4.6) we infer the existence of A_1, A_2, $\cdots$, A_{m-1} doubly stochastic matrices in $\mathcal{M}_{n,n}(\mathbb{R})$ such that

$$\begin{aligned}
\mathbf{z}_2 &= A_1\mathbf{z}_1, \\
\mathbf{z}_3 &= A_2\mathbf{z}_2, \\
&\vdots \\
\mathbf{z}_m &= A_{m-1}\mathbf{z}_{m-1}.
\end{aligned}$$

Hence, we have that

$$\bar{\mathbf{z}} = p_1\mathbf{z}_1 + \cdots + p_m\mathbf{z}_m$$

$$= \left(p_1 I_n + p_2 A_1 + \cdots + p_{n-1}A_{n-2}\cdots A_1 + p_m A_{m-1}\cdots A_1\right)\mathbf{z}_1,$$

hence, we deduce that $\bar{z} \prec \mathbf{z}_1$, based on the fact that the matrix $p_1 I_n + p_2 A_1 + \cdots + p_{n-1}A_{n-2}\cdots A_1 + p_m A_{m-1}\cdots A_1$ is doubly stochastic.

On the other hand, we have that

$$\mathbf{z}_m = p_m\mathbf{z}_m + \left(p_1 + \cdots + p_{m-1}\right)\mathbf{z}_m = p_m\mathbf{z}_m + p_{m-1}A_{m-1}\mathbf{z}_{m-1} + \left(p_1 + \cdots + p_{m-2}\right)\mathbf{z}_m = \ldots$$

$$= p_m \mathbf{z}_m + p_{m-1} A_{m-1} \mathbf{z}_{m-1} + \cdots + p_1 A_{m-1} \cdots A_1 \mathbf{z}_1.$$

Hence, we have that

$$p_m \mathbf{z}_m + p_{m-1} A_{m-1} \mathbf{z}_{m-1} + \cdots + p_1 A_{m-1} \cdots A_1 \mathbf{z}_1 \prec p_1 \mathbf{z}_1 + p_2 \mathbf{z}_2 + \cdots + p_m \mathbf{z}_m.$$

More precisely, if we choose in Lemma 2.4.1, $\mathbf{x}_1 = p_1 \mathbf{z}_1$, $\mathbf{x}_2 = p_2 \mathbf{z}_2, \ldots, \mathbf{x}_m = p_m \mathbf{z}_m$, $U_m = I_m$, $U_{m-1} = A_{m-1}$, $\ldots, U_1 = A_{m-1} \cdots A_1$, we get $\mathbf{z}_m \prec \bar{\mathbf{z}}$.

Thus, we just have proved that

$$\mathbf{z}_m \prec \bar{\mathbf{z}} \prec \mathbf{z}_1.$$

Inspired from [115, Theorem 1.5.6.], f is convex and invariant under permutation of coordinates, then by using [126, Theorem 5] we have that

$$f(\mathbf{z}_m) \leq f(\bar{\mathbf{z}}) \leq f(\mathbf{z}_1),$$

hence, we infer the existence of an index l such that

$$f(\mathbf{z}_m) \leq \cdots \leq f(\mathbf{z}_{l+1}) \leq f(\bar{\mathbf{z}}) \leq f(\mathbf{z}_l) \leq \cdots \leq f(\mathbf{z}_1). \tag{2.4.7}$$

Following the idea in the proof of [115, Theorem 1.5.6.], for any $\mathbf{d} \in \partial f(\bar{\mathbf{z}})$, using (2.4.9), we have

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) - \sum_{i=1}^{m} p_i f(\mathbf{z}_i)$$

$$\leq \sum_{i=1}^{l-1} S_i \left(\langle d, \mathbf{z}_i - \mathbf{z}_{i+1} \rangle - f(\mathbf{z}_i) + f(\mathbf{z}_{i+1})\right)$$

$$+ S_l \left(\langle \mathbf{d}, \mathbf{z}_l - \bar{\mathbf{z}} \rangle - f(\mathbf{z}_l) + f(\bar{\mathbf{z}})\right)$$

$$+ \bar{S}_{l+1} \left(f(\bar{\mathbf{z}}) - f(\mathbf{z}_{l+1}) - \langle \mathbf{d}, \bar{\mathbf{z}} - \mathbf{z}_{l+1} \rangle\right)$$
$$+ \sum_{i=l+1}^{m-1} \bar{S}_{i+1} \left(f(\mathbf{z}_i) - f(\mathbf{z}_{i+1}) - \langle \mathbf{d}, \mathbf{z}_i - \mathbf{z}_{i+1} \rangle\right),$$

where $\bar{S}_i = S_m - S_i$.

Using (2.4.4)-(2.4.5) from Remark 13 we get that all the above terms are nonpositive real numbers and the conclusion follows easily.

$\square$

For the convenience of the reader we also present some details for the case of increasing sequences with respect to the majorization relation.

Theorem 2.4.3. *If $f : I^n \to \mathbb{R}$ is a convex function invariant under permutation of coordinates, then for every $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$, which is monotonic increasing with respect to majorization relation, and every real m-tuple $\mathbf{p} = (p_1, \ldots, p_m)$ such that, for every $i \in \{1, 2, \ldots, m\}$ we have*

$$0 \leq P_i \leq P_m = 1,$$

then the following inequality holds

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f(\mathbf{z}_i).$$

Proof. From (2.4.6) we infer the existence of A_1, A_2, $\cdots$, A_{m-1} doubly stochastic matrices in $\mathcal{M}_{n,n}(\mathbb{R})$ such that

$$\begin{aligned}
\mathbf{z}_{m-1} &= A_m \mathbf{z}_m, \\
\mathbf{z}_{m-2} &= A_{m-1} \mathbf{z}_{m-1}, \\
&\vdots \\
\mathbf{z}_2 &= A_3 \mathbf{z}_3, \\
\mathbf{z}_1 &= A_2 \mathbf{z}_2.
\end{aligned}$$

Hence, we have that

$$\begin{aligned}
\bar{\mathbf{z}} &= p_1 \mathbf{z}_1 + \cdots + p_m \mathbf{z}_m \qquad (2.4.8) \\
&= (p_1 A_2 \cdots A_m + p_2 A_3 \cdots A_m + \ldots p_{m-1} A_m + p_m I_n) \mathbf{z}_m,
\end{aligned}$$

hence, we deduce that $\bar{z} \prec \mathbf{z}_m$, based on the fact that the matrix $p_1 A_2 \cdots A_m + p_2 A_3 \cdots A_m + \ldots p_{m-1} A_m + p_m I_n$ is doubly stochastic.

On the other hand, we have that

$$\begin{aligned}
\mathbf{z}_1 = p_1 \mathbf{z}_1 + (p_2 + \cdots + p_m) \mathbf{z}_1 &= p_1 \mathbf{z}_1 + p_2 A_2 \mathbf{z}_2 + (p_3 + \cdots + p_m) \mathbf{z}_1 \\
&= \cdots = p_1 \mathbf{z}_1 + p_2 A_2 \mathbf{z}_2 + \cdots + p_m A_2 \cdots A_m \mathbf{z}_m,
\end{aligned}$$

It follows that

$$p_1 \mathbf{z}_1 + p_2 A_2 \mathbf{z}_2 + \cdots + p_m A_2 \cdots A_m \mathbf{z}_m \prec p_1 \mathbf{z}_1 + p_2 \mathbf{z}_2 + \cdots + p_m \mathbf{z}_m,$$

where we have used Lemma 2.4.1, for $\mathbf{x}_1 = p_1 \mathbf{z}_1$, $\mathbf{x}_2 = p_2 \mathbf{z}_2, \ldots, \mathbf{x}_m = p_m \mathbf{z}_m$, $U_1 = I_m$, $U_2 = A_2$, $\ldots, U_m = A_2 \cdots A_m$. Hence, we get $\mathbf{z}_1 \prec \bar{\mathbf{z}}$. Thus, we just have proved that

$$\mathbf{z}_1 \prec \bar{\mathbf{z}} \prec \mathbf{z}_m.$$

Inspired from [115, Theorem 1.5.6.], f is convex and invariant under permutation of coordinates, then by using [126] we have

$$f(\mathbf{z}_1) \leq f(\bar{\mathbf{z}}) \leq f(\mathbf{z}_m),$$

we infer the existence of an index l such that

$$f(\mathbf{z}_1) \leq \cdots \leq f(\mathbf{z}_l) \leq f(\bar{\mathbf{z}}) \leq f(\mathbf{z}_{l+1}) \leq \cdots \leq f(\mathbf{z}_m). \qquad (2.4.9)$$

Now, using the similar argument as in the proof of Theorem 2.4.2 the same conclusion holds.

□

Example 2. *Let us consider the following vectors in $\mathbb{R}^n$ which verify*

$$\mathbf{z}_1 \prec \mathbf{z}_2 \prec \cdots \prec \mathbf{z}_{n-1} \prec \mathbf{z}_n,$$

where

$$\mathbf{z}_1 = \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}, \frac{1}{n}\right),$$
$$\mathbf{z}_2 = \left(0, \frac{1}{n}, \dots, \frac{1}{n}, \frac{2}{n}\right),$$
$$\dots$$
$$\mathbf{z}_{n-1} = \left(0, 0, \dots, \frac{1}{n}, \frac{n-1}{n}\right),$$
$$\mathbf{z}_n = (0, 0, \dots, 0, 1).$$

By choosing the weights

$$(p_1, p_2, \dots p_n) = \left(\frac{n-1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n}, \frac{n-1}{n}\right),$$

which can be nonpositive and verifies the hypotheses in Theorem 2.4.2. Hence, for any convex function invariant under permutation of coordinates $f : I^n \to \mathbb{R}$, we have

$$f\left(\frac{n-1}{n^2}, \frac{n-2}{n^2}, \frac{n-3}{n^2} \dots, \frac{1}{n^2}, \frac{-3n^2+7n-2}{n^2}\right)$$
$$\leq \frac{n-1}{n}\left(f\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}, \frac{1}{n}\right) + f(0, 0, \dots, 0, 1)\right) - \frac{1}{n}\sum_{i=2}^{n-1} f(\mathbf{z}_i).$$

2.4.2 Sherman's type inequalities with nonpositive weights

In this subsection we develop the result from the previous section for the case of nonpositive weights, in different situations. The first step is introduce the weighted concept of majorization between two n-tuples $\mathbf{x} = (x_1, ..., x_l)$, $\mathbf{y} = (y_1, \dots, y_m)$, where $\mathbf{z}_1, \dots, \mathbf{z}_l \in I^n$, $\mathbf{y}_1, \dots, \mathbf{y}_m \in I^n$, with real weights $\mathbf{a} = (a_1, \dots, a_l) \in \mathbb{R}^l$ (which can be nonpositive) and $\mathbf{b} = (b_1, \dots, b_m) \in [0, \infty)^m$, where I is an interval in $\mathbb{R}$ and $m, l \geq 2$.

We define the concept of weighted majorization $(\mathbf{x}, \mathbf{a}) \prec (\mathbf{y}, \mathbf{b})$ by considering any matrix $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{ml}(\mathbb{R})$, verifying

$$0 \leq A_k^i \leq A_k^m = 1, \qquad (1 \leq k, i \leq m) \tag{2.4.10}$$

where

$$A_k^i = \alpha_{1i} + \cdots + \alpha_{ki} \qquad (k \in \{1, 2, \dots, m\}) \qquad (1 \leq k \leq m), \tag{2.4.11}$$

such that

$$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \dots, m), \tag{2.4.12}$$

$$\mathbf{x}_i = \sum_{j=1}^{m} \mathbf{y}_j \alpha_{ji}, \quad (i = 1, \dots, l). \tag{2.4.13}$$

We can present now the extension of Sherman's inequality in $\mathbb{R}^n$, when the weights are allowed to be nonpositive.

Theorem 2.4.4. *If*

$$\mathbf{x}_m \prec \mathbf{x}_{m-1} \prec \cdots \prec \mathbf{x}_2 \prec \mathbf{x}_1. \tag{2.4.14}$$

and let us suppose that conditions (2.4.10)-(2.4.13) *are satisfied. Then, the following inequality*

$$\sum_{i=1}^{l} a_i f(\mathbf{x}_i) \leq \sum_{j=1}^{m} b_j f(\mathbf{y}_j)$$

holds for every convex function $f : I^n \to \mathbb{R}$ which is invariant under permutation of coordinates.

Proof. As in the proof of Theorem 2.4.2 we can deduce that

$$\mathbf{x}_m \prec \bar{\mathbf{y}}_j \prec \mathbf{x}_1 \qquad (j = 1, \ldots, m),$$

which means that we can use Jensen-Steffensen's inequality and we obtain

$$f\left(\sum_{j=1}^{m} \alpha_{ji}\mathbf{y}_j\right) \leq \sum_{j=1}^{m} \alpha_{ji} f(\mathbf{y}_j) \qquad (i = 1, \ldots, l).$$

Taking into account (2.4.2) – (2.4.3) and applying Theorem 2.4.2 for each $\mathbf{z}_i$, $i = 1, ..., l$, where $z_i = \sum_{j=1}^{m} p_j \mathbf{y}_j$, $p_j = \alpha_{ji}$, we get

$$\sum_{i=1}^{l} a_i f(\mathbf{x}_i) = \sum_{i=1}^{l} a_i f\left(\sum_{j=1}^{m} \mathbf{y}_j \alpha_{ji}\right)$$

$$\leq \sum_{i=1}^{l} a_i \left(\sum_{j=1}^{m} \alpha_{ji} f(\mathbf{y}_j)\right) = \sum_{j=1}^{m} f(\mathbf{y}_j) \sum_{i=1}^{l} a_i \alpha_{ji}.$$

Consequently, since $b_j = \sum_{i=1}^{l} a_i \alpha_{ji}$ we have

$$\sum_{i=1}^{l} a_i f(\mathbf{x}_i) \leq \sum_{j=1}^{m} b_j f(\mathbf{y}_j).$$

□

2.4.3 The case of h_2 strongly convex functions

The second topic we address in this section is related to the study of a perturbed family of convex functions by complete homogeneous symmetric polynomials with even degree, which are positive.

Inspired from the strategy used in [1, 2, 3, 30, 83, 15] we have the following result.

Theorem 2.4.5. *(Jensen-Steffensen's type inequality) Let $C > 0$ and let I be an interval in $\mathbb{R}$. If $f : I^n \to \mathbb{R}$ is h_2 strongly convex with modulus C and invariant under permutation of coordinates, then* *for every monotonic sequence* $\mathbf{z}_1, \ldots, \mathbf{z}_m \in I^n$*, as in* (2.4.6)*, and every real n-tuple* $\mathbf{p} = (p_1, \ldots, p_m)$ *such that, for every* $i \in \{1, 2, \ldots, m\}$*,* $0 \leq P_i \leq P_m = 1$ *the following inequality holds:*

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f\left(\mathbf{z}_i\right) - C \sum_{i=1}^{m} p_i h_2\left(\mathbf{z}_i - \bar{\mathbf{z}}\right),$$

where $\bar{\mathbf{z}}$ *is defined in* (2.4.8)*.*

Proof. As in the proof of Theorem 2.4.2 we can obtain that

$$\mathbf{z}_m \prec \bar{\mathbf{z}} \prec \mathbf{z}_1,$$

which means that $g : I^n \to \mathbb{R}$, where $g\left(\bar{\mathbf{z}}\right) = g\left(\sum_{i=1}^{n} p_i \mathbf{z}_i\right)$ is well defined.

Using the convexity of the function $g(\cdot) = f(\cdot) - ch_2(\cdot)$, as in Definition of h_2 strongly convex function, and applying Jensen-Steffensen's inequality, we obtain

$$g\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i g(\mathbf{z}_i).$$

Going back to f, we get

$$\begin{aligned} f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) - Ch_2\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) &\leq \sum_{i=1}^{m} p_i (f(\mathbf{z}_i) - Ch_2(\mathbf{z}_i)) \\ &= \sum_{i=1}^{m} p_i f(\mathbf{z}_i) - C \sum_{i=1}^{m} p_i h_2(\mathbf{z}_i), \end{aligned}$$

or written differently

$$\begin{aligned} f\left(\sum_{i=1}^{n} p_i \mathbf{z}_i\right) &\leq \sum_{i=1}^{m} p_i f(\mathbf{z}_i) - C\left[\sum_{i=1}^{m} p_i h_2(\mathbf{z}_i) - h_2\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right)\right] \\ &= \sum_{i=1}^{m} p_i f\left(\mathbf{z}_i\right) - C \sum_{i=1}^{m} \frac{p_i}{2} \left(\left(\sum_{k=1}^{n} z_i^k\right)^2 + \sum_{k=1}^{n} \left(z_i^k\right)^2\right) \\ &+ \frac{C}{2} \left(\left(\sum_{i=1}^{n} \sum_{k=1}^{m} p_k z_k^i\right)^2 + \sum_{i=1}^{n} \left(\sum_{k=1}^{m} p_k z_k^i\right)^2\right) \\ &= \sum_{i=1}^{m} p_i f\left(\mathbf{z}_i\right) - \frac{C}{2} \left(\sum_{i=1}^{m} p_i \left(\sum_{k=1}^{n} z_i^k\right)^2 - \left(\sum_{i=1}^{m} p_i \sum_{k=1}^{n} z_i^k\right)^2\right) \\ &- \frac{C}{2} \left(\sum_{i=1}^{n} \left(\sum_{k=1}^{m} p_k \left(z_k^i\right)^2\right) - \sum_{i=1}^{n} \left(\sum_{k=1}^{m} p_k z_k^i\right)^2\right). \end{aligned}$$

Now, inspired from [15, Theorem 2] we can treat the above last two terms as follows: for any

$$u_i = \sum_{k=1}^{n} z_i^k$$

and

$$\bar{u} = \sum_{i=1}^{m} p_i u_i,$$

we have the following identity

$$\sum_{i=1}^{m} p_i (u_i)^2 - (\bar{u})^2 = \sum_{i=1}^{m} p_i (u_i - \bar{u})^2,$$

hence, it follows that

$$f\left(\sum_{i=1}^{m} p_i \mathbf{z}_i\right) \leq \sum_{i=1}^{m} p_i f(\mathbf{z}_i) - C \sum_{i=1}^{m} p_i h_2(\mathbf{z}_i - \bar{\mathbf{z}}).$$

□

Using our extension of Sherman's results (for nonpositive weights) we can deduce Sherman's inequality for h_2 strongly convex functions with modulus C.

Theorem 2.4.6. *(Sherman's type inequality) Let $C > 0$ and let I be an interval in $\mathbb{R}$. Let $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_l)$, $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$, where $\mathbf{z}_1, \ldots, \mathbf{z}_l \in I^n$, $\mathbf{y}_1, \ldots, \mathbf{y}_m \in I^n$ and let $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{R}^l$ and $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$ be such that $(\mathbf{y}, \mathbf{b}) \prec (\mathbf{z}, \mathbf{a})$. If in addition we assume that*

$$\mathbf{z}_m \prec \mathbf{z}_{m-1} \prec \cdots \prec \mathbf{z}_2 \prec \mathbf{z}_1. \tag{2.4.15}$$

then for every $f : I^n \to \mathbb{R}$ h_2 strongly convex with modulus C and invariant under permutation of coordinates we have

$$\sum_{i=1}^{l} b_i f(\mathbf{y}_i) \leq \sum_{j=1}^{m} a_j f(\mathbf{z}_j) - C \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} h_2(\mathbf{z}_j - \mathbf{y}_i).$$

Proof. From (2.4.2) – (2.4.3) and using Theorem 2.4.5 for each $\mathbf{y}_i$, $i = 1, \ldots, l$, where $\mathbf{y}_i = \sum_{j=1}^{l} p_j \mathbf{z}_j$, $p_j = \alpha_{ji}$, we have

$$\begin{aligned}
\sum_{i=1}^{l} b_i f(\mathbf{y}_i) &= \sum_{i=1}^{l} b_i f\left(\sum_{j=1}^{m} \mathbf{z}_j \alpha_{ji}\right) \\
&\leq \sum_{i=1}^{l} b_i \left(\sum_{j=1}^{m} \alpha_{ji} f(\mathbf{z}_j) - C \sum_{j=1}^{m} \alpha_{ji} h_2(\mathbf{z}_j - \mathbf{y}_i)\right) \\
&= \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} f(\mathbf{z}_j) - C \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} h_2(\mathbf{z}_j - \mathbf{y}_i) \\
&= \sum_{j=1}^{m} f(\mathbf{z}_j) \sum_{i=1}^{l} b_i \alpha_{ji} - C \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} h_2(\mathbf{z}_j - \mathbf{y}_i).
\end{aligned}$$

Hence, using that $a_j = \sum_{i=1}^{l} b_i \alpha_{ji}$ we get

$$\sum_{i=1}^{l} b_i f(\mathbf{y}_i) \leq \sum_{j=1}^{m} a_j f(\mathbf{z}_j) - C \sum_{i=1}^{l} b_i \sum_{j=1}^{m} \alpha_{ji} h_2 (\mathbf{z}_j - \mathbf{y}_i).$$

□

2.4.4 Conclusions and applications

The point of departure in this subsection is given by the possibility to define the weighted concept of majorization within a class of spaces with curved geometry (in which compare the the length of a median of a triangle to the lengths of its sides).

Using a similar strategy as in the previous sections our future aim will be to prove that convex type inequalities hold even in global NPC spaces, for some nonpositive weights. This could pe done taking into account the following remarks and properties of this spaces.

Definition 2.4.2. *A global NPC space is a complete metric space $M = (M, d)$ for which the following inequality holds true: for every pair of points $x_0, x_1 \in M$ there exists a point $y \in M$ such that for all points $z \in M$,*

$$d^2(z, y) \leq \frac{1}{2} d^2(z, x_0) + \frac{1}{2} d^2(z, x_1) - \frac{1}{4} d^2(x_0, x_1). \tag{2.4.16}$$

Here "NPC" stands for "nonpositive curvature". Global NPC spaces are also known as CAT(0) spaces or Hadamard spaces. For more details, the interested reader may consult the excellent survey of Sturm [163] (and also the books of Ballman [17], Bridson and Haefliger [32], and Jost [80]).

Not that in a global NPC space, each pair of points $x_0, x_1 \in M$ can be connected by a geodesic (that is, by a rectifiable curve $\gamma : [0,1] \to M$ such that the length of $\gamma|_{[s,t]}$ is $d(\gamma(s), \gamma(t))$ for all $0 \leq s \leq t \leq 1$). Moreover, this geodesic is unique.

The point y that appears in Definition 2.4.1 is the *midpoint* of x_0 and x_1 and has the property

$$d(x_0, y) = d(y, x_1) = \frac{1}{2} d(x_0, x_1).$$

An important role here is played by the inequality (2.4.16), which assures the uniform convexity of the square distance. See Bhatia [25]. Every Hilbert space is a global NPC space. Its geodesics are the line segments and $y = \frac{x_0 + x_1}{2}$. In general, a Riemannian manifold is a global NPC space if and only if it is complete, simply connected and of nonpositive sectional curvature. Besides manifolds, other important examples of global NPC spaces are the Bruhat-Tits buildings (in particular, the trees). See [32].

Definition 2.4.3. *A set $C \subset M$ is called* convex *if $\gamma([0,1]) \subset C$ for each geodesic $\gamma : [0,1] \to M$ joining the points $\gamma(0), \gamma(1) \in C$.*

A function $f : C \to \mathbb{R}$ is called convex if C is a convex set and for each geodesic $\gamma : [0,1] \to C$ the composition $\varphi \circ \gamma$ is a convex function in the usual sense, that is,

$$f(\gamma(t)) \leq (1-t) f(\gamma(0)) + t f(\gamma(1))$$

for all $t \in [0,1]$.

The function f *is called concave if* $-f$ *is convex.*

We remark that Jensen's inequality works in the context of global NPC spaces (despite the fact that the property of associativity of convex combinations fails). A probabilistic approach is made available by the paper of Sturm [163]. The main ingredient here is the *barycenter* of a discrete probability measures $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ is defined by the formula

$$\operatorname{bar}(\lambda) = \arg\min_{z \in M} \frac{1}{2} \sum_{i=1}^{n} \lambda_i d^2(z, x_i).$$

In the case of Hilbert spaces, this coincides with the usual definition of barycenter in flat spaces, that is, $\sum_{i=1}^{n} \lambda_i x_i$.

Theorem 2.4.7. (The discrete form of Jensen's Inequality). *For every continuous convex function* $f : M \to \mathbb{R}$ *and every discrete probability measure* $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ *on* M*, we have the inequality*

$$f(\operatorname{bar}(\lambda)) \le \sum_{i=1}^{n} \lambda_i f(x_i).$$

The result of Theorem 2.4.7 is a particular case of the integral form of Jensen's Inequality, which was first noticed by Jost [79] (and later extended by Eells and Fuglede [55]).

In what follows we shall deal with the relation of weighted majorization $\prec$, for pairs of discrete probability measures. See [126], for an extension of the Hardy-Littlewood-Pólya Theorem to the context of global NPC spaces.

Taking into account the *barycenter* of a discrete probability measures $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$, which works in the context of global NPC spaces only for positive weights $(\lambda_i)_i$ with $\sum_{i=1}^{n} \lambda_i = 1$, our future aim is to extend the above concept of barycenter to real weights, chosen in a similar way as in Jensen-Steffensen's inequality from this section.

The second future aim is related cu a relaxed concept of convexity, namely relative convexity. In [118] we discuss the availability of Jensen's inequality in a nonconvex context, in which we emphasize the usefulness of the concept of point of convexity. Even in the case of spaces with a curved geometry we have successfully introduced the point of convexity. In [124] we have discussed the meaning of the relative convexity notion.

We briefly present here the main ideas which will be used to treat the case of nonpositive weights in this context.

Definition 2.4.4. *Let* $f : M \to \mathbb{R}$ *be a continuous function. A point* $a \in M$ *is a* point of convexity *of the function* f *if*

$$f(a) \le \sum_{i=1}^{n} \lambda_i f(x_i), \tag{2.4.17}$$

for every family of points $x_1, \ldots, x_n$ *in* M *and every family of positive weights* $\lambda_1, \ldots, \lambda_n$ *with* $\sum_{i=1}^{n} \lambda_i = 1$ *and* $\operatorname{bar}\left(\sum_{i=1}^{n} \lambda_i \delta_{x_i}\right) = a$.

The point a is a point of concavity *if it is a point of convexity for $-f$ (equivalently, if the above inequality works in the reversed way).*

An example which illustrate the meaning of the above concept is offered by the function $f(x) = xe^x$. This function is concave on $(-\infty, -2]$ and convex on $[-2, \infty)$ (attaining a global minimum at $x = -1$). Every point $a \geq -1$ is a point of convexity because the tangent line $y = T(x)$ at the point (a, ae^a) gives supporting line for the graph of the function. A simple computation show that

$$f(a) = T(a) = T\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i T(x_i) \leq \sum_{i=1}^{n} \lambda_i f(x_i),$$

for every $x_1, \ldots, x_n \in \mathbb{R}$ and $\lambda_1, \ldots, \lambda_n \in [0,1]$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and $a = \sum_{i=1}^{n} \lambda_i x_i \geq -1$.

Other interesting connections between the subdifferential of the function and the notion of relative convexity can be found in [118]. More precisely, if a function admits a supporting hyperplane at a point a, then a is a point of convexity. In other words, every point at which the subdifferential is nonempty is a point of convexity. For other details, see for instance, [46, 82, 170, 171]. In this context, another future aim is to put in the same context the existence of the *point of convexity* of a function with the weakly or strongly convexity property. This could be done via the notion of subdifferential, which is somehow connected with the modulus C from the definition of weakly or strongly convexity. The choose of the optimal constant here is also another interesting purpose to be done.

Chapter 3

Jensen Steffensen's inequalities on spaces with curved geometry

3.1 On metric spaces of nonpositive curvature

In the first part of this chapter, we present some preliminary notions related to metric spaces of nonpositive curvature ("NPC spaces"), but also a discussion of barycenters of probability measures on such spaces. We will concentrate on analytic and stochastic aspects of nonpositive curvature. We are inspired by [163].

3.1.1 Geodesic spaces

We say that a curve in a metric space (N, d) is a continuous map $\rho : I \to N$ where $I \subset \mathbb{R}$ is some interval and we define its length $L_d(\rho)$ as the supremum of $\sum_{k=1}^{n} d(\rho_{\lambda_k}, \rho_{\lambda_{k-1}})$ where $\lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_n$ and $\lambda_0, \ldots, \lambda_n \in I$.

A curve is called geodesic if and only if

$$d(\rho_\mu, \rho_\lambda) = d(\rho_\mu, \rho_\nu) + d(\rho_\nu, \rho_\lambda),$$

for all $\mu, \nu, \lambda \in I$ with $\mu < \nu < \lambda$. Or otherwise written, iff $L_d(\rho|_{[\nu,\lambda]}) = d(\rho_\nu, \rho_\lambda)$ for all $\nu, \lambda \in I$ with $\nu < \lambda$.

In the sense of Reimannian geometry, that geodesics are only required to minimize locally the length (i.e. the above holds true only if $\mid \nu - \lambda \mid$ is sufficiently small) whereas geodesics in our sense are always globally minimizing the length.

A curve $\rho : [a, b] \to N$ connects the points $y, z \in N$ if and only if $\rho_a = y$ and $\rho_b = z$ and this implies that $L_d(\rho) \geq d(y, z)$.

Definition 3.1.1. *A metric space is called a length space (or inner metric space) if and only if for all $y, z \in N$ we have*

$$d(y, z) = \inf_\rho L_d(\rho),$$

where the infimum is taken over all curves which connect y and z. It is called a geodesic space if and only if each pair of points $y, z \in N$ is connected by a curve ρ of length $L_d(\rho) = d(y, z)$. This curve is not required to be unique.

Proposition 3.1.1. *A complete metric space (N, d) is a geodesic space if and only if $\forall z_0, z_1 \in N$, $\exists t \in N$ such that*

$$d(z_0, t) = d(z_1, t) = \frac{1}{2}(z_0, z_1).$$

In this context, any point $t \in N$ with the above properties will be called midpoint of z_0 and z_1.

Proof. Given $z_0, z_1 \in N$, we obtain their midpoint $z_{1/2} \in N$. Then the points $x_{1/4}$ and $z_{3/4}$ are obtained as midpoints of z_0 and $z_{1/2}$ or $z_{1/2}$ and z_1 respectively. Using this algorithm, we obtain the points z_λ for all dyadic $\lambda \in [0, 1]$ and obviously

$$d(z_\mu, z_\lambda) = d(z_\mu, z_\nu) + d(z_\nu, z_\lambda),$$

for all dyadic $0 \leq \mu < \nu < \lambda \leq 1$. By completeness of N, it yields the existence $z_\lambda \in N$ for all $\lambda \in [0, 1]$ such that $z : [0, 1] \to N$ is a geodesic. $\square$

Remark 14. *Note that a characterization in terms of "approximate midpoints" similar to Proposition 3.1.1 holds true for length spaces:*

A complete metric space (N, d) is a length space or geodesic space if and only if for all $z_0, z_1 \in N$ and $\epsilon > 0$ (or for $\epsilon = 0$ respectively) there exists $y \in N$ such that

$$d^2(z_0, y) + d^2(z_1, y) \leq \frac{1}{2} d^2(z_0, z_1) + \epsilon.$$

Let (N, d) be a geodesic space.

Definition 3.1.2. *A set $N_0 \subset N$ is called convex iff $\rho([0, 1]) \subset N_0$ for each geodesic $\rho : [0, 1] \to N$ with $\rho_0, \rho_1 \in N_0$. A function $f : N \to \mathbb{R}$ is called convex iff the function $f \circ \rho : [0, 1] \to \mathbb{R}$ is convex for each geodesic $\rho : [0, 1] \to N$, that is iff $\forall \lambda \in [0, 1]$*

$$f(\rho_\lambda) \leq (1 - \lambda) f(\rho_0) + \lambda f(\rho_1).$$

Proposition 3.1.2. *For $f : N \to \mathbb{R}$ define its epigraph*

$$Epi_f = \{(z, \mu) \in N \times \mathbb{R} : f(z) \leq \mu\} \subset N \times \mathbb{R}.$$

Then

(i) *f is convex if and only if N_f is convex.*

(ii) *f is lower semicontinuous if and only if N_f is closed.*

Proof. (i) Let N_f a subset of the space $\hat{N} = N \times \mathbb{R}$ with the metric

$$d((z, \mu), (y, \nu)) = (d^2(z, y) + \mid \mu - \nu \mid^2)^{1/2}.$$

Thus, $\hat{\rho} : [0,1] \to \hat{N}$ is a geodesic if and only if $\hat{\rho}(\lambda) = (\rho(\lambda), c_0 + c_1\lambda)$ with a geodesic $\rho : [0,1] \to N$ and $c_0, c_1 \in \mathbb{R}$. Further consider $\hat{\rho}$ be a geodesic with $\hat{\rho}(0), \hat{\rho}(1) \in N_f$, that is with

$$f \circ \rho(0) \leq c_0 \text{ and } f \circ \rho(1) \leq c_0 + c_1.$$

Convexity of $f : N \to \mathbb{R}$ implies convexity of $f \circ \rho : [0,1] \to \mathbb{R}$ and this in turn

$$f \circ \rho(\lambda) \leq c_0 + c_1\lambda,$$

or, in other words, $\hat{f}(\lambda) \in N_f$. This provides the convexity on N_f.

Conversely, we assume that N_f is convex. Let $\rho : [0,1] \to N$ be any geodesic. Choose

$$c_0 = f \circ \rho(0), c_1 = f \circ \rho(1) - f \circ \rho(0) \text{ and } \hat{\rho}(\lambda) := (\rho(\lambda), c_0 + c_1\lambda).$$

Then $\hat{\rho}(0)$, $\hat{\rho}(1) \in N_f$ and thus also $\hat{\rho} \in N_f$. Previous results states that

$$f \circ \rho(\lambda) \leq c_0 + c_1\lambda = (1-\lambda) f \circ \rho(0) + \lambda f \circ \rho(1),$$

for all $\lambda \in [0,1]$. That is, $f \circ \rho : [0,1] \to \mathbb{R}$ is convex for each geodesic $\rho : [0,1] \to N$ and thus $f : N \to \mathbb{R}$ is convex.

(ii) N_f is closed $\iff$ $(\hat{z}_n \to \hat{z}, \hat{z}_n \in N_f \implies \hat{z} \in N_f)$ $\iff$ $(z_n \to z, \mu_n \to \mu \implies f(z) \leq \mu)$ $\iff$ f is lower semicontinuous. □

Definition 3.1.3. *A function $f : N \to \mathbb{R}$ is called uniformly convex if and only if there exists a strictly increasing function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ such that for any geodesic $\rho : [0,1] \to N$*

$$f(\rho_{1/2}) \leq \frac{1}{2}(f(\rho_0) + f(\rho_1)) - \eta(d(\rho_0, \rho_1)).$$

A function f is called strictly convex iff for any geodesic $\rho : [0,1] \to N$ with $\rho_0 \neq \rho_1$

$$f(\rho_{1/2}) < \frac{1}{2}(f(\rho_0) + f(\rho_1)).$$

Proposition 3.1.3. *Let $f : N \to \mathbb{R}$ be a uniformly convex, lower semicontinuous function on a complete geodesic space (N, d). Then there exists a unique minimizer $t \in N$, i.e. a unique point $t \in N$ with $f(t) = \inf_{w \in N} f(w)$. We write*

$$t = \arg\min_{w \in N} f(w).$$

Proof. (i) Existence: Let t_n be a sequence of points in N with $\lim_n f(t_n) = \inf_t f(t) := \alpha$ and let $t_{n,k}$ the midpoint between t_n and t_k.

For $n, k \to \infty$ we have that

$$\alpha \leq f(t_{n,k}) \leq \frac{1}{2} f(t_n) + \frac{1}{2} f(t_k) - \eta(d(t_n, t_k)).$$

Hence, $d(t_n, t_k) \to 0$ for $n, k \to \infty$. In other words, $(t_n)_n$ is a Cauchy sequence, so there exists $t^* = \lim_{n\to\infty} t_n \in N$ since N is complete. Moreover, $f(t^*) = \inf_t f(t)$ by lower semicontinuity of f.

(ii) Uniqueness: Assume that $f(t_0) = f(t_1) = \inf_t f(t) = \alpha$ and $t_0 \neq t_1$. We get the contradiction $\alpha \leq f(t_{\frac{1}{2}}) < \frac{1}{2}\alpha + \frac{1}{2}\alpha$, for $t_{\frac{1}{2}}$ the midpoint between t_0 and t_1. □

Remark 15. *For the uniqueness of the minimizer it is suffices to require that f is strictly convex. If N is compact then for the existence of the minimizer it is suffices to require that f is convex and lower semicontinuous.*

Definition 3.1.4. *A geodesic space (N, d) is called doubly convex if and only if the function d : $(z, y) \to d(z, y)$ is convex on $N \times N$ or, in other words, iff the function $\lambda \to d(\rho_\lambda, \eta_\lambda)$ is convex for each pair of geodesics $\rho, \eta : [0, 1] \to N$. It is called strictly doubly convex if and only if d is strictly convex on $N \times N$.*

Remark 16. (*i*) *In a doubly convex geodesic space, any two of its points are joined by a unique geodesic and this geodesics depends continuously on its endpoints.*

(*ii*) *If a geodesic space is locally doubly convex and simply connected then it is doubly convex [55, 64].*

Let $(M, \mathcal{M})$ be a measurable space and (N, d) be a metric space. A map $v : M \to N$ is called measurable if and only if it is measurable with respect to the given σ-field $\mathcal{M}$ on M and the Borel σ-field $\mathcal{B}(N)$ on N, i.e. iff $v^{-1}(N') \in \mathcal{M}$ for all $N' \in \mathcal{B}(N)$. Note that for the latter it suffices that $v^{-1}(N') \in \mathcal{M}$ for all open $N' \subset N$.

3.1.2 Global NPC spaces

Inspired by Sturm [163], we present an introduction to metric spaces of nonpositive curvature, NPC spaces, with emphasis on analytic and stochastic aspects of nonpositive curvature. In this context, we use the explicit estimates for the distance function, we do not deal with triangle or angle comparison and we do not introduce the tangent cone or the space of directions.

For the many and deep geometric aspects we refer to the huge literature on NPC spaces. The whole development started with the investigations of A. D. Alexandrov [7] and Yu. G. Reshetnyak [147] and was strongly influenced by the work of M. Gromov [63]. Recently, there appeared various monographs devoted exclusively to NPC spaces: [17], [32] and [80]. Also the monographs [16, 33, 55] contain much material on this subject. Moreover, we recommend the articles [6, 79, 87].

Definition 3.1.5. *A metric space (N, d) is called global NPC space if it is complete and if for each pair of points $z_0, z_1 \in N$ there exists a point $y \in N$ with the property that for all points $t \in N$ we have that*

$$d^2(t, y) \leq \frac{1}{2}d^2(t, z_0) + \frac{1}{2}d^2(t, z_1) - \frac{1}{4}d^2(z_0, z_1). \tag{3.1.1}$$

In this context, "NPC" means nonpositive curvature. Global NPC spaces are also called *Hadamard spaces.* Property (3.1.1) is called the NPC inequality and it can be weakened.

Remark 17. *A complete metric space (N, d) is a global NPC space if and only if for all $z_0, z_1 \in N$ and $\epsilon > 0$ there exists $y \in N$ such that for all $t \in N$ we have that*

$$d^2(t, y) \leq \frac{1}{2}d^2(t, z_0) + \frac{1}{2}d^2(t, z_1) - \frac{1}{4}d^2(z_0, z_1) + \epsilon. \tag{3.1.2}$$

Proposition 3.1.4. *If (N, d) is a global NPC space then is a geodesic space. Even more, for any pair of points $z_0, z_1 \in N$ there exists a unique geodesic $z : [0, 1] \to N$ connecting them. For $\lambda \in [0, 1]$ the intermediate points z_λ depend continuously on the endpoints z_0, z_1. Finally, for any $t \in N$ we have*

$$d^2(t, z_\lambda) \leq (1 - \lambda)d^2(t, z_0) + \lambda d^2(t, z_1) - \lambda(1 - \lambda)d^2(z_0, z_1). \tag{3.1.3}$$

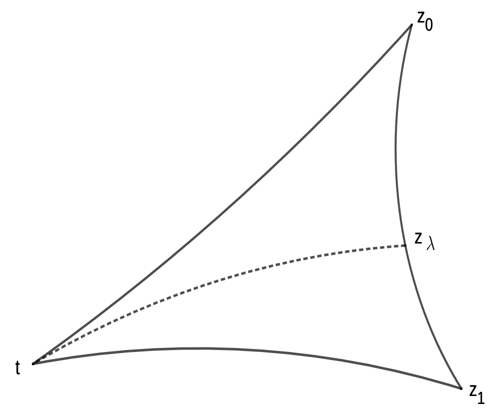

Figure 3.1: NPC inequality

For the convenience of the reader we present the proof given in [163].

Proof. (i) Choosing $t = z_0$ or $t = z_1$ in (3.1.1) yields

$$d(z_0, z_{1/2}) \leq \frac{1}{2} d(z_0, z_1) \text{ and } d(z_2, z_1) \leq \frac{1}{2} d(z_0, z_1).$$

Whence, z_2 is a midpoint and (N, d) is a geodesic space. Choosing t to be any other midpoint of z_0 and z_1 yields $d(t, z_{1/2}) = 0$. That is, midpoints are unique and thus also geodesic are unique.

(ii) Given any geodesic $z : [0, 1] \to N$ it suffices to prove (3.1.3) for all dyadic $\lambda \in [0, 1]$. It obviously holds for $\lambda = 0$ and $\lambda = 1$. Assume that it holds for all $\lambda = k2^{-n}$ with $k = 0, 1, ..., 2^n$. We want to prove that then (3.1.3) also holds for all $\lambda = k2^{-(n+1)}$ for all $\lambda = k2^{-n}$ with $k = 0, 1, ..., 2^{n+1}$. For even k this is just the assumption. Fix $\lambda = k2^{-(n+1)}$ with an odd k and put $\Delta\lambda = 2^{-(n+1)}$. Then by (3.1.1) we have that

$$d^2(t, z_{1/2}) \leq \frac{1}{2} d^2(t, z_{\lambda-\Delta\lambda}) + \frac{1}{2} d^2(t, z_{\lambda+\Delta\lambda}) - \frac{1}{4} d^2(z_{\lambda-\Delta\lambda}, z_{\lambda+\Delta\lambda})$$

and by (3.1.3)

$$d^2(t, z_{\lambda\pm\Delta\lambda}) \leq (1 - \lambda \mp \Delta\lambda) d^2(t, z_0) + (\lambda \pm \Delta\lambda) d^2(t, z_1) - (1 - \lambda \mp \Delta\lambda)(\lambda \pm \Delta\lambda) d^2(z_0, z_1)$$

Thus

$$\begin{aligned} d^2(t, z_\lambda) &\leq (1 - \lambda) d^2(t, z_0) + \lambda d^2(t, z_1) \\ &- \left[(\Delta\lambda)^2 - \frac{1}{2}(1 - \lambda - \Delta\lambda)(\lambda + \Delta\lambda) - \frac{1}{2}(1 - \lambda + \Delta\lambda)(\lambda - \Delta\lambda) \right] d^2(z_0, z_1) \\ &= (1 - \lambda) d^2(t, z_0) + \lambda d^2(t, z_1) - \lambda(1 - \lambda) d^2(z_0, z_1). \end{aligned}$$

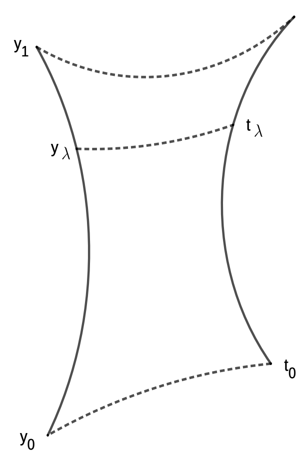

Figure 3.2: Geodesic comparison

(iii) Now let $z, y : [0,1] \to N$ be two geodesics. Then applying (3.1.3) twice yields

$$d^2(z_\lambda, y_\lambda) \le (1-\lambda)d^2(z_0, y_\lambda) + \lambda d^2(z_1, y_\lambda) - \lambda(1-\lambda)d^2(z_0, z_1)$$

$$\le (1-\lambda)^2 d^2(z_0, y_0) + \lambda^2 d^2(z_1, y_1)$$

$$+\lambda(1-\lambda)[d^2(z_0, y_1) + d^2(z_1, y_0) - d^2(z_0, z_1) - d^2(y_0, y_1)].$$

Obviously, the right hand side converges to 0 if $y_0 \to z_0$ and $y_1 \to z_1$, and thus $y_\lambda \to z_\lambda$, that is z_λ depends continuously on z_0 and z_1. □

Corollary 5. *(Geodesic Comparison). Let (N, d) be a global NPC space, $\rho, \eta : [0,1] \to N$ be geodesics and $\lambda \in [0,1]$. Then*

$$d^2(\rho_\lambda, \eta_\lambda) \le (1-\lambda)d^2(\rho_0, \eta_0) + \lambda d^2(\rho_1, \eta_1) - \lambda(1-\lambda)[d(\rho_0, \rho_1) - d(\eta_0, \eta_1)^2 \tag{3.1.4}$$

and

$$d(\rho_\lambda, \eta_\lambda) \le (1-\lambda)d(\rho_0, \eta_0) + \lambda d(\rho_1, \eta_1).$$

Proof. We use what we demonstrated in part (iii) of Proposition 3.1.4 and we have that

$$d^2(\rho_\lambda, \eta_\lambda) - (1-\lambda)^2 d^2(\rho_0, \eta_0) - \lambda^2 d^2(\rho_1, \eta_1)$$

$$\le \lambda(1-\lambda)[d^2(\rho_0, \eta_1) + d^2(\rho_1, \eta_0) - d^2(\rho_0, \rho_1) - d^2(\eta_0, \eta_1)].$$

By quadruple comparison, the right hand side is

$$\le \lambda(1-\lambda)\big[d^2(\rho_0, \eta_0) + d^2(\rho_1, \eta_1)$$

$$-\mu(d(\rho_0, \eta_0) - d(\rho_1, \eta_1))^2 - (1-\mu)(d(\rho_0, \rho_1) - d(\eta_0, \eta_1))^2\big]$$

for each $\mu \in [0,1]$. For $\mu = 0$ this yields the Reshetnyak's quadruple comparison and for $\mu = 1$ it yields

$$\begin{aligned} d^2(\rho_\lambda, \eta_\lambda) &\leq (1-\lambda)d^2(\rho_0, \eta_0) + \lambda d^2(\rho_1, \eta_1) - \lambda(1-\lambda)[d(\rho_0, \eta_0) - d(\rho_0, \eta_1)]^2 \\ &= [(1-\lambda)d(\rho_0, \eta_0) + \lambda d(\rho_1, \eta_1)]^2. \end{aligned}$$

□

Inequality (3.1.4) assert that d is doubly convex, i.e. $(z, y) \to d(z, y)$ is a convex function on $N \times N$. We have some obvious consequences:

(i) For each $t \in N$ the function $z \to d(z,t)$ is convex; in particular, all balls $B_r(t) \subset N$ are convex.

(ii) Geodesics depend continuously on their endpoints in the following quantitative way:

$$d_\infty(\eta, \rho) = \sup\{d(\eta_0, \rho_0), d(\eta_1, \rho_1)\},$$

where for any curves $\eta, \rho : [0,1] \to N$ we put $d_\infty(\eta, \rho) := sup\{d(\eta_\lambda, g_\lambda) : \lambda \in [0,1]\}$.

(iii) N is contractible and, in particular, simply connected.

Proposition 3.1.5. *(Convex Projection).*

(i) *For each convex closed set $K \subset N$ in a global NPC space (N, d) there exists a unique map $\pi_K : N \to K$ ("projection onto K") with*

$$d(\pi_K(t), t) = \inf_{w \in K} d(w, t) \;\; (t \in N);$$

(ii) *π_K is "orthogonal":*

$$d^2(t, w) \geq d^2(t, \pi_K(t)) + d^2(\pi_K(t), w) \quad (t \in N, w \in K);$$

(iii) *π_K is a contraction:*

$$d(\pi_K(t), \pi_K(w)) \leq d(t, w) \quad\quad (t, w \in N).$$

Proof. (i) Fix $t \in N$ and a closed convex set $K \subset N$. Then K is a global NPC space and the function $f : K \to R$, $z \to d^2(z, t)$ is continuous and uniformly convex on K. Hence there exists a unique minimizer in K.

(ii) Let $\lambda \to w_\lambda$ be a geodesic joining $w_0 := \pi_K(t)$ and $w_1 := w$. Then $w_\lambda \in K$ for all $\lambda \in [0,1]$ by convexity and closedness of K. Hence, by the NPC inequality

$$d^2(\pi_k(t), t) \leq d^2(w_\lambda, t) \leq (1-\lambda)d^2(\pi_K(t), t) + \lambda d^2(w, t) - \lambda(1-\lambda)d^2(\pi_K(t), w)$$

and therefore

$$d^2(\pi_k(t), t) + (1-\lambda)d^2(\pi_K(t), w) \leq d^2(w, t).$$

(iii) Put $t' = \pi_K(t)$, $w' = \pi_K(w)$. Then (ii) and quadruple comparison imply

$$d^2(t, w) + d^2(w, w') + d^2(w', t') + d^2(t', t) \geq d^2(t, w') + d^2(t', w) \geq d^2(t, t') + d^2(w, w') + 2d^2(w', t'),$$

which yields the claim. □

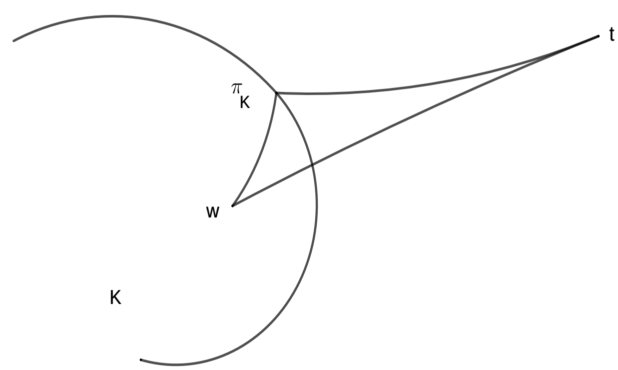

Figure 3.3: Convex projection

The important fact here is the existence of a unique projection without assuming any kind of compactness of K.

Remark 18. *(i) Given any subset $A \subset N$ in a global NPC space (N, d), there exists a unique smallest convex set $C(A)$ containing A, called convex hull of A. It can be constructed as $C(A) = \bigcup_{n=0}^{\infty} A_n$ where $A_0 := A$ and for $n \in N$, the set A_n consists of all points in N which lie on geodesics which start and end in A_{n-1}.*

(ii) Given any bounded subset $A \subset N$ in a global NPC space (N, d) there exists a unique closed ball of minimal radius which contains A. In other words, there exists a unique point $z \in N$ (the circumcenter of A) such that

$$r(z, A) = \inf_{t \in N} r(t, A),$$

where $r(t, A) := \sup_{y \in A} d(t, y)$. This is an immediate consequence of Proposition 3.1.3 since the function $t \to r^2(t, A)$ is uniformly convex.

3.1.3 Examples of global NPC spaces

This subsection, inspired by Sturm [163] gives some examples of global NPC spaces. The main examples, in our context, are manifolds, trees and Hilbert spaces. Other examples are cones, buildings and surfaces of revolution. New global NPC spaces can be built out of given global NPC spaces as subsets, images, gluings, products or L^2-spaces.

Proposition 3.1.6. *(Manifolds). Let (N, d) be a Reimannian manifold and let d be its Reimannian distance. Then (N, d) is a global NPC space if and only if it is complete, simply connected and of nonpositive (sectional) curvature.*

Besides manifolds, the most important examples of NPC spaces are trees, in particular, spiders.

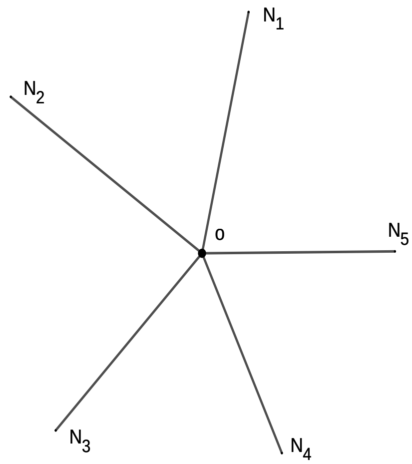

Figure 3.4: The 5-spider

Example 3. *(Spiders). Let K be an arbitrary set and for each $i \in K$ let $N_i = \{(i, \mu) : \mu \in \mathbb{R}_+\}$ be a copy of $\mathbb{R}_+$ (equipped with the usual metric). Define the spider over K or K-spider (N, d) by gluing together all these spaces N_i, $i \in K$, at their origins, i.e.*

$$N = \{(i, \mu) : i \in K, \mu \in \mathbb{R}_+\}/\sim \textit{ where } (i, 0) \sim (j, 0)(\forall i, j)$$

and

$$d((i, \mu), (j, \nu)) = \begin{cases} |\mu - \nu| & \textit{if } i = j \\ |\mu| + |\nu|, & \textit{else.} \end{cases}$$

The rays N_i can be regarded as closed subsets of N. Any two rays N_i and N_j with $i \neq j$ intersect at the origin $o := (i, 0) = (j, 0)$ of N.
The K-spider N depends (upto isometry) only on the cardinality of K. If $K = \{1, \ldots, k\}$ for some $k \in \mathbb{N}$ then it is called k-spider. It can be realized as a subset of the complex plane

$$\left\{\mu \cdot exp\left(\frac{l}{k} 2\pi i\right) \in \mathbb{C} : \mu \in \mathbb{R}_+, l \in \{1, \ldots, k\}\right\},$$

however, equipped with a non-Euclidian metric. If $k = 1$ or $k = 2$ then it is isometric to $\mathbb{R}_+$ or $\mathbb{R}$, respectively. The 3-spider is also called tripod.

Proposition 3.1.7. *(Trees). Each metric tree is a global NPC space.*

Proof. We have to prove the NPC inequality (3.1.1) for each triple of points $z_0, z_1, t \in N$. Without restriction, we may replace N by the convex hull of these three points which is isometric to the convex hull of three points in the tripod. That is, without restriction N is the tripod.

Firstly, consider the case where z_0, z_1, t lie on one geodesic $\rho : I \to N$. Then ρ is an isometry between $I \subset \mathbb{R}$ and $\rho(I) \subset N$. Since I is globally NPC, so is $\rho(I)$. Actually, I and thus $\rho(I)$ are even "flat", i.e.

$$d^2(t, z_{1/2}) = \frac{1}{2}d^2(t, z_0) + \frac{1}{2}d^2(t, z_1) - \frac{1}{4}d^2(z_0, z_1),$$

for all $z_0, z_1, t \in \rho(I)$ and with $z_{1/2}$ being the midpoint of z_0, z_1.

Secondly, consider the non-degenerate case $z_0 = (i, \mu)$, $z_1 = (j, \nu)$ and $t = (k, \lambda)$ with $\mu \cdot \nu \cdot \lambda > 0$ and different $i, j, k \in \{1, 2, 3\}$. Assume without restriction $\mu \geq \nu$ and put $t' = (j, \lambda)$. Then $z_0, z_{1/2} \in N_i$ and $t' \in N_j$. The points z_0, z_1, t' lie on one geodesic. Therefore, by the previous considerations

$$d^2(t', z_{1/2}) = \frac{1}{2}d^2(t', z_0) + \frac{1}{2}d^2(t', z_1) - \frac{1}{4}d^2(z_0, z_1).$$

Moreover, $d(z_0, t) = d(z_0, t')$ and $d(z_{1/2}, t) = d(z_{1/2}, t')$ whereas $d(z_1, t) \geq d(z_1, t')$. Hence, finally,

$$\begin{aligned} d^2(t, z_{1/2}) = d^2(t', z_{1/2}) &= \frac{1}{2}d^2(t', z_0) + \frac{1}{2}d^2(t', z_1) - \frac{1}{4}d^2(z_0, z_1) \\ &\leq \frac{1}{2}d^2(t, z_0) + \frac{1}{2}d^2(t, z_1) - \frac{1}{4}d^2(z_0, z_1). \end{aligned}$$

□

Proposition 3.1.8. *(Hilbert spaces).*

(i) *Each Hilbert space is a global NPC space.*

(ii) *A Banach space is a global NPC space if and only if it is a Hilbert space.*

(iii) *A metric space is (derived from) a Hilbert space is and only if it is a nonempty, geodesically complete global NPC space with curvature ≥ 0. One possible (of many equivalent) definitions for the latter is to require that in* (3.1.1) *also the reverse inequality holds true.*

Proof. (i) Choosing $z_{1/2} = \frac{1}{2}(z_0 + z_1)$ yields equality in (3.1.1):

$$\left| t - \frac{z_0 + z_1}{2} \right|^2 = \frac{1}{2}|t - z_0|^2 + \frac{1}{2}|t - z_1|^2 - \frac{1}{4}|z_0 - z_1|^2.$$

(ii) Assume that N is a Banach and global NPC space. Given $z_0, z_1 \in N$, one (and hence the unique) midpoint is $z_{1/2} = \frac{z_0+z_1}{2}$. Then choosing $t = 0$ in (3.1.1) yields

$$|z_0 - z_1|^2 + |z_0 + z_1|^2 \leq 2|z_0|^2 + 2|z_1|^2,$$

which is a "parallelogram inequality". Replacing z_0 and z_1 in this inequality by $z_0 + z_1$ and $z_1 - z_1$, respectively, yields the opposite inequality and thus proves the parallelogram equality.

(iii) The "only if"-implication is easy. For the "if"-implication fix an arbitrary point $o \in N$. Then for each $z \in N$ there exists a unique geodesic $z : \mathbb{R} \to N$ with $z_0 = o$ and $z_1 = z$. Using these geodesics we define a scalar multiplication by $\beta \cdot z := z_\beta$ ($\forall \beta \in \mathbb{R}, z \in N$), an addition by $z + y :=$ midpoint of $2 \cdot z$ and $2 \cdot y$ ($\forall z, y \in N$), and an inner product by

$$\langle z, y \rangle := \frac{1}{2}(d^2(z, y) - d^2(o, z) - d^2(o, y)) \quad (\forall z, y \in N).$$

For details, see [87]. □

Lemma 3.1.1. *(Subsets). A subset $N_0 \subset N$ of a global NPC space N is a global NPC space if and only if it is closed and convex.*

3.1.4 Barycenters on global NPC spaces

Inspired by Sturm [163], we present some theoretical aspects related to barycenters on global NPC spaces.

Let (N,d) be a complete metric space and let $\mathcal{P}(N)$ denote the set of all probability measures p on $(N,\mathcal{B}(N))$ with separable support $\text{supp}(p) \subset N$. For $1 \le \theta < \infty$, $\mathcal{P}^\theta(N)$ will denote the set of $p \in \mathcal{P}(N)$ with $\int d^\theta(z,y)p(dy) < \infty$ for some $z \in N$, and $\mathcal{P}^\infty(N)$ will denote the set of all $p \in \mathcal{P}(N)$ with bounded support. Finally, we denote by $\mathcal{P}_0(N)$ the set of all $p \in \mathcal{P}(N)$ of the form $p = \frac{1}{n}\sum_{i=1}^n \delta_{z_i}$ with suitable $z_i \in N$. Here and henceforth, $\delta_z : A \to \mathbf{1}_A(z)$ denote the Dirac measure in the point $z \in N$. Obviously,

$$\mathcal{P}_0(N) \subset \mathcal{P}^\infty(N) \subset \mathcal{P}^\theta(N) \subset \mathcal{P}^1(N).$$

For $q \in \mathcal{P}(N)$ the number $var(q) := \inf_{t\in N} \int_N d^2(t,z)q(dz)$ is called variance of q. Of course, $q \in \mathcal{P}^2(N)$ if and only if $\text{var}(q) < \infty$.

Given $p,q \in \mathcal{P}(N)$ we say that $\gamma \in \mathcal{P}(N^2)$ is a coupling of p and q iff its marginals are p and q, that is, iff $\forall A \in \mathcal{B}(N)$

$$\gamma(A \times N) = p(A) \text{ and } \gamma(N \times A) = q(A). \tag{3.1.5}$$

One such coupling γ is the product measure $p \otimes q$.

Definition 3.1.6. *We define the (L^1–) Wasserstein distance or Kantorovich-Rubinstein distance d^W on $\mathcal{P}^1(N)$ by*

$$d^W(p,q) = \inf\left\{ \int_N\int_N d(z,y)\gamma(dzdy) : \gamma \in \mathcal{P}(N^2) \textit{ is coupling of } p \textit{ and } q\right\}.$$

Proposition 3.1.9. *Let (N,d) be a global NPC space and fix $y \in N$. For each $q \in \mathcal{P}^1(N)$ there exists a unique point $t \in N$ which minimizes the uniformly convex, continuous function $t \to \int_N [d^2(t,z) - d^2(y,z)]q(dz)$. This point is independent of y; it is called barycenter (or, more precisely, d^2-barycenter) of q and denoted by*

$$b(q) = \arg\min_{t\in N} \int_N [d^2(t,z) - d^2(y,z)]q(dz).$$

If $q \in \mathcal{P}^2(N)$ then $b(q) = \arg\min_{t\in N} d^2(t,z)q(dz)$.

Proof. Let $F_y(t) = \int [d^2(t,z) - d^2(y,z)]q(dz)$. Then

$$F_y(t) - F_{y'}(t) = \int [d^2(y',z) - d^2(y,z)]q(dz)$$

is dependent of t. Moreover, $|F_y(t) < \infty|$ since

$$|F_y(t)| = \left| \int_N [d(t,z) - d(y,z)] \cdot [d(t,z) + d(y,z)]q(dz) \right|$$
$$\le d(t,y) \cdot \left[\int_N d(t,z)q(dz) + \int_N d(y,z)q(dz) \right].$$

The uniform convexity of $t \to d^2(t,z)$ as stated in Proposition 3.1.4 implies that $t \to F_y(t)$ is uniformly convex: For any two points $t_0, t_1 \in N$ let $\lambda \to t_\lambda$ denote the geodesic. Application of (3.1.3) gives

$$F_y(t_\lambda) = \int [d^2(t_\lambda,z) - d^2(y,z)]q(dz)$$

$$\leq (1-\lambda)\int [d^2(t_0,z)-d^2(y,z)]q(dz)+\lambda\int [d^2(t_1,z)-d^2(y,z)]q(dz)$$
$$-\lambda(1-\lambda)d^2(t_0,t_1)$$
$$=(1-\lambda)F_y(t_0)+\lambda F_y(t_1)-\lambda(1-\lambda)d^2(t_0,t_1).$$

Moreover, continuity of $t\to F_y(t)$ is obvious from

$$|F_y(t)-F_y(t')|\leq \int_N |d^2(t,z)-d^2(t',z)|q(dz).$$

According to Proposition 3.1.3, uniform convexity and lower semicontinuity of F_y implies existence and uniqueness of a minimizer. □

Proposition 3.1.10. *(Variance Inequality). Let (N,d) be a global NPC space. For any probability measure $q\in\mathcal{P}^1(N)$ and for all $t\in N$:*

$$\int_N [d^2(t,z)-d^2(b(q),z)]q(dz)\geq d^2(t,b(q)). \tag{3.1.6}$$

Proof. Given q and t, apply the estimate from the previous proof with $t_1:=t$, $t_0:=b(q)$ and $y:=b(q)$. The fact that $b(q)$ is minimizer yields

$$0\leq F(t_\lambda)\leq 0+\lambda\cdot F(t)-\lambda(1-\lambda)d^2(t,b(q)).$$

That is, for all $\lambda>0$

$$\int_N [d^2(t,z)-d^2(b(q),z)]q(dz)\geq (1-\lambda)d^2(t,b(q)).$$

For $\lambda\to 0$ this yields the claim. □

Now, we want to present another natural way to define the "expectation" $\mathbf{EY}$ of a random variable $\mathbf{Y}$ is to use generalization of the law of large numbers. This requires to give a meaning to $\frac{1}{n}\sum_{i=1}^n \mathbf{Y}_i$. Our definition below only uses the fact that any two points in N are joined by unique geodesics. Our law of large numbers for global NPC spaces gives convergence towards the expectation defined as minimizer of the L^2 distance.

Definition 3.1.7. *Given any sequence $(y_i)_{i\in N}$ of points in N we define a new sequence $(v_n)_{n\in\mathbb{N}}$ of points $v_n\in N$ by induction on n as follows:*

$$v_1:=y_1 \text{ and } v_n:=\left(1-\frac{1}{n}\right)v_{n-1}+\frac{1}{n}y_n,$$

where the RHS should denote the point $\rho_{1/n}$ on the geodesic $\rho:[0,1]\to N$ connecting $\rho_0=v_{n-1}$ and $\rho_1=y_n$. The point v_n will be denoted by $\frac{1}{n}\overset{\rightarrow}{\sum}_{i=1,\dots,n} y_i$ and called inductive mean value of $y_1,\dots,y_n$.

Note that in general the point $\frac{1}{n}\overset{\rightarrow}{\sum}_{i=1,\dots,n} y_i$ will strongly depend on permutations of the y_i.

Theorem 3.1.1. *(Law of Large Numbers). Let $(\mathbf{Y}_i)_{i\in\mathbb{N}}$ be a sequence of independent, identically distributed random variables $\mathbf{Y}_i\in L^2(\Omega,N)$ on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ with values in a global NPC space (N,d). Then*

$$\frac{1}{n}\overset{\rightarrow}{\sum_{i=1,\dots,n}} Y_i\to \mathbb{E}Y_1 \text{ for } n\to\infty,$$

in $L^2(\Omega, N)$ and in probability ("weak low of a large numbers").
If moreover $Y_i \in L^\infty(\Omega, N)$ then for $\mathbb{P}$-almost every $\omega \in \Omega$

$$\frac{1}{n} \overset{\rightarrow}{\sum_{i=1,\ldots,n}} Y_i(\omega) \to \mathbb{E}Y_1 \text{ for } n \to \infty.$$

Finally, we will give various characterizations of nonpositive curvature in terms of properties of probability measures on the spaces. For instance, the validity of a variance inequality turns out to characterize NPC spaces. Similarly, an inequality between two kind of variances as well as a weighted quadruple inequality.

Theorem 3.1.2. *Let (N, d) be a complete metric space. Then the following properties are equivalent:*

(i) (N, d) is a global NPC space.

(ii) For any probability measure $q \in \mathcal{P}^2(N)$ there exists a point $t_q \in N$ such that for all $t \in N$

$$\int_N d^2(t, z)q(dz) \geq d^2(t, t_q) + \int_N d^2(t_q, z)q(dz). \tag{3.1.7}$$

(iii) For any probability measure $q \in \mathcal{P}(N)$

$$var(q) \leq \frac{1}{2} \int_N \int_N d^2(z, y)q(dz)q(dy).$$

(iv) (N, d) is a length space with the property that for any $z_1, z_2, z_3, z_4 \in N$ and $\nu, \lambda \in [0, 1]$

$$\nu(1-\nu)d^2(z_1, z_3) + \lambda(1-\lambda)d^2(z_2, z_4)$$
$$\leq \nu\lambda d^2(z_1, z_2) + (1-\nu)\lambda d^2(z_2, z_3) + (1-\nu)(1-\lambda)d^2(z_3, z_4) + \nu(1-\lambda)d^2(z_4, z_1)$$

The proof will show that in (iii) it suffices to consider probability measures q which are supported by four points and in (iv) it suffices to consider $\lambda = \frac{1}{2}$.

Proof. $(i) \implies (ii)$: We use corollary 3.1.6.

$(ii) \implies (i)$: Given points $\rho_0, \rho_1 \in N$ and $\lambda \in [0, 1]$, choose the probability measure $q = (1-\lambda)\delta_{\rho_0} + \lambda\delta_{\rho_1}$ and denote the point t_q by ρ_λ. Then (ii) implies for all $t \in N$

$$(1-\lambda)d^2(t, \rho_0) + \lambda d^2(t, \rho_1) \geq (1-\lambda)d^2(\rho_\lambda, \rho_0) + \lambda d^2(\rho_\lambda, \rho_1) + d^2(\rho_\lambda, t)$$
$$\geq (1-\lambda)\lambda d^2(\rho_0, \rho_1) + d^2(\rho_\lambda, t),$$

where the last inequality is a simple consequence of the triangle inequality. This proves (i).

$(ii) \implies (iii)$: If $\mathrm{var}(q) = \infty$ then $\int_N d^2(z, y)q(dz) = \infty$ for all $y \in N$ and the claim follows. Therefore, we may assume $\mathrm{var}(q) < \infty$. In this case, the claim follows from integrating (3.1.9) against $q(dt)$.

$(iii) \implies (iv)$: Let ρ_0, ρ_1 be any two points in N and $\epsilon > 0$. Choose the probability measure $q = \frac{1}{2}\delta_{\rho_0} + \frac{1}{2}\delta_{\rho_1}$. Then (iii) implies that there exists a point $t \in N$ with

$$d^2(t, \rho_0) + d^2(t, \rho_1) \leq \frac{1}{2}d^2(\rho_0, \rho_1) + \epsilon.$$

According to Remark 14, this already implies that (N,d) is a length space.

To see the second claim, choose

$$q = \frac{1}{2}[\nu\delta_{z_1} + \lambda\delta_{z_2} + (1-\nu)\delta_{z_3} + (1-\lambda)\delta_{z_4}].$$

Then for each $\epsilon > 0$, (iii) implies that for suitable $t \in N$

$$\frac{1}{8}[\nu\lambda d^2(z_1,z_2) + (1-\nu)\lambda d^2(z_2,z_3) + (1-\nu)(1-\lambda)d^2(z_3,z_4)$$

$$+\nu(1-\lambda)d^2(z_4,z_1) + \nu(1-\nu)d^2(z_1,z_3) + \lambda(1-\lambda)d^2(z_2,z_4)] + \epsilon$$

$$\geq \frac{1}{4}[\nu d^2(t,z_1) + \lambda d^2(t,z_2) + (1-\nu)d^2(t,z_3) + (1-\lambda)d^2(t,z_4)]$$

$$\geq \frac{1}{4}[\nu(1-\nu)d^2(z_1,z_3) + \lambda(1-\lambda)d^2(z_2,z_4)],$$

where again the last inequality is a simple consequence of the triangle inequality. Since this holds for any $\epsilon > 0$ it proves the claim.

$(iv) \implies (i)$: The fact that (N,d) is a length space implies that, given $\rho_0, \rho_1 \in N$ and $\nu > 0$, there exists $y \in N$ such that

$$d^2(\rho_0,y) + d^2(\rho_1,y) \leq \frac{1}{2}d^2(\rho_0,\rho_1) + \nu^2.$$

For arbitrary $t \in N$, apply (iv) to $z_1 = t, z_2 = \rho_1, z_3 = y, z_4 = \rho_0$ and $\lambda = \frac{1}{2}$. It yields

$$\begin{aligned}\nu(1-\nu)d^2(t,\rho_\lambda) &\leq \frac{\nu}{2}d^2(t,\rho_1) + \frac{\nu}{2}d^2(t,\rho_0) + \frac{1-\nu}{2}d^2(y,\rho_1) + \frac{1-\nu}{2}d^2(y,\rho_0) - \frac{1}{4}d^2(\rho_0,\rho_1)\\ &\leq \frac{\nu}{2}d^2(t,\rho_1) + \frac{\nu}{2}d^2(t,\rho_0) - \frac{\nu}{4}d^2(\rho_0,\rho_1) + \frac{\nu^2}{2}(1-\nu).\end{aligned}$$

Dividing by ν and the letting $\nu \to 0$ this yields the claim.

□

Proposition 3.1.11. *(Hilbert spaces). If N is a Hilbert space then for each $q \in \mathcal{P}^1(N)$*

$$b(q) = \int_N zq(dz)$$

in the sense that

$$\langle b(q), y\rangle = \int_N \langle z, y\rangle q(dz) \quad (y \in N).$$

Note that this identity is true for probability measures. Let m be a measure on $(N, \mathcal{B}(N))$ with $0 < m(N) < \infty$. Then the barycenter $b(m)$ of m can be defined as before by

$$b(m) = \arg\min_{t\in N} \int_N [d^2(t,z) - d^2(0,z)]m(dz),$$

which yields

$$b(m) = \frac{1}{m(N)} \int_N zm(dz).$$

Proof. Recall that $b(q)$ is a unique minimizer of

$$F : t \to \int_N [d^2(t,z) - d^2(0,z)]q(dz) = \int_N ||t-z||^2 - ||z||^2 q(dz).$$

Hence, $t = b(q)$ if and only if

$$\frac{d}{d_\epsilon} F(t + \epsilon y)|_{\epsilon=0} = 2\int_N \langle y, t-z\rangle q(dz) = 0,$$

for all $y \in N$. □

Recall that every separable Hilbert space is either isomorphic to some Euclidian space $\mathbb{R}^k$ or to the space l^2. In other words, it is isomorphic to $\bigotimes_{i\in K}\mathbb{R}$ with a finite or countable set K. By the preceding $b(q) = (b(q_i))_{i\in K}$ with $b(q_i) = \int_{\mathbb{R}} zq_i(dz) = \int_N z_i q(dz)$ where z_i and q_i denote the projection of z and q, respectively, onto the i-th factor of N.

Before studying arbitrary trees, we will have a look on spiders. Let K be an arbitrary set and N be the corresponding K-spider. Given $q \in \mathcal{P}^1(N)$ we define numbers

$$\nu_i(q) := \int_{N_i} d(o,z)q(dz), \qquad b_i(q) := \nu_i(q) - \sum_{j\neq i}\nu_j(q)$$

for $i \in K$. (The point $b_i(q)$ is the usual mean value of the image of q in $\mathbb{R}$ in N_i is identified with $\mathbb{R}_+$ and all the other N_j are glued together and identified with $\mathbb{R}_-$.) Note that $b_i(q) > 0$ for at most one $i \in K$.

Proposition 3.1.12. *(Spiders). If $b_i(q) > 0$ for some $i \in K$ then $b(q) = (i, b_i(q))$. Otherwise, $b(q) = o$.*

Proof. Fix q and i. If $b(q) = (i, \nu_0)$ for some $\nu_0 > 0$ then $\nu \to F(\nu)$, where

$$F(\nu) := \int_N d^2((i,\nu),z)q(dz) = \int_{N_i}(\nu - d(o,z))^2 q(dz) + \sum_{j\neq i}\int_{N_j}(\nu + d(o,z))^2 q(dz),$$

attains its minimum on $]0,\infty[$ in $\nu = \nu_0$. The latter implies

$$0 = \frac{1}{2}F'(\nu_0) = \int_{N_i}(\nu_0 - d(0,z))q(dz) + \sum_{j\neq i}\int_{N_j}(\nu_0 + d(0,z))q(dz)$$

$$= \nu_0 - \nu_i(q) + \sum_{j\neq i}\nu_j(q) = \nu_0 - b_i(q),$$

and thus $\nu_0 = b_i(q)$. Similarly, $b(q) = o$ implies $F'(0) \geq 0$ and thus $0 \geq b_i(q)$. □

Remark 19. *The k-spider has the following remarkable property:*

Let $p = \sum_{i=1}^k \alpha_i \cdot p_i \in \mathcal{P}^1(N)$ be a convex combination of $p_i \in \mathcal{P}(N_i), i = 1,\dots,k$, for suitable $\alpha_i > 0$ with $\sum_{i=1}^k \alpha_i = 1$ and put $\bar{p} = \sum_{i=1}^k \alpha_i \cdot \delta_{b(p_i)}$. Then

$$b(p) = b(\bar{p}).$$

Indeed, with the notations from above

$$\nu_i(\bar{p}) = \int_N d(o,z)\bar{p}(dz) = \alpha_i \cdot d(o, b(p_i)) = \alpha_i \int_N d(o,z)p_i(dz) = \nu_i(p),$$

for each i and hence the claim follows. Here the crucial point is that each p_i is supported by a flat space N_i.

3.1.5 Jensen's inequality and L^1 contraction property

Note that throughout this subsection (N,d) will always be a global NPC space.

Proposition 3.1.13. *If a probability measure $q \in \mathcal{P}^1(N)$ is supported by a convex closed set $K \subset N$ then its barycenter $b(q)$ lies in K. In particular, if supp$(q) \subset \bar{B}_r(z)$ then $b(q) \in \bar{B}_r(z)$.*

Proof. Assume that $b(q) \notin K$. Then by Proposition 3.1.5

$$\int [d^2(b(q),z) - d^2(y,z)]q(dz) \geq \int [d^2(\pi_K(b(q)),z) - d^2(y,z)]q(dz)$$

which contradicts the minimizing property of $b(q)$. □

Theorem 3.1.3. *(Jensen's inequality). For any lower semicontinuous convex function $f : N \to \mathbb{R}$ and any $q \in \mathcal{P}^1(N)$*

$$f(b(q)) \leq \int_N f(z)q(dz),$$

provided the RHS is well-defined.

The above RHS is well-defined if either $\int f^+ dq < \infty$ or $\int f^- dq < \infty$. In particular, it is well-defined if f is Lipschitz continuous.

If $\int f dq < \infty$ is well-defined then in Jensen's inequality we may assume without restriction that f is bounded from below and $\int |f| dq < \infty$. Indeed, the assumption implies that $\int f dq = \lim_{k\to\infty} \int f_k dq$ with $f_k := f \vee (-k)$ being bounded from below and convex. Furthermore, $\int f^+ dq = \infty$ would imply $\int f dq = \infty$ in which case Jensen's inequality is trivially true.

Inspired by [163], we will present two entirely different, elementary proofs.

Proof. (First proof following [55]). Given f and q as above, let $\hat{N} = N \times \mathbb{R}$ and $N_f = \{(z,\lambda) \in \hat{N} : f(z) \leq \lambda\}$ which is a closed convex subset of the global NPC space $\hat{N}$.

Put $\hat{f} : N \to \hat{N}$, $z \to (z, f(z))$ and let $\hat{q} = q \circ \hat{f}^{-1}$ be the image of the probability measure q under the map $\hat{f}$. Without restriction, we may assume $\int_N |f(z)|q(dz) < \infty$. Then $\hat{f} \in \mathcal{P}^1(\hat{N})$ since for $\hat{t} = (t, \lambda) \in \hat{N}$

$$\int_N d(\hat{t},\hat{z})\hat{q}(d\hat{z}) \leq \int_N [d(t,z) + |\lambda - f(z)|]q(dz) < \infty.$$

We have that

$$b(\hat{q}) = \Big(b(q), \int_N f(z)q(dz)\Big).$$

Moreover, supp$(\hat{q}) \in N_f$; hence we have that $b(\hat{q}) \in N_f$. That is, $f(b(q)) \leq \int_N f(z)q(dz)$. □

Proof. (Second proof). Now for simplicity assume $q \in \mathcal{P}^2(N)$ and $\int_N f^2(z)q(dz) < \infty$. The general case follows by an approximation argument. Choose a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and an iid sequence $(Y_i)_i$ of a random variables $Y_i : \Omega \to N$ with distribution $\mathbb{P}_{Y_i} = q$. Put

$$Z_i := f(Y_i)$$

$$S_n := \frac{1}{n} \sum_{i=1,\ldots,n}^{\rightarrow} Y_i$$

$$T_n := \frac{1}{n} \sum_{i=1}^{n} Z_i.$$

Then by the weak law of large numbers (for N-valued and for $\mathbb{R}$-valued random variables, respectively)

$$S_n \to \mathbb{E}Y_1 = b(q), \quad T_n \to \mathbb{E}f(Y_1) = \int f dq$$

in probability. Further, we claim that

$$f(S_n) \leq T_n.$$

Indeed, this is true for $n = 1$ and follows for general n by induction:

$$\begin{aligned} f(S_{n+1}) &= f\left(\frac{n}{n+1}S_n + \frac{1}{n+1}Y_{n+1}\right) \\ &\leq \frac{n}{n+1}f(S_n) + \frac{1}{n+1}f(Y_{n+1}) \\ &\leq \frac{n}{n+1}T_n + \frac{1}{n+1}Z_{n+1} = T_{n+1}, \end{aligned}$$

where we only used the convexity of f along geodesics. Therefore, by lower semicontinuity of f

$$f(b(q)) \leq \liminf_{n\to\infty} f(S_n) \leq \liminf_{n\to\infty} T_n = \int f dq.$$

□

Theorem 3.1.4. *(Fundamental Contraction Property), For all $p, q \in \mathcal{P}^1(N)$:*

$$d(b(p), b(q)) \leq d^W(p, q). \tag{3.1.8}$$

Proof. Given $p, q \in \mathcal{P}^1(N)$ consider $\phi \in \mathcal{P}^1(N^2)$ with marginals p and q. Then $b(\phi) = (b(p), b(q))$. Thus Jensen's inequality with the convex function $d : N^2 \to \mathbb{R}$ yields

$$d(b(p), b(q)) = d(b(\phi)) \leq \int_{N^2} d(z)\phi(dz).$$

Therefore, $d(b(p), b(q)) \leq d^W(p, q)$. □

Example 4. *(Barycenter Map of Es-Sahib & Heinich). Let (N, d) be a locally compact, global NPC spaces. Then one can define recursively for each $n \in \mathbb{N}$ a unique map $\beta_n : N^n \to N$ satisfying*

(*i*) $\beta_n(z_1, \ldots, z_1) = z_1$,

(ii) $d(\beta_n(z_1,\ldots,z_n),\beta_n(y_1,\ldots,y_n)) \leq \frac{1}{n}\sum_{i=1}^n d(z_i,y_i)$,

(iii) $\beta_n(z_1,\ldots,z_n) = \beta_n(\check{z}_1,\ldots,\check{z}_n)$ *where* $\check{x}_i := \beta_{n-1}(z_1,\ldots,z_{i-1},z_{i+1},\ldots,z_n)$.

This map is invariant under permutation of coordinates and satisfies

$$d(t,\beta_n(z_1,\ldots,z_n)) \leq \frac{1}{n}\sum_{i=1}^n d(t,z_i) \qquad (t \in N).$$

3.1.6 The integral form of Jensen's inequality

We want to extend Jensen's inequality to the general framework of finite measure spaces. Note that this subsection is inspired from [116].

Remember that a *finite measure space* is any triplet (Ω,Σ,μ) consisting of an abstract nonempty set Ω, a σ-algebra Σ of subsets of Ω and a σ-additive measure $\mu : \Sigma \to \mathbb{R}_+$ such that $\mu(\Omega) > 0$. We can reduce the study of finite measure spaces to that of *probability spaces*, characterized by the fact that $\mu(\Omega) = 1$ by replacing μ with $\mu/\mu(\Omega)$.

Remark 20. *The language of traditional measure theory is slightly different from that of measure-theoretic probability theory. In the books of probability theory we find the notation* (Ω,Σ,P) *for a probability space consisting of a sample space* Ω*, a* σ*-algebra* Σ *of events (viewed as subsets of* Ω*) and a probability measure* $P : \Sigma \to \mathbb{R}_+$*. Note that in the probabilistic context, the real measurable functions* $X : \Omega \to \mathbb{R}$ *are called random variables. Their definitory property is that* $X^{-1}(A) \in \Sigma$ *for every Borel subset* A *of* $\mathbb{R}$*. Remark that, in probability theory, an important point is the use of the notion of independence.*

Jensen's inequality gives us two important concepts that can be attached to a finite measure space (Ω,Σ,μ): the integral arithmetic mean and the barycenter. The *integral arithmetic mean* or the *mean value* of a μ-integrable function $f : \Omega \to \mathbb{R}$ is defined by the formula

$$M_1(f) = \frac{1}{\mu(\Omega)}\int_\Omega f(z)d\mu(z).$$

This number is also called the *expectation* or *expected value* of f and is denoted $E(f)$. The expectation generates a functional $E : L^1(\mu) \to \mathbb{R}$ having the following three properties:

(i) $E(\alpha f + \beta g) = \alpha E(f) + \beta E(g)$, (Linearity);

(ii) $f \geq 0$ implies $E(f) \geq 0$, (Positivity);

(iii) $E(1) = 1$ (Calibration).

We introduce the expectation of a real-valued function $f \in L^1(\mu)$ using the Riemann-Stieltjes integral and the key ingredient is the cumulative distribution function of f, which is defined by the formula

$$F : \mathbb{R} \to [0,1], \quad F(z) = \mu(\{\omega : f(\omega) \leq z\});$$

on several occasions we will use the notation F_f instead of F, to specify the integrable function under attention.

Remark that, this function is increasing and its limits at infinity are $\lim_{z\to-\infty} F(z) = 0$ and $\lim_{z\to\infty} F(z) = 1$. Also, the cumulative distribution function is right continuous, that is,

$$\lim_{z\to z_0^+} F(z) = F(z_0) \text{ at every } z_0 \in \mathbb{R}.$$

The distribution function allows us to introduce the expectation of a random variable by a formula that avoids the use of measure theory:

$$E(f) = \int_{-\infty}^{\infty} z dF_f(z).$$

For convenience, the concept of barycenter will be introduced in the context of probability measures μ defined on the σ-algebra $\mathcal{B}(I)$ of Borel subsets of an interval I. More exactly, we will consider the class

$$\mathcal{P}^1(I) = \{\mu : \mu \text{ Borel probability measure on I and } \int_I |z|d\mu(z) < \infty\}.$$

This class includes all Borel probability measures null outside a bounded subinterval.

Definition 3.1.8. *The barycenter of a Borel probability measure $\mu \in \mathcal{P}^1(I)$ is the real point*

$$bar(\mu) = \int_I z d\mu(z).$$

Necessarily, when the barycenter $bar(\mu)$ exists, it must be in I. Indeed, if $bar(\mu) \notin I$, then either $bar(\mu)$ is an upper bound for I or it is a lower bound. Since

$$\int_I [z - bar(\mu)]d\mu(z) = 0.$$

this situation will impose $z - bar(\mu) = 0$ μ-almost everywhere, which is not possible because $\mu(I) = 1$.

Using Definition 3.1.8 we can present a first example concerns the case of a discrete probability measure $\lambda = \sum_{k=1}^n \lambda_k \delta_{z_k}$ concentrated at the points $z_1, \ldots, z_n \in \mathbb{R}$. Here δ_t represents the Dirac measure concentrated at t, that is, the measure given by

$$\delta_t(A) = 1 \text{ it } t \in A \text{ and } \delta_t(A) = 0 \text{ otherwise.}$$

In this case,

$$E(f) = \int_{\mathbb{R}} f(z)d\lambda(z) = \sum_{k=1}^n \lambda_k f(z_k),$$

for every continuous function $f : \mathbb{R} \to \mathbb{R}$. Hence

$$bar(\lambda) = \int_{\mathbb{R}} z d\lambda(z) = \sum_{k=1}^n \lambda_k z_k.$$

The barycenter of the restriction of Lebesque measure to an interval $[a, b]$ is the middle point because

$$\frac{1}{b-a}\int_a^b z dz = \frac{a+b}{2}.$$

The barycenter of the Gaussian probability measure $\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$ is the origin. Indeed,

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ze^{-z^2/2}dz = 0.$$

Theorem 3.1.5. *(The integral form of Jensen's inequality) Let (Ω, Σ, μ) be a probability space and let $g : \Omega \to \mathbb{R}$ be a μ-integrable function. If f is a convex function defined on an interval I that includes the image of g, then $E(g) \in I$ and*

$$f(E(g)) \leq \int_{\Omega} f(g(z))d\mu(z).$$

Notice that the right hand side integral always exists but it might be ∞ if the μ-integrable of $f \circ g$ is not expressly asked. When both functions g and $f \circ g$ are μ-integrable, then the above inequality becomes

$$f(E(g)) \leq E(f \circ g).$$

Furthermore, if in addition f is strictly convex, then this inequality becomes an equality if and only if g is constant μ-almost everywhere.

Corollary 6. *If $\mu \in \mathcal{P}^1(I)$ and $f : I \to \mathbb{R}$ is a μ-integrable convex function, then*

$$f(bar(\mu)) \leq \int_I f(z)d\mu(z).$$

Corollary 7. *(The Hermite-Hadamard inequality for arbitrary Borel probability measures; [56]) Suppose that $f : [a, b] \to \mathbb{R}$ is a convex function and μ is Borel probability measure on $[a, b]$. Then*

$$f(bar(\mu)) \leq \int_a^b f(z)d\lambda(z) \leq \frac{b - bar(\mu)}{b - a} \cdot f(a) + \frac{bar(\mu) - a}{b - a} \cdot f(b).$$

Remark 21. *(Differential entropy) We assume that (Ω, Σ, P) is a probability space. Let $f : [0, \infty) \to \mathbb{R}$, $f(z) = z \log z$ a convex function. By applying Jensen's inequality, we obtain the following upper estimate for the differential entropy $h(g) = -\int_\Omega g \log g d\mu$ of a positive μ-integrable function g:*

$$h(g) \leq -\left(\int_\Omega g d\mu\right) \log\left(\int_\Omega g d\mu\right).$$

If $\int_\Omega g d\mu = 1$, then $h(g) \leq 0$ even though the function f has a variable sign, it attains the minimum value $-\frac{1}{e}$ at $t = \frac{1}{e}$.

A natural question is how large is the Jensen gap

$$E(f \circ g) - f(E(g)).$$

The following result noticed by O. Hölder [73] provides the answear in an important spacial case.

Proposition 3.1.14. *Suppose that $f : [a, b] \to \mathbb{R}$ is a twice differentiable function for which there exist real constants m and M such that $m \leq f'' \leq M$. Then*

$$\begin{aligned}\frac{m}{2}\sum_{1\leq i<j\leq n} \lambda_i\lambda_j(z_i - z_j)^2 &\leq \sum_{k=1}^{n} \lambda_k f(z_k) - f\left(\sum_{k=1}^{n} \lambda_k z_k\right) \\ &\leq \frac{M}{2}\sum_{1\leq i<j\leq n} \lambda_i\lambda_j(z_i - z_j)^2,\end{aligned}$$

whenever $z_1, \ldots, z_n \in [a, b]$, $\lambda_1, \ldots, \lambda_n \in [0, 1]$ and $\sum_{k=1}^n \lambda_k = 1$.

Proof. This is a consequence of the discrete Jensen's inequality when applied to the convex functions $f - mz^2/2$ and $Mz^2/2 - f$. □

Corollary 8. *(The gap in the AM-GM inequality) If $0 < m \leq z_1, \ldots, z_n \leq M$, $\lambda_1, \ldots, \lambda_n \in [0,1]$ and $\sum_{k=1}^{n} \lambda_k = 1$, then*

$$\begin{aligned}\frac{m}{2n^2} \sum_{1\leq j<k\leq n} (\log z_j - \log z_k)^2 &\leq \frac{1}{n}\sum_{k=1}^{n} z_k - \left(\prod_{k=1}^{n} z_k\right)^{1/n} \\ &\leq \frac{M}{2n^2} \sum_{1\leq j<k\leq n} (\log z_j - \log z_k)^2.\end{aligned}$$

Proposition 3.1.14 exhibits the role of the variance in estimating the precision in Jensen's inequality.

If (Ω, Σ, μ) is a probability space, the *variance* of a function $g \in L^2(\mu)$ is defined by the formula

$$\operatorname{var}(g) = E((g - E(g))^2) = E(g^2) - (E(g))^2.$$

The variance is an indicator of how much the values of g are spread out. A variance of zero indicates that all the values of g are identical, except possibly for a subset of Ω of probability zero.

The square root of the variance is called the *standard deviation*.

Since a probability measure is a finite measure, the space $L^2(\mu)$ is included in $L^1(\mu)$. Thus expectation and variance applies to every function that belongs to $L^2(\mu)$.

Let us consider the probability space $\Omega = \{1, \ldots, n\}$, $\Sigma = \mathcal{P}(\Omega)$ and $\mu = \sum_{k=1}^{n} \lambda_k \delta_k$. The variance of the function $g : \{1, \ldots, n\} \to \mathbb{R}$ defined by $g(k) = z_k$ for $k = 1, \ldots, n$, is

$$\begin{aligned}\operatorname{var}(g) &= E((g - E(g))^2) \\ &= E(g^2) - (E(g))^2 \\ &= \sum_{n=1}^{n} \lambda_n z_n^2 - \left(\sum_{n=1}^{n} \lambda_n z_n\right)^2 \\ &= \frac{1}{2} \sum_{1\leq i,j\leq n} \lambda_i \lambda_j (z_i - z_j)^2 = \sum_{1\leq i<j\leq n} \lambda_i \lambda_j (z_i - z_j)^2,\end{aligned}$$

and thus the result of Proposition 3.1.14 can be reformulated as

$$\frac{m}{2}\operatorname{var}(g) \leq E(f(g)) - f(E(g)) \leq \frac{M}{2}\operatorname{var}(g).$$

Notice that this double estimate works in general. In probability theory and statistics an important role is played by the so called continuous random variables. A random variable X attached to a probability measure space (Ω, Σ, P) is called continuous if its cumulative distribution is of the form

$$F_X(z) = P(\{\omega : X(\omega) < z\}) = \int_{-\infty}^{z} w(u)du,$$

for a suitable Lebesque integrable function $w \in L^1(\mathbb{R})$, called the density of F_X. In this case, the probability that X takes a value α is 0 and

$$P(\{\omega : a \leq X(\omega) \leq b\}) = \int_a^b w(u)du \text{ for all } -\infty \leq a \leq b \leq \infty.$$

Furthermore, the computation of the expectation and of the variance of X reduces to the computation of certain Lebesque integrals:

$$E(X) = \int_{-\infty}^{\infty} zw(z)dz \text{ and } \operatorname{var}(X) = \int_{-\infty}^{\infty} (z - E(X))^2 w(z)dz.$$

A continuous random variable X is called *normal* if its distribution function is associated to a density of the form

$$w(u, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-\mu)^2}{2\sigma^2}}.$$

In this case, the values of the parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ are precisely the expectation and the standard derivation of X.

Remark 22. *(Upper bounds on the variance) As was noticed by D. S. Mitrinović, J. E. Pečarić and A. M. Fink in [107], p. 296, if X is a random variable such that $\alpha \leq X \leq \beta$ for two suitable constants α and β, then*

$$var(X) \leq (\beta - E(X))(E(X) - \alpha).$$

This remark improves the upper bound previously indicated by T. Popoviciu [145]

$$var(X) \leq \frac{(\beta - \alpha)^2}{4}.$$

The bound found by Mitrinović, Pečarić and Fink follows easily from the general properties of expectation:

$$\begin{aligned} 0 \leq E((\beta - X)(X - \alpha)) &= -\alpha\beta + (\alpha + \beta)E(X) - E(X^2) \\ &= (\beta - E(X))(E(X) - \alpha) - var(X). \end{aligned}$$

Remark 23. *(Chebyshev's probabilistic inequality) If X is a random variable associated to a probability measure space (Ω, Σ, μ), then*

$$\mu(\{|X - E(X)| \geq \epsilon\}) \leq \frac{var(X)}{\epsilon^2},$$

for all $\epsilon > 0$.

The covariance of two real random variables $X, Y \in L^2(P)$ is defined by

$$\begin{aligned} \operatorname{cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Two random variables X and Y whose covariance is zero are called *uncorrelated.* If X and Y are independent, then their covariance is zero. This follows because under independence, $E(XY) = E(X)E(Y)$. However, uncorrelation does not imply in general independence.

The concept of covariance allows us to indicate a new upper estimate of the Jensen gap.

Theorem 3.1.6. *(The covariance form of Jensen's inequality) Let* (Ω, Σ, μ) *be a probability space and let* $g : \Omega \to \mathbb{R}$ *be a* μ*-integrable function. Suppose also that* f *is a convex function defined on an interval* I *that includes the image of* g *and* $\eta : I \to \mathbb{R}$ *is a function such that*

(i) $\eta(z) \in \partial f(z)$ *for every* $z \in I$*;*

(ii) $\eta \circ g$ *and* $g \cdot (\eta \circ g)$ *are* μ*-integrable functions.*

Then

$$0 \leq E(f \circ g) - f(E(g)) \leq cov(g, \eta \circ g).$$

If f *is concave, then the last two inequalities work in the reserved direction.*

Proof. The first inequality is motivated by Theorem 3.1.5. The second inequality follows by integrating the inequality

$$f(E(g)) \geq f(g(z)) + (E(g) - g(z)) \cdot \eta(g(z)) \text{ for all } z \in \Omega.$$

□

Corollary 9. *(S. S. Dragomir and N. M. Ionescu [51]) If* f *is a differentiable convex function defined on an open interval* I*, then*

$$\begin{aligned} 0 &\leq \sum_{k=1}^{n} \lambda_k f(z_k) - f\left(\sum_{k=1}^{n} \lambda_k z_k\right) \\ &\leq \sum_{k=1}^{n} \lambda_k z_k f'(z_k) - \left(\sum_{k=1}^{n} \lambda_k z_k\right)\left(\sum_{k=1}^{n} \lambda_k f'(z_k)\right) \end{aligned}$$

for all $z_1, \ldots, z_n \in I$ *and all* $\lambda_1, \ldots, \lambda_n \in [0, 1]$*, with* $\sum_{k=1}^{n} \lambda_k = 1$*.*

The covariance defines a Hermitian product and it differs from a scalar product by the fact that $\mathrm{cov}(X, X) = 0$ implies only that X is constant almost everywhere. Since the Cauchy-Bunyakovsky-Schwarz inequality still works for such products, we have the inequality

$$|\mathrm{cov}(X, Y)| \leq (\mathrm{var}(X))^{1/2} (\mathrm{var}(Y))^{1/2},$$

knows as the *covariance form of Cauchy-Bunyakovsky-Schwarz inequality.* This inequality shows that Pearson's correlation coefficient of X and Y, that is,

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{(\mathrm{var}(X))^{1/2} (\mathrm{var}(Y))^{1/2}}$$

takes values in the interval $[-1, 1]$. A value of 1, respectively -1 for $\rho_{X,Y}$ implies that the relationship between X and Y is described by a linear equation, for which Y increases, respectively decreases as X increases. A value of 0 implies that there is no linear correlation between the two random variables. From the covariance form of Cauchy-Bunyakovsky-Schwartz inequality and Remark 23 we infer the following classical result:

Theorem 3.1.7. *(Grüss Inequality [65]) Suppose that the random variables* X *and* Y *are bounded, precisely,* $\alpha \leq X \leq \beta$ *and* $\delta \leq Y \leq \gamma$*. Then*

$$|cov(X, Y)| \leq \frac{1}{4}(\beta - \alpha)(\gamma - \delta).$$

and the constant $\frac{1}{4}$ *being sharp.*

3.1.7 Convex sets in real linear spaces

In this subsection we review some basic facts, necessary for a deep understanding of the concept of convexity in real linear spaces. The natural domain for a convex function is a convex set.

A subset C of a linear space E is said to be *convex* if it contains the line segment

$$[x, y] = \{(1-\lambda)x + \lambda y : \lambda \in [0, 1]\},$$

connecting any of its points x and y. Convexity is a weak form of rotundity. Besides line segments, some others simple examples of convex sets in the Euclidian space $\mathbb{R}$ are the lines, the planes, the open disc, plus any part of their boundary and the N-dimensional rectangles (Cartesian products of N nonempty intervals).

New examples from the old ones can be obtained by considering arbitrary intersections and/or the following two algebraic operations with sets:

$$A + B = \{x + y : x \in A, y \in B\},$$

$$\lambda A = \{\lambda x : x \in A\},$$

for $A, B \subset E$ and $\lambda \in \mathbb{R}$. The addition of sets is also known as the *Minkowski addition*. Addition of sets is commutative and associative. One can prove easily that $\lambda A + \mu B$ is a convex set provided that A and B are convex and $\lambda, \mu \geq 0$.

A subset A of E is said to be affine if it contains the whole line through any two of its points. Algebraically, this means that

$$x, y \in A \text{ and } \lambda \in \mathbb{R} \text{ imply } (1-\lambda)x + \lambda y \in A.$$

Cleary, any affine subset is also convex, but the converse is not true. It is important to notice that any affine subset A is just the translate of a unique linear subspace L and all translations of a linear space represent affine sets. In fact, for every $a \in A$, the translate

$$L = A - a$$

is a linear space and it is clear that $A = L + a$. For the uniqueness part, notice that if L and M are linear subspaces of E and $a, b \in E$ verify

$$L + a = M + b,$$

then necessarily $L = M$ and $a - b \in L$. This remark allows us to introduce the concept of dimension for an affine set (as the dimension of the linear subspace of which it is a translate). Given a finite family $x_1, \ldots, x_n$ of points in E, an affine combination of them is any point of the form

$$x = \sum_{k=1}^{n} \lambda_k x_k,$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and $\sum_{k=1}^{n} \lambda_k = 1$. If in addition $\lambda_1, \ldots, \lambda_n \geq 0$, then x is called a *convex combination* of $x_1, \ldots, x_n$.

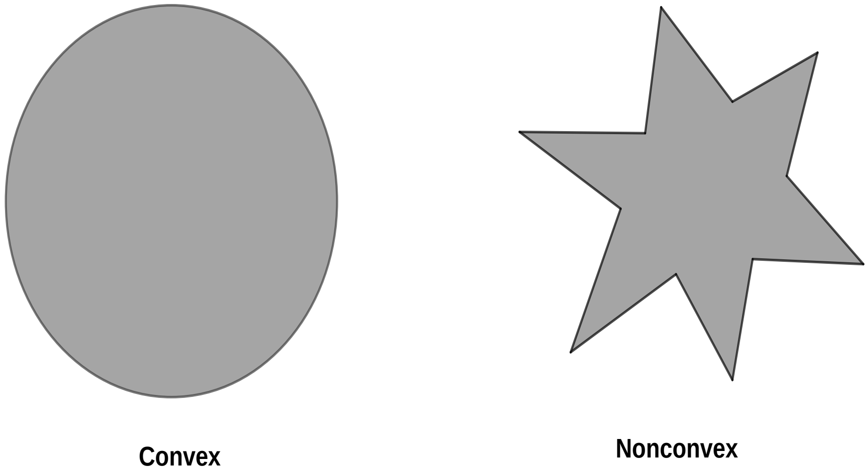

Figure 3.5: Convex and nonconvex planar sets

Lemma 3.1.2. *A subset C of E is convex (respectively affine) if and only if it contains every convex (respectively affine) combination of points of C.*

Proof. The sufficiency part is clear, while the necessity part can be proved by mathematical induction. $\square$

Given a subset A of E, the intersection $\mathrm{conv}(A)$ of all convex subsets of E containing A is convex and thus it is the smallest set of this nature containing A. We call it the *convex hull* of A. By using Lemma 3.1.2, one can verify easily that $\mathrm{conv}(A)$ consists of all convex combinations of elements of A.

The affine variant of this construction yields the *affine hull* of A, denoted $\mathrm{aff}(A)$. As a consequence we can introduce the concept of dimension for convex sets to be the dimension of their affine hulls.

A nice example of convex hull is offered by the Gauss-Lucas theorem on the distribution of the critical points of a polynomial: the roots $(\mu_k)_{k=1}^{n-1}$ of the derivative P' of any complex polynomial P of degree $n \geq 2$ lie in the smallest convex polygon containing the roots $(\lambda_i)_{j=1}^{n}$ of the polynomial P. Indeed assuming that w is a root of P' and $P(w) \neq 0$, we have

$$0 = \frac{P'(w)}{P(w)} = \sum_{k=1}^{n} \frac{1}{w - \lambda_k} = \sum_{k=1}^{n} \frac{\overline{w} - \overline{\lambda}_k}{|w - \lambda_k|^2},$$

whence

$$w = \sum_{k=1}^{n} \frac{1}{|w - \lambda_k|^2} \lambda_k \Bigg/ \sum_{k=1}^{n} \frac{1}{|w - \lambda_k|^2}.$$

If $(S_i)_{i \in \mathcal{I}}$ is a finite family of subsets of an N-dimensional linear space E, then the convex hull of their

Minkovski addition equals the Minkovski addition of their convex hulls:

$$\operatorname{conv}\Bigg(\sum_{i\in\mathcal{I}} S_i\Bigg) = \sum_{i\in\mathcal{I}} \operatorname{conv}(S_i). \tag{3.1.9}$$

Surprisingly, this simple remark has deep consequently to the geometry of convex sets. The clue is provided by the following unifying lemma, used by R. M. Anderson in his course on Economic Theory, taught in Spring 2010 at Berkeley.

Lemma 3.1.3. *Consider a finite family $(S_i)_{i\in\mathcal{I}}$ of nonempty subsets of $\mathbb{R}^N$. Then every $x \in$ conv$\Big(\sum_{i\in\mathcal{I}} S_i\Big)$ admits a representation of the form*

$$x = \sum_{i\in\mathcal{I}} \Bigg(\sum_{1\le j\le n_i} \lambda_{ij} x_{ij}\Bigg)$$

such that

(i) $\sum_{i\in\mathcal{I}} \le |\mathcal{I}| + N$;

(ii) $x_{ij} \in S_i$ *and* $\lambda_{ij} > 0$ *for all* i, j;

(iii) $\sum_{j=1}^{n_i} \lambda_{ij} = 1$ *for all* $i \in \mathcal{I}$.

Proof. According to formula (3.1.9), every point $x \in \operatorname{conv}\Big(\sum_{i\in\mathcal{I}} S_i\Big)$ admits a representation of the form $x = \sum_{i\in\mathcal{I}} x_i$ with $x_i \in \operatorname{conv}(S_i)$ for all i. Therefore,

$$x = \sum_{i\in\mathcal{I}} \Bigg(\sum_{1\le j\le n_i} \lambda_{ij} x_{ij}\Bigg), \tag{3.1.10}$$

for suitable $x_{ij} \in S_i$ and $\lambda_{ij} > 0$ with $\sum_{j=1}^{n_i} = 1$. Clearly, one can choose such a representation for which $n = \sum_{i\in\mathcal{I}} n_i$ is minimal. If $n > |\mathcal{I}| + N$, then the vectors $x_{ij} - x_{i1}$ for $i \in \mathcal{I}$ and $j \in [2, n_i]$ are linearly dependent in $\mathbb{R}^N$. Then

$$\sum_{i\in\mathcal{I}} \sum_{2\le j\le n_i} c_{ij}(x_{ij} - x_{i1}) = 0,$$

for some real coefficients, not all zero. Adding to equation (3.1.10) the last equation multiplied by a real number λ we obtain

$$x = \sum_{i\in\mathcal{I}} \sum_{1\le j\le n_i} \tilde{\lambda}_{ij} x_{ij}, \tag{3.1.11}$$

where

$$\tilde{\lambda}_{ij} = \lambda_{ij} + \lambda c_{ij} \qquad \text{if } j \ge 2,$$
$$\tilde{\lambda}_{i1} = \lambda_{i1} - \lambda \sum_{2\le k\le n_i} c_{ik} \qquad \text{if } j = 1.$$

By a suitable choice of λ one can ensure that $\tilde{\lambda}_{ij} \ge 0$ for all indices i,j and that all least one coefficient $\tilde{\lambda}_{ij}$ is zero. The representation (3.1.11) eliminates one of x_{ij}, contrary to the minimality of n. Consequently, $n \le |\mathcal{I}| + N$ and the proof is complete. $\square$

Theorem 3.1.8. *(Carathéodoory's theorem) Suppose that S is a subset of a linear space E and its convex hull $conv(S)$ has dimension m. Then each point x of $conv(S)$ is the convex combination of at most $m+1$ points of S.*

Proof. Clearly, we may assume that $E = \mathbb{R}^m$. Then apply Lemma 3.1.3 for $\mathcal{I} = \{1\}$ and $S_1 = S$. $\square$

The sets of the form $C = \mathrm{conv}(\{x_0, \ldots, x_n\})$ are usually called *polytopes*. If $x_1 - x_0, \ldots, x_n - x_0$ are linearly independent, then C is called an n-simplex with vertices $x_0, \ldots, x_n$. In this case, $dimC = n$ and every point x of C has a unique representation $x = \sum_{k=0}^{n} \lambda_k x_k$, as a convex combination of vertices; the numbers $\lambda_0, \ldots, \lambda_n$ are called the barycentric coordinates of x.

The standard n-simplex or unit n-simplex is the simplex Δ^n whose vertices are the elements of the canonical algebraic basis of $\mathbb{R}^{n+1}$, that is,

$$\Delta^n = \left\{ (\lambda_0, \ldots, \lambda_n) \in \mathbb{R}^{n+1} : \sum_{k=0}^{n} \lambda_k = 1 \text{ and } \lambda_k \geq 0 \text{ for all } k \right\}.$$

Given an arbitrary n-simplex C with vertices $(x_0, \ldots, x_n)$, the map

$$\omega : \Delta^n \to C, \quad \omega(\lambda_0, \ldots, \lambda_n) = \sum_{k=0}^{n} \lambda_k x_k$$

is affine and bijective. Notice that any polytope $\mathrm{conv}(\{x_0, \ldots, x_n\})$ is a union of simplices whose vertices belong to $\{x_0, \ldots, x_n\}$.

Remark 24. *(Lagrange's barycenter identity) Consider a finite system $\mathcal{S}$ of mass points (x_k, m_k) in $\mathbb{R}^{n+1}$, for $k = 1, \ldots, n$; x_k indicates position and m_k the mass. In mechanics and physics, one defines the barycenter or center of mass of the system by*

$$bar(\mathcal{S}) = \frac{\sum_{k=1}^{n} m_k x_k}{\sum_{k=1}^{n} m_k}.$$

The mass point $(bar\mathcal{S}, \sum_{k=1}^{n} m_k)$ represents the resultant of the system $\mathcal{S}$. Notice that $bar\mathcal{S} \in conv(\{x_0, \ldots, x_n\})$. A practical way to determine the barycenter was found by J. K. Lagrange [91], who proved the following identity: For every family of points $x, x_1, \ldots, x_n$ in $\mathbb{R}^n$ and every family of real weights $m_1, \ldots, m_n$ with $M = \sum_{k=1}^{n} m_k > 0$, we have

$$\sum_{k=1}^{n} m_k ||x - x_k||^2 = M \left|\left| x - \frac{1}{M} \sum_{k=1}^{n} m_k x_k \right|\right|^2 + \frac{1}{M} \cdot \sum_{1 \leq i < j \leq n} m_i m_j ||x_i - x_j||^2.$$

For the proof, use the formula $||z||^2 = \langle z, z \rangle$. The previous formula yields the following variational definition of barycenter: $bar(\mathcal{S})$ is the unique point that minimizes the function $x \to \sum_{k=1}^{n} m_k ||x - x_k||^2$, that is,

$$bar(\mathcal{S}) = \arg\min_{x \in \mathbb{R}^n} \sum_{k=1}^{n} m_k ||x - x_k||^2.$$

3.2 On the barycenter for discrete Steffensen Popoviciu measures in global NPC spaces

In this section we put in a new light the concept of barycenter for discrete Steffensen Popoviciu measures supported in some points belonging to a space with curved geometry. More precisely, we ensure the existence of the barycenter if we relax the restrictions imposed to the weights of the measure. As applications, even in the case of nonpositive weights we deduce Jensen-Steffensen's, Hardy-Littlewood-Polya's and Sherman's type inequalities on global NPC spaces.

In the last decades numerous authors performed an intense research activity to extend majorization theory beyond classical case of probability measures, i.e. Steffensen Popoviciu measures. The main point of interest into this topic of research is to offer a large framework under which Jensen's type inequalities works. Jensen Steffensen's inequality (see [116, Theorem 2.4.4]) reveals an important case when Jensen's inequality works beyond the framework of positive measures. In fact, this is our aim, to relax the concept of barycenter in spaces with curved geometry, in order to provide more insight into the relation between signed measures and Jensen's type inequalities.

Based on the fact that most of weighted inequalities from the theory of convex functions are dealing with positive weights we consider here the challenging case of nonpositive weights. In this context, we recall the so called Jensen Steffensen inequality (we refer, for instance, to [115]).

Theorem 3.2.1. *Let $x_n \leq x_{n-1} \leq \cdots \leq x_1$ in an interval $[a, b]$ and let $p_1, \ldots, p_n$ be some real numbers such that the partial sums $S_k = \sum_{i=1}^{k} p_i$ verify the relations*

$$0 \leq S_k \leq S_n \quad and \quad S_n > 0.$$

Then, for every convex functions $f : [a, b] \to \mathbb{R}$ we have the inequality

$$f\left(\frac{1}{S_n}\sum_{k=1}^{n} p_k x_k\right) \leq \frac{1}{S_n}\sum_{k=1}^{n} p_k f(x_k).$$

In fact, the above result is related to the general concept of Steffensen Popoviciu's measure, as it is presented in [114, 115, 116].

Definition 3.2.1. *Let K be a compact convex subset of a real locally convex Hausdorff space E. A Steffensen Popoviciu measure on K is any real Borel measure μ on K such that $\mu(K) > 0$ and*

$$\int_K f(x)\, d\,\mu(x) \geq 0,$$

for every positive, continuous and convex function $f : K \to \mathbb{R}$.

The characterization of discrete Steffensen Popoviciu's measures is presented in [116, Corollary 9.14].

Proposition 3.2.1. *Suppose that $x_1 \leq \cdots \leq x_n$ are real points and $p_1, \ldots, p_n$ are real weights. Then, the discrete measure $\mu = \sum_{k=1}^{n} p_k \delta_{x_k}$ is a Steffensen Popoviciu measure if*

$$\sum_{k=1}^{n} p_k > 0 \;\; and \;\; 0 \leq \sum_{k=1}^{m} p_k \leq \sum_{k=1}^{n} p_k \quad (m \in \{1, \ldots, n\}).$$

The concept of barycenter for Steffensen Popoviciu measures was fully discussed in [116, Lemma 9.2.3 and Theorem 9.2.4]. But, our aim is to give a new perspective to the barycenter concept on more general spaces, namely global NPC spaces, via the majorization techniques.

It is worth noticing that the concept of weighted majorization in $\mathbb{R}^N$ is related to an optimization problem. Indeed, we have

$$x_i = \arg\min_{z\in\mathbb{R}^N} \frac{1}{2}\sum_{j=1}^{n} a_{ij}\left\| z - y_j \right\|^2, \quad \text{for } i = 1, \ldots, m.$$

In what follows we shall deal with the relation of weighted majorization $\prec$, for pairs of discrete probability measures. In the context of Euclidean space $\mathbb{R}^N$,

$$\sum_{i=1}^{l} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{m} \mu_j \delta_{y_j} \tag{3.2.1}$$

means the existence of a $m \times l$-dimensional matrix $A = (a_{ij})_{i,j}$ such that the following four conditions are fulfilled:

$$a_{ij} \geq 0, \text{ for all } i, j, \tag{3.2.2}$$

$$\sum_{j=1}^{m} a_{ji} = 1, \quad i = 1, \ldots, l, \tag{3.2.3}$$

$$\mu_j = \sum_{i=1}^{l} a_{ji}\lambda_i, \quad j = 1, \ldots, m, \tag{3.2.4}$$

and

$$x_i = \sum_{j=1}^{m} a_{ji} y_j, \quad i = 1, \ldots, l. \tag{3.2.5}$$

See Borcea [27] and Marshal, Olkin and Arnold [104]. The matrices verifying the conditions (3.2.2) and (3.2.3) are called *stochastic on rows*. When $l = m$ and all weights λ_i and μ_j are equal to $1/m$, the condition (3.2.4) assures the *stochasticity on columns*, so in that case we deal with doubly stochastic matrices.

Under the above settings, S. Sherman [160] use the concept of weighted majorization and proved that, the following inequality

$$\sum_{i=1}^{l} \lambda_i f(x_i) \leq \sum_{j=1}^{m} \mu_j f(y_j)$$

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

The aim of this section from the thesis is to extend the above majorization theory and the classical inequalities for Steffensen Popoviciu measures, in spaces with curved geometry. More precisely, our scope is to extend Theorem 3.2.1 in the framework of global NPC spaces and then to derive HLP's, Sherman's and Jensen Steffensen's type inequalities.

Subsection 3.1.1 is devoted to the concept of barycenter and Jensen Steffensen's inequalities in the framework of global NPC spaces; in subsection 3.1.2 we present some applications related to HLP's, Sherman's and Jensen Steffensen's type inequalities for the weighted majorization of discrete measures (with nonpositive weights); Subsection 3.1.3 present some conclusions, extensions and further applications related to relative convexity in global NPC spaces.

3.2.1 Jensen Steffensen's type inequalities on global NPC spaces

Inspired from [94], we present an extension of barycenter for Steffensen Popoviciu discrete measures, where the most important ingredient in NPC spaces is the *barycenter* of a discrete probability measures $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$. Thus, in what follows we relax the concept of barycenter by considering nonpositive weights for the discrete measures.

Definition 3.2.2. *Let $X := \{x_1, \ldots, x_n\}$ be a family of points in a global NPC space M, all these points belonging to the same geodesic $[x_1, x_n]$ and, in addition the following assumptions are verified*

$$x_i \in [x_{i-1}, x_{i+1}] \qquad (i \in \{2, 3, \ldots, n-1\}).$$

For any family of real weights $\Lambda := \{\lambda_1, \ldots, \lambda_n\}$ which verify

$$0 \le S_i \le S_n = 1 \qquad (i \in \{1, 2, \ldots, n\}),$$

where

$$S_k = \lambda_1 + \cdots + \lambda_k \qquad (k \in \{1, 2, \ldots, n\}),$$

we define the notion of weak barycenter of the family of points X with respect to the family of real weights Λ as the unique point $\bar{\mathbf{x}}$ on the geodesic $[x_1, x_n]$ satisfying

$$d(\bar{\mathbf{x}}, x_1) = \bar{S}_2 d(x_2, x_1) + \bar{S}_3 d(x_3, x_2) + \cdots + \bar{S}_n d(x_n, x_{n-1}), \tag{3.2.6}$$

or, equivalenty,

$$d(x_n, \bar{\mathbf{x}}) = S_1 d(x_2, x_1) + S_2 d(x_3, x_2) + \cdots + S_{n-1} d(x_n, x_{n-1}), \tag{3.2.7}$$

where

$$\bar{S}_k = \lambda_k + \cdots + \lambda_n \qquad (k \in \{1, 2, \ldots, n\}).$$

Remark 25. *Note that, the weak barycenter $\bar{\mathbf{x}}$ from* (3.2.10) *and* (3.2.11) *is well defined and we have that*

$$d(\bar{\mathbf{x}}, x_1) + d(x_n, \bar{\mathbf{x}}) = d(x_2, x_1) + d(x_3, x_2) + \cdots + d(x_n, x_{n-1}) = d(x_n, x_1),$$

which confirm the fact that $\bar{\mathbf{x}}$ lies on the geodesic $[x_1, x_n]$. Moreover, using (3.2.10) *or* (3.2.11) *in flat spaces some computations goes back to the following classical formula*

$$\bar{\mathbf{x}} = \lambda_1 x_1 + \cdots + \lambda_n x_n.$$

We are now in position to present a completely new proof of Jensen-Steffensen's inequality in the most relevant case, where we have considered the maximum possible number of nonpositive weights.

Theorem 3.2.2. (The discrete form of Jensen-Steffensen's Inequality) *Let X and Λ be given as in Definition 3.2.2, but with nonpositive weights $\lambda_2, \lambda_3, \ldots, \lambda_{n-1} \le 0$.*

Then, for every continuous convex function $f : M \to \mathbb{R}$ we have the inequality

$$f(\bar{\mathbf{x}}) \le \sum_{i=1}^{n} \lambda_i f(x_i).$$

Proof. Taking into acount (3.2.10) and (3.2.11) a moment of reflections shows that

$$\bar{\mathbf{x}} = \operatorname{bar}((1-t_0)\delta_{x_1} + t_0\delta_{x_n}),$$

where

$$t_0 = \frac{(\lambda_2 + \cdots + \lambda_n)\, d(x_2, x_1) + (\lambda_3 + \cdots + \lambda_n)\, d(x_3, x_2) + \cdots + \lambda_n d(x_n, x_{n-1})}{d(x_n, x_1)}.$$

Hence, from convexity property along geodesic of the function f we get

$$f(\bar{\mathbf{x}}) \leq \frac{\lambda_1 d(x_2, x_1) + (\lambda_1 + \lambda_2)\, d(x_3, x_2) + \ldots (\lambda_1 + \cdots + \lambda_{n-1})\, d(x_n, x_{n-1})}{d(x_n, x_1)} f\,(x_1)$$

$$+ \frac{(\lambda_2 + \cdots + \lambda_n)\, d(x_2, x_1) + (\lambda_3 + \cdots + \lambda_n)\, d(x_3, x_2) + \cdots + \lambda_n d(x_n, x_{n-1})}{d(x_n, x_1)} f\,(x_n)$$

$$= \lambda_1 f\,(x_1) + \lambda_n f\,(x_n) + \lambda_2 \frac{f\,(x_1)\, d(x_n, x_2) + f\,(x_n)\, d(x_2, x_1)}{d(x_n, x_1)}$$

$$+\lambda_3 \frac{f\,(x_1)\, d(x_n, x_3) + f\,(x_n)\, d(x_3, x_1)}{d(x_n, x_1)}$$

$$\cdots$$

$$+\lambda_{n-1} \frac{f\,(x_1)\, d(x_n, x_{n-1}) + f\,(x_n)\, d(x_{n-1}, x_1)}{d(x_n, x_1)}.$$

Using again the convexity of f and the fact that

$$\operatorname{bar}\left(\frac{d(x_n, x_i)}{d(x_n, x_1)}\delta_{x_1} + \frac{d(x_i, x_1)}{d(x_n, x_1)}\delta_{x_n}\right) = x_i \qquad (i = 2, \ldots, n-1),$$

we deduce that the following inequalities hold true

$$f(x_2) \leq \frac{f\,(x_1)\, d(x_n, x_2) + f\,(x_n)\, d(x_2, x_1)}{d(x_n, x_1)},$$

$$f(x_3) \leq \frac{f\,(x_1)\, d(x_n, x_3) + f\,(x_n)\, d(x_3, x_1)}{d(x_n, x_1)},$$

$$\cdots$$

$$f(x_{n-1}) \leq \frac{f\,(x_1)\, d(x_n, x_{n-1}) + f\,(x_n)\, d(x_{n-1}, x_1)}{d(x_n, x_1)}.$$

Finally, since $\lambda_2, \lambda_3, \ldots, \lambda_{n-1}$ are nonpositive we get the desired conclusion.

□

3.2.2 Applications to HLP's and Sherman's type inequalities with nonpositive weights

In order to obtain Sherman's type inequalities with nonpositive weights we firstly introduce the relaxed concept of majorization between two n-tuples of points in a global NPC space (M, d).

Definition 3.2.3. *Let* $\mathbf{x} = (x_1, \ldots, x_n) \in M^n$, $\mathbf{y} = (y_1, \ldots, y_n) \in M^n$, $n \geq 2$.

We define the concept of majorization $\mathbf{x} \prec \mathbf{y}$ *by asking the existence of a matrix* $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$ *such that*

- $\alpha_{ji} \leq 0 \qquad (i \neq 1$ *or* $i \neq l)$;
- $y_i \in [y_{i-1}, y_{i+1}] \qquad (i \in \{2, 3, \ldots, l-1\})$ *verify that all these points belong to the same geodesic* $[y_1, y_m]$;
- x_i *is the weak barycenter of the family of points* $X := \{y_1, \ldots, y_m\}$ *with respect to the family of real weights* Λ^j *as the unique point* x_i *on the geodesic* $[y_1, y_m]$ *satisfying*

$$d(x_i, y_1) = \bar{S}_2^j d(y_2, y_1) + \bar{S}_3^j d(y_3, y_2) + \cdots + \bar{S}_n^j d(y_n, y_{n-1}), \tag{3.2.8}$$

or, equivalenty,

$$d(y_n, x_i) = S_1^j d(y_2, y_1) + S_2^j d(y_3, y_2) + \cdots + S_{n-1}^j d(y_n, y_{n-1}), \tag{3.2.9}$$

where

$$\begin{gathered}
\Lambda^j := \{\alpha_{1j}, \ldots, \alpha_{nj}\} \qquad (j \in \{1, \ldots, m\}), \\
\bar{S}_k^j = \alpha_{kj} + \cdots + \alpha_{nj} \qquad (k \in \{1, 2, \ldots, l\}), \\
S_k^j = \alpha_{1j} + \cdots + \alpha_{kj} \qquad (k \in \{1, 2, \ldots, l\}), \\
0 \leq S_k^j \leq S_n^j = 1 \qquad (k \in \{1, 2, \ldots, l\}).
\end{gathered}$$

We can present now the extension of HLP's inequality in a global NPC space (M, d), when the weights are allowed to be nonpositive.

Theorem 3.2.3. *In the hypotheses from Definition 3.2.3 let us suppose that conditions* (3.2.8) *are satisfied. Then, the following inequality*

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

holds for every convex function $f : M \to \mathbb{R}$.

Proof. Since the hypotheses of Theorem 3.2.2 are satisfied for each $\bar{\mathbf{x}} = x_i$, we get

$$f(x_i) \leq \sum_{j=1}^{m} \alpha_{ij} f(y_j) \qquad (i = 1, \ldots, l).$$

Taking into account (3.2.8) – (3.2.9) and applying Theorem 3.2.2 for each y_i, $i = 1, \ldots, m$, where $y_i = \sum_{j=1}^{l} p_j x_j$, $p_j = \alpha_{ij}$, we get

$$\sum_{i=1}^{m} a_i f(x_i) = \sum_{i=1}^{m} a_i f\left(\sum_{j=1}^{m} y_j \alpha_{ji}\right)$$

$$\leq \sum_{i=1}^{l} a_i \left(\sum_{j=1}^{m} \alpha_{ji} f\left(y_j\right) \right) = \sum_{j=1}^{m} f\left(y_j\right) \sum_{i=1}^{l} a_i \alpha_{ji}.$$

Consequently, since $b_j = \sum_{i=1}^{l} a_i \alpha_{ji}$ we have

$$\sum_{i=1}^{l} a_i f\left(x_i\right) \leq \sum_{j=1}^{m} b_j f\left(y_j\right).$$

$\square$

We are in position to introduce the relaxed weighted concept of majorization between two n-tuples of points in a global NPC space M.

Definition 3.2.4. *Let* $\mathbf{x} = (x_1, \ldots, x_l) \in M^l$, $\mathbf{y} = (y_1, \ldots, y_m) \in M^m$, $m, l \geq 2$. *We consider some real weights* $\mathbf{a} = (a_1, \ldots, a_l) \in \mathbb{R}^l$ *(which can be nonpositive) and* $\mathbf{b} = (b_1, \ldots, b_m) \in [0, \infty)^m$.

We define the concept of weighted majorization $(\mathbf{x}, \mathbf{a}) \prec (\mathbf{y}, \mathbf{b})$ *by asking the existence of a matrix* $\mathbf{A} = (\alpha_{ij}) \in \mathcal{M}_{lm}(\mathbb{R})$ *such that*

- $\alpha_{ji} \leq 0 \qquad (i \neq 1 \text{ or } i \neq l)$;
- $y_i \in [y_{i-1}, y_{i+1}] \qquad (i \in \{2, 3, \ldots, l-1\})$ *verify that all these points belong to the same geodesic* $[y_1, y_m]$;
- x_i *is the weak barycenter of the family of points* $X := \{y_1, \ldots, y_m\}$ *with respect to the family of real weights* Λ^j *as the unique point* x_i *on the geodesic* $[y_1, y_m]$ *satisfying*

$$d(x_i, y_1) = \bar{S}_2^j d(y_2, y_1) + \bar{S}_3^j d(y_3, y_2) + \cdots + \bar{S}_n^j d(y_n, y_{n-1}), \tag{3.2.10}$$

or, equivalenty,

$$d(y_n, x_i) = S_1^j d(y_2, y_1) + S_2^j d(y_3, y_2) + \cdots + S_{n-1}^j d(y_n, y_{n-1}), \tag{3.2.11}$$

where

$$\begin{aligned} \Lambda^j &:= \{\alpha_{1j}, \ldots, \alpha_{nj}\} \qquad (j \in \{1, \ldots, m\}), \\ \bar{S}_k^j &= \alpha_{kj} + \cdots + \alpha_{nj} \qquad (k \in \{1, 2, \ldots, l\}), \\ S_k^j &= \alpha_{1j} + \cdots + \alpha_{kj} \qquad (k \in \{1, 2, \ldots, l\}), \\ 0 \leq S_k^j &\leq S_n^j = 1 \qquad (k \in \{1, 2, \ldots, l\}), \end{aligned}$$

- $$b_j = \sum_{i=1}^{l} a_i \alpha_{ji}, \quad (j = 1, \ldots, m), \tag{3.2.12}$$

We can now present the extension of Sherman's inequality in a global NPC space (M, d), when the weights are allowed to be nonpositive.

Theorem 3.2.4. *In the hypotheses from Definition 3.2.4 let us suppose that conditions* (3.2.10) *are satisfied. Then, the following inequality*

$$\sum_{i=1}^{m} a_i f(x_i) \leq \sum_{j=1}^{l} b_j f(y_j)$$

holds for every convex function $f : M \to \mathbb{R}$.

Proof. Since the hypotheses of Theorem 3.2.2 are satisfied for each $\bar{\mathbf{x}} = x_i$, we get

$$f(x_i) \leq \sum_{j=1}^{m} \alpha_{ij} f(y_j) \qquad (i = 1, \dots, l).$$

Taking into account (3.2.10) – (3.2.11) and applying Theorem 3.2.2 for each y_i, $i = 1, \dots, m$, where $y_i = \sum_{j=1}^{l} p_j x_j$, $p_j = \alpha_{ij}$, we get

$$\sum_{i=1}^{m} a_i f(x_i) = \sum_{i=1}^{m} a_i f\left(\sum_{j=1}^{m} y_j \alpha_{ji}\right)$$

$$\leq \sum_{i=1}^{l} a_i \left(\sum_{j=1}^{m} \alpha_{ji} f(y_j)\right) = \sum_{j=1}^{m} f(y_j) \sum_{i=1}^{l} a_i \alpha_{ji}.$$

Consequently, since $b_j = \sum_{i=1}^{l} a_i \alpha_{ji}$ we have

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

□

3.2.3 Conclusions and further open problems

Note that, in this section we extend the notion of barycenter for discrete Steffensen Popoviciu measures. In essence, we allow the case of nonpositive weights, but with additional conditions imposed to the points in which the measure is supported.

Hence, in this section of the thesis we assume that all the points in which is supported the discrete measure should be on the same geodesic. But, we can give some general conditions, in which the same results hold, but without imposing the same geodesic support.

More precisely, if the points $x_1, \dots, x_n \in M$ satisfy

$$\text{bar}\left(\frac{d(x_n, x_i)}{d(x_n, x_1)}\delta_{x_1} + \frac{d(x_i, x_1)}{d(x_n, x_1)}\delta_{x_n}\right) = x_i \qquad (i = 2, \dots, n-1),$$

then, all the inequalities from Theorems 3.2.2 and 3.2.3 still hold true.

Hence, one future aim will consists of looking for a characterization of the points verifying the above conditions. On the other hand, in [118, 124] the availability of Jensen's inequality in a certain nonconvex context is discussed. We outline the usefulness of the concept of point of convexity. We briefly present here the main ideas which will be used to treat the case of nonpositive weights in this context.

Definition 3.2.5. *Let $f : M \to \mathbb{R}$ be a continuous function. A point $a \in M$ is a* point of convexity *of the function f if*

$$f(a) \leq \sum_{i=1}^{n} \lambda_i f(x_i), \tag{3.2.13}$$

for every family of points $x_1, \ldots, x_n$ in M and every family of positive weights $\lambda_1, \ldots, \lambda_n$ with $\sum_{i=1}^{n} \lambda_i = 1$ and $\operatorname{bar}\left(\sum_{i=1}^{n} \lambda_i \delta_{x_i}\right) = a$.

The point a is a point of concavity *if it is a point of convexity for $-f$ (equivalently, if the above inequality works in the reversed way).*

In [118] we discuss the availability of Jensen's inequality in a nonconvex context, in which we emphasize the usefulness of the concept of point of convexity. Thus, even in the case of spaces with a curved geometry we have successfully introduced the point of convexity. See [124].

Hence, another future aim is to define a relaxed notion of "point of relative convexity", based on the barycenter discussed in this subsection. Moreover, our aim is to recover that all the convex type inequalities hold true if there exists such a point of relative convexity, all of this being open problems in global NPC spaces.

Chapter 4

Final remarks and open problems

In this chapter we present some final remarks and further open problems related to the results obtained in the present doctoral thesis. These problems are dealing with the possibility to continue to develop new results within the topic of this thesis. More precisely, we discuss about norm properties of the complete homogeneous symmetric polynomials, some error estimates with respect to euclidean norms and other symmetric inequalities related with discrete Korn's type inequalities. These norms, their unusual construction, and their potential applications suggest some open problems. Moreover, the topic of weak majorization in a global NPC spaces and its applications is also discussed.

4.1 Norms on complex matrices included by complete homogeneous symmetric polynomials

Note that this section is inspired from [4]. In this section we discuss about a family of norms on the space of $n \times n$ complex matrices which are initially defined in terms of certain symmetric functions of eigenvalues of complex Hermitian matrices. The fact that we deal with eigenvalues, as opposed to their absolute values, is notable. It prevents standard machinery, such as the theory of symmetric gauge functions, from applying and the techniques used to establish that we indeed have norms are more complicated than one might expect. For example, combinatorics, probability theory, and Lewis' framework for group invariance in convex matrix analysis each play key roles.

These norms on the Hermitian matrices are of independent interest, because they can be computed recursively or directly read from the characteristic polynomial. They can be extended in a natural and nontrivial manner to all complex matrices. Such extensions of original norms involve partition combinatorics and trace polynomials in noncommuting variables.

A Schur convexity argument permits our norms to be bounded below in terms of the mean eigenvalue of a matrix. Denote by $H_n(\mathbb{C})$ the set of $n \times n$ complex Hermitian matrices, $M_n(\mathbb{C})$ the set of $n \times n$ complex matrices and the eigenvalues of $A \in H_n(\mathbb{C})$ by

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$$

and define

$$\lambda(A) = (\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)) \in \mathbb{R}^n.$$

We may use λ and $\lambda_1, \lambda_2, \ldots, \lambda_n$ if the matrix A is clear from context. Denote by $\operatorname{diag}(x_1, x_2, \ldots, x_n) \in M_n(\mathbb{C})$ the $n \times n$ diagonal matrix with diagonal entries $x_1, x_2, \ldots, x_n$, in that order. If $x = (x_1, x_2, \ldots, x_n)$ is understood from context, we may write $\operatorname{diag}(x)$ for brevity. For the convenience of the reader we recall the complete homogeneous symmetric polynomials of degree d in the n variables $x_1, x_2, \ldots, x_n$ given by

$$h_d(x_1, x_2, \ldots, x_n) = \sum_{1 \le i_1 \le \cdots \le i_d \le n} x_{i_1} x_{i_2} \ldots x_{i_d}, \tag{4.1.1}$$

the sum of all degree d monomials in $x_1, x_2, \ldots, x_n$. For example,

$$\begin{aligned} h_0(x_1, x_2) &= 1, \\ h_1(x_1, x_2) &= x_1 + x_2, \\ h_2(x_1, x_2) &= x_1^2 + x_1 x_2 + x_2^2, \\ h_3(x_1, x_2) &= x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3. \end{aligned}$$

When the degree d is even and $x \in \mathbb{R}^n$, Hunter proved that $h_d(x) \ge 0$, with equality if and only if $x = 0$ [74]. This is not obvious because some of the summands that comprise $h_d(x)$ for d even may be negative.

Definition 4.1.1. *A partition of $d \in \mathbb{N}$ is an r-tuple $\pi = (\pi_1, \pi_2, \ldots, \pi_r) \in \mathbb{N}^r$ such that $\pi_1 \ge \pi_2 \ge \cdots \ge \pi_r$ and $\pi_1 + \pi_2 + \cdots + \pi_r = d$, the number of terms r depends on the partition π. We say that $\pi \vdash d$ if π is a partition of d.*

For any $\pi \vdash d$, we define

$$p_\pi(x_1, x_2, \ldots, x_n) = p_{\pi_1} p_{\pi_2} \ldots p_{\pi_r},$$

where $p_k(x_1, x_2, \ldots, x_n) = x_1^k + x_2^k + \cdots + x_n^k$ are the power sum symmetric polynomials. In another form, we can write

$$h_d(x_1, x_2, \ldots, x_n) = \sum_{\pi \vdash d} \frac{p_\pi(x_1, x_2, \ldots, x_n)}{z_\pi}, \tag{4.1.2}$$

in which the sum runs over all partitions $\pi = (\pi_1, \pi_2, \ldots, \pi_r)$ of d and

$$z_\pi = \prod_{i \ge 1} i^{m_i} m_i!,$$

where m_i is the multiplicity of i in π. For example, if $\pi = (4, 4, 2, 1, 1, 1)$ then $z_\pi = (1^3 3!)(2^1 1!)(4^2 2!) = 384$. The integer z_π is precisely the Hall inner product of p_π with itself, in symmetric function theory.

If $A \in H_n(\mathbb{C})$ has eigenvalues $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, then

$$p_\pi(\lambda) = p_{\pi_1}(\lambda) p_{\pi_2}(\lambda) \cdots p_{\pi_r}(\lambda) = (trA^{\pi_1})(trA^{\pi_2}) \cdots (trA^{\pi_r}). \tag{4.1.3}$$

Thus, the previous relations connects eigenvalues, traces and partitions to symmetric polynomials.

The following theorem provides a family of novel norms on the space $H_n(\mathbb{C})$ of $n \times n$ Hermitian matrices. See [4].

Theorem 4.1.1. *For even $d \ge 2$, the following is a norm on $H_n(\mathbb{C})$:*

$$|||A|||_d = (h_d(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)))^{1/d}.$$

For example, using equations (4.1.2) and (4.1.3) we obtain trace-polynomial representations

$$|||A|||_2^2 = \frac{1}{2}(\mathrm{tr}(A^2) + (\mathrm{tr}A)^2), \tag{4.1.4}$$

$$|||A|||_4^4 = \frac{1}{24}((\mathrm{tr}A)^4 + 6(\mathrm{tr}A)^2\mathrm{tr}(A^2) + 3(\mathrm{tr}(A^2))^2 + 8(\mathrm{tr}A)\mathrm{tr}(A^3) + 6\mathrm{tr}(A^4)). \tag{4.1.5}$$

The above theorem needs some useful remarks.

(1) The sums (4.1.1) and (4.1.2) that characterize $h_d(\lambda(A))$ may contain negative summands.

(2) The sums that define these norms do not involve the absolute values of the eigenvalues of A.

(3) The relationship between the spectra of (Hermitian) A, B and $A+B$, conjectured by A. Horn in 1962 [72], was only established in 1998-9 by Klyachko [84] and Knutson-Tao [85]. Therefore, the triangle inequality is difficult to establish. Even if A and B are diagonal, the result is not obvious, but also in the case of positive diagonal matrices this result has been rediscovered.

(4) In the general Hermitian case is not straight-forward: standard techniques like symmetric gauge functions are not applicable, we need to involves Lewis' framework for group invariance in convex matrix analysis.

(5) Another genuine approach to norms on $\mathbb{R}^n$ is due to Ahmadi, de Klerk, and Hall [[5], Thm. 2.1]. Note that, using Theorem 4.1.1 together with [[5], Thm. 2.1] we get convexity property.

4.2 Complete homogeneous symmetric polynomials as expectations

In this section we present a change of the perspective, by connecting complete homogeneous symmetric polynomials with expectations. Let us consider $\zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n)$ be a vector of independent standard exponential random variables and let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Since $\mathbb{E}[\zeta_i^k] = k!$ for $i = 1, 2, \ldots, n$ we have that

$$\begin{aligned}
\mathbb{E}[\langle \zeta, x\rangle^d] &= \mathbb{E}[(\zeta_1 x_1 + \zeta_2 x_2 + \cdots + \zeta_n x_n)^d] \\
&= \mathbb{E}\Bigg[\sum_{k_1+k_2+\ldots k_n = d} \frac{d!}{k_1!k_2!\ldots k_n!}\zeta_1^{k_1}\zeta_2^{k_2}\ldots\zeta_n^{k_n}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\Bigg] \\
&= \sum_{k_1+k_2+\ldots k_n = d} \frac{d!}{k_1!k_2!\ldots k_n!}\mathbb{E}\Big[\zeta_1^{k_1}\zeta_2^{k_2}\ldots\zeta_n^{k_n}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\Big] \\
&= d! \sum_{k_1+k_2+\ldots k_n = d} \frac{\mathbb{E}[\zeta_1^{k_1}]\mathbb{E}[\zeta_2^{k_2}]\ldots\mathbb{E}[\zeta_n^{k_n}]}{k_1!k_2!\ldots k_n!}x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \\
&= d! \sum_{k_1+k_2+\ldots k_n = d} x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \\
&= d! h_d(x),
\end{aligned}$$

for integral $d \geq 1$, where for the last estimates we have used the linearity of expectation and the independence of the $\zeta_1, \zeta_2, \ldots, \zeta_n$.

Sincer for any d an even natural number, we have

$$h_d(x) = \frac{1}{d!}\mathbb{E}[|\langle \zeta, x\rangle|^d] \geq 0, \tag{4.2.1}$$

Minkowski's inequality implies that

$$\left(\mathbb{E}[|\langle \zeta, x+y\rangle|^d]\right)^{1/d} \leq \left(\mathbb{E}[|\langle \zeta, x\rangle|^d]\right)^{1/d} + \left(\mathbb{E}[|\langle \zeta, y\rangle|^d]\right)^{1/d},$$

$$[h_d(x+y)]^{1/d} \leq [h_d(x)]^{1/d} + [h_d(y)]^{1/d}.$$

for $x, y \in \mathbb{R}^n$.

We consider now the inner product $\langle X, Y\rangle = \mathrm{tr}(XY)$ on $H_n(\mathbb{C})$, which is the restriction of the Frobenius inner product to $H_n(\mathbb{C})$. The inequality

$$\mathrm{tr}(XY) \leq \mathrm{tr}\delta(X)\delta(Y) \qquad (X, Y \in H_n(\mathbb{C})) \tag{4.2.2}$$

is due to von Neumann and, for diagonal matrices, is equivalent to a classical rearrangement result

$$\langle x, y\rangle \leq \langle \tilde{x}, \tilde{y}\rangle,$$

where $\tilde{x} \in \mathbb{R}^n$ has the components of $x = (x_1, x_2, \ldots, x_n)$ in decreasing order.

For even d, the nonnegativity of the polynomials has a probabilistic approach which appears in [165], and in [[164], Lem. 12], which cites [20]. There are many other proofs of the nonnegativity of the even-degree CHS polynomials. Of course, there is Hunter's inductive proof [74]. Rovenţa and Temereancă used divided differences [[153], Thm. 3.5]. Recently, Böttcher, Garcia, Omar and O'Neill [29] employed a spline-based approach suggested by Olshansky after Garcia, Omar, O'Neill, and Yih obtained it as a byproduct of investigations into numerical semigroups [[60], Cor. 17].

The CHS norm of a Hermitian matrix can be exactly computed from its characteristic polynomial. The following theorem involves only formal series manipulations. See [4].

Theorem 4.2.1. *Let $p_A(x)$ denote the characteristic polynomials of $A \in H_n(\mathbb{C})$. For $d \geq 2$ even, $|||A|||_d^d$ is the d-th coefficient in the Taylor expansion around the origin of*

$$\frac{1}{det(I - xA)} = \frac{1}{x^n p_A(1/x)}.$$

Example 5. *Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Then $p_A(z) = x^2 - x - 1$ and*

$$\frac{1}{x^2 p_A(1/x)} = \frac{1}{1 - x - x^2} = \sum_{n=0}^{\infty} f_{n+1}x^n,$$

in which f_n is the n-th Fibonacci number; given by $f_{n+2} = f_{n+1} + f_n$ and $f_0 = 0$ and $f_1 = 1$. It follows that $|||A|||_d^d = f_d$ for even $d \geq 2$.

Note that, for any $A \in H_n(\mathbb{C})$, the sequence $h_d(\lambda_1, \lambda_2, \ldots, \lambda_n)$ satisfies a constant-coefficient recurrence of order n since its generating function is a rational function whose denominator has degree n. Using this method we can solve such a recurrence and we can compute explicitly $||A||_d$, for $d = 2, 4, 6, \ldots$.

On the other hand, for any $d \geq 1$, the Newton-Gerard identities [158] give

$$h_d(x_1, x_2, \ldots, x_n) = \frac{1}{d}\sum_{i=1}^{d} h_{d-i}(x_1, x_2, \ldots, x_n)p_i(x_2, x_2, \ldots, x_n).$$

Hence, for $A \in H_n(\mathbb{C})$ and $d \geq 2$ even, we have that

$$h_d(\lambda(A)) = \frac{1}{d}\sum_{i=1}^{d} h_{d-i}(\lambda(A))tr(A^i),$$

and then we can recursively compute $|||A|||_d^d = h_d(\lambda(A))$.

In the following we try to show that each CHS norm on $M_n(\mathbb{C})$ is bounded below by an explicit positive multiple of the trace seminorm. That is, the CHS norms of a matrix can be related to its mean eigenvalue. See [4].

Theorem 4.2.2. *For $A \in M_n(\mathbb{C})$ and $d \geq 2$ even,*

$$|||A|||_d \geq \binom{n+d-1}{d}^{1/d} \frac{|trA|}{n},$$

with equality if and only if A is a multiple of the identity.

Proof. Let $d \geq 2$ be even. The even-degree complete homogeneous symmetric polynomials are Schur convex [165].

Let $A \in M_n(\mathbb{C})$ and define $B(t) = e^{it}A + e^{-it}A^*$ for $t \in \mathbb{R}$. Then $\lambda(B(t))$ majorizes $\mu(t) = (\mu(t), \mu(t), \ldots, \mu(t)) \in \mathbb{R}^n$, where $\mu(t) = trB(t)/n$.

We have that

$$|||B(t)|||_d^d = h_d(\lambda(B(t))) \geq h_d(\mu(t)) = \mu(t)^d \binom{n+d-1}{d},$$

$$|||A|||_d \geq \left(\frac{\binom{n+d-1}{d}}{2\pi\binom{d}{d/2}} \int_0^{2\pi} \mu(t)^d dt \right)^{1/d}. \tag{4.2.3}$$

Taking into account that

$$\int_0^{2\pi} \mu(t)^d dt = \int_0^{2\pi} \left(\frac{trB(t)}{n}\right)^d dt = \frac{1}{n^d}\int_0^{2\pi} (e^{it}trA + e^{-it}tr(A^*))^d dt$$

$$= \frac{1}{n^d}\sum_{k=0}^{d} \binom{d}{k} (trA^*)^{d-k}(trA)^k \int_0^{2\pi} e^{i(2k-d)t} dt$$

$$= \frac{2\pi}{n^d}\binom{d}{d/2}|trA|^d,$$

which gives the conclusion. The Fourier expansion gives that

$$e^{it}A + e^{-it}A^* = \left(\sum_{n\in\mathbb{Z}} \hat{\mu}(n)e^{int}\right) I,$$

thus $A = \hat{\mu}(1)I$. The converse also holds. $\square$

Note also that, for $A \in H_n(\mathbb{C})$ and even $d \geq 2$, we can deduce that

$$\left(\frac{1}{2^{\frac{d}{2}}(\frac{d}{2})!}\right)^{1/d} ||A||_{op} \leq |||A|||_d \leq \binom{n+d-1}{d}^{1/d} ||A||_{op}.$$

4.3 Polynomial norms

Note that this section is inspired from [5]. In this section, we present some polynomial norms, which means norms that are the d^{th} root of a homogeneous polynomial with degree d. An interesting connection between convexity and norm is given in the following theorem. See [5].

Theorem 4.3.1. *Let f be a form of degree d in n variables. The following statements are equivalent:*

(1) *The function $f^{\frac{1}{d}}$ is a norm on $\mathbb{R}^n$.*

(2) *The function f is convex and positive definite.*

(3) *The function f is strictly convex.*

Proof. (1) $\implies$ (2) Since $f^{1/d}$ is a norm, then $f^{1/d}$ is positive definite, and so is f. Notice that any norm is convex and the d^{th} power of a nonnegative convex function remains convex.
(2) $\implies$ (3) Suppose that f is convex, positive definite, but not strictly convex and we can say that there exists $\bar{x}, \bar{y} \in \mathbb{R}^n$ with $\bar{x} \neq \bar{y}$, and $\gamma \in (0,1)$ such that

$$f(\gamma\bar{x} + (1-\gamma)\bar{y}) = \gamma f(\bar{x}) + (1-\gamma)f(\bar{y}).$$

Let $g(\alpha) := f(\bar{x} + \alpha(\bar{y} - \bar{x}))$. Note that g is a restriction of f to a line and, consequently, g is a convex, positive definite, univariate polynomial in α. We now define

$$h(\alpha) := g(\alpha) - (g(1) - g(0))\alpha - g(0). \tag{4.3.1}$$

Similarly to g, h is a convex univariate polynomial as it is the sum of two convex univariate polynomials. We also know that $h(\alpha) \geq 0, \forall \alpha \in (0,1)$. Indeed, by convexity of g, we have that

$$g(\alpha x + (1-\alpha)y) \geq \alpha g(x) + (1-\alpha)g(y),\ \forall x, y \in \mathbb{R} \text{ and } \alpha \in (0,1).$$

This inequality holds in particular for $x = 1$ and $y = 0$, which proves the claim. Observe now that $h(0) = h(1) = 0$. By convexity of h and its nonnegativity over $(0,1)$, we have that $h(\alpha) = 0$ on $(0,1)$ which further implies that $h = 0$. Hence, from (4.3.1), g is an affine function. As g is positive

definite, it cannot be that g has a nonzero slope, so g has to be a constant. But this contradicts that $\lim_{\alpha\to\infty} g(\alpha) = \infty$.
To see why this limit must be infinite, we show that $\lim_{||x||\to\infty} f(x) = \infty$. As

$$\lim_{\alpha\to\infty} ||\bar{x} + \alpha(\bar{y} - \bar{x})|| = \infty \text{ and } g(\alpha) = f(\bar{x} + \alpha(\bar{y} - \bar{x})),$$

implies that $\lim_{\alpha\to\infty} g(\alpha) = \infty$. To show that $\lim_{||x||\to\infty} f(x) = \infty$, let

$$x^* = \arg\min_{||x||=1} f(x).$$

By positive definiteness of f, $f(x^*) > 0$. Let M be any positive scalar and define $R := (M/f(x^*))^{1/d}$. Then for any x such that $||x|| = R$, we have

$$f(x) \geq \min_{||x||=R} f(x) \geq R^d f(x^*) = M,$$

where the second inequality holds by homogeneous of f. Thus $\lim_{||x||\to\infty} f(x) = \infty$.
(3) $\implies$ (1) Homogeneity of $f^{1/d}$ is immediate. Positivity follows from the first-order characterization of strict convexity:

$$f(y) > f(x) + \bigtriangledown f(x)^T(y - x),\ \forall y \neq x.$$

Indeed, for $x = 0$, the inequality becomes $f(x) > 0$, $\forall y \neq 0$, as $f(0) = 0$ and $\bigtriangledown f(0) = 0$. Hence, f is positive definite, and so is $f^{1/d}$. It remains to prove the triangle inequality. Let $g := f^{1/d}$. Denote by S_f and S_g the 1-sublevel sets of f and g respectively. It is clear that

$$S_g = \{x | f^{1/d}(x) \leq 1\} = \{x | f(x) \leq 1\} = S_f,$$

and as f is strictly convex (and hence quasi-convex), S_f is convex and so is S_g. Let $x, y \in \mathbb{R}^n$. We have that $\frac{x}{g(x)} \in S_g$ and $\frac{y}{g(y)} \in S_g$. From convexity of S_g,

$$g\left(\frac{g(x)}{g(x)+g(y)} \cdot \frac{x}{g(x)} + \frac{g(y)}{g(x)+g(y)} \cdot \frac{y}{g(y)}\right) \leq 1.$$

Homogeneity of g then gives us

$$\frac{1}{g(x)+g(y)} g(x+y) \leq 1,$$

which shows that triangle inequality holds. $\square$

Taking into account that not all norms are polynomial norms we are asking if we can generally approximate the norms by polynomial norms.

In the following, we show that, though not every norm is a polynomial norm, but any norm can be approximated to arbitrary precision by a polynomial norm. A related result is given by Barvinok in [18]. In that section, he shows that any norm can be approximated by the d-th root of a nonnegative degree-d form, and quantifies the quality of the approximation as a function of n and d. The form he obtains however is not shown to be convex. In fact, in a later work [[19], Section 2.4], Barvinok points out that it would be an interesting question to know whether any norm can be approximated by the d^{th} root of a convex form with the same quality of approximation as for d-th roots of nonnegative forms.

The result below is a step in that direction, although the quality of approximation is weaker than that by Barvinok [18]. We note that the form in Barvinok's construction is a sum of squares of other forms. Such forms are not necessarily convex. By contrast, the form that we construct is a sum of powers of linear forms and hence always convex.

Theorem 4.3.2. *Let $||\cdot||$ be any norm on $\mathbb{R}^n$. Then, for any even integer $d \geq 2$:*
(i) There exists an n-variable convex positive definite form f_d of degree d such that

$$\frac{d}{n+d}\left(\frac{n}{n+d}\right)^{n/d}||x|| \leq f_d^{1/d}(x) \leq ||x||, \; (x \in \mathbb{R}^n). \tag{4.3.2}$$

In particular, for any sequence $\{f_d\}$ $(d = 2, 4, 6, ...)$ of such polynomials one has

$$\lim_{d\to\infty} \frac{f_d^{1/d}(x)}{||x||} = 1 \; \forall x \in \mathbb{R}^n.$$

(ii) One may assume without loss of generality that f_d in (i) is a nonnegative sum of d^{th} powers of linear forms.

Moreover, we also present recent concavity and convexity results for symmetric polynomials and their ratios. Note that this part of this section is inspired from [164].

More precisely, our aim is to present the proof of the following convex type inequalities:

$$[e_k((x+y)^p)]^{1/pk} \geq [e_k(x^p)]^{1/pk} + [e_k(y^p)]^{1/pk} \tag{4.3.3}$$

$$[h_k((x+y)^{1/p})]^{p/k} \geq [h_k(x^{1/p})]^{p/k} + [h_k(y^{1/p})]^{p/k} \tag{4.3.4}$$

$$\left[\frac{e_k((x+y)^p)}{e_{k-l}((x+y)^p)}\right]^{\frac{1}{lp}} \leq \left[\frac{e_k(x^p)}{e_{k-l}(x^p)}\right]^{\frac{1}{lp}} + \left[\frac{e_k(y^p)}{e_{k-l}(y^p)}\right]^{\frac{1}{lp}} \tag{4.3.5}$$

$$\left[\frac{h_k((x+y)^{1/p})}{h_1((x+y)^{1/p})}\right]^{\frac{p}{k-1}} \leq \left[\frac{h_k(x^{1/p})}{h_1(x^{1/p})}\right]^{\frac{p}{k-1}} + \left[\frac{h_k(y^{1/p})}{h_1(y^{1/p})}\right]^{\frac{p}{k-1}}, \tag{4.3.6}$$

for $x, y \in \mathbb{R}^n_+$, $p \in (0,1)$ and $1 \leq l \leq k \leq n$. In these inequalities e_k denotes the k-th elementary symmetric polynomial

$$e_k(x) = \sum_{S\subseteq[n],[S]=k} \prod_{i\in S} x_i$$

and x^p we mean the vector $(x_1^p, \ldots, x_n^p)$.

We firstly present an important concavity result for e_k, the Marcus-Lopes inequality [102]:

$$\frac{e_k(x+y)}{e_{k-1}(x+y)} \geq \frac{e_k(x)}{e_{k-1}(x)} + \frac{e_k(y)}{e_{k-1}(y)}, \quad 1 \leq k \leq n, \;\; x, y \in \mathbb{R}^n_+. \tag{4.3.7}$$

This inequality is used by Marcus and Lopes to prove the concavity of $e_k(x)^{1/k}$. Now, we give the following concavity inequality:

$$\left[\frac{e_k((x+y)^p)}{e_{k-1}((x+y)^p)}\right]^{\frac{1}{p}} \geq \left[\frac{e_k(x^p)}{e_{k-1}(x^p)}\right]^{\frac{1}{p}} + \left[\frac{e_k(y^p)}{e_{k-1}(y^p)}\right]^{\frac{1}{p}}. \tag{4.3.8}$$

We introduce the parallel sum operation which is an important element of the proof:

$$x : y := (x^{-1} + y^{-1})^{-1}, \quad x, y > 0.$$

Lemma 4.3.1. *The parallel sum is jointly concave on $\mathbb{R}^2_+$. Moreover, for every $p \geq -1$, the p-parallel sum $x :_p y := [x^p : y^p]^{1/p}$ is jointly concave. Also, both $:$ and $:_p$ are monotonic in both arguments.*

Proof. The Hessian equals

$$(p+1)\begin{pmatrix} -x^{p-1}y^{p+1}(x^p+y^p)^{-2-\frac{1}{p}} & x^p y^p (x^p+y^p)^{-2-\frac{1}{p}} \\ x^p y^p (x^p+y^p)^{-2-\frac{1}{p}} & -x^{p+1}y^{p-1}(x^p+y^p)^{-2-\frac{1}{p}} \end{pmatrix},$$

which is clearly negative definite for $p \geq -1$. Monotonicity is clear from first derivatives. $\square$

Observe that $x : y = y : x$ and $(x : y) : z = x : (y : z)$, thus we extend the above notation and simply write $x_1 : x_2 : \cdots : x_n \equiv x_1 : [x_2 : [\cdots : (x_{n-1} : x_n)]]$. However, the operation $:_p$ is not associative if generalized the same way. Thus, a more preferable multivariate generalization is the following:

$$(x_1, \ldots, x_n) \to [x_1^p : x_2^p : \cdots : x_n^p]^{1/p}.$$

Lemma 4.3.2. *Let $f_1 : \mathbb{R}^{m_1} \to \mathbb{R}_{++}$ and $f_2 : \mathbb{R}^{m_2} \to \mathbb{R}_{++}$ be continuous concave functions. Then, $f_1(x) :_p f_2(x)$ is jointly concave on $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$.*

Proof. It is suffices to establish midpoint concavity. Since f_1 and f_2 are concave, we have

$$f_1\left(\frac{x_1+x_2}{2}\right) \geq \frac{1}{2}f_1(x_1) + \frac{1}{2}f_1(x_2), \text{ and } f_2\left(\frac{y_1+y_2}{2}\right) \geq \frac{1}{2}f_2(y_1) + \frac{1}{2}f_2(y_2).$$

The function $:_p$ is monotonically increasing in each of its arguments and is jointly concave, therefore

$$\begin{aligned} f_1\left(\frac{x_1+x_2}{2}\right) :_p f_2\left(\frac{y_1+y_2}{2}\right) &\geq \left(\frac{f_1(x_1)+f_1(x_2)}{2}\right) :_p \left(\frac{f_2(y_1)+f_2(y_2)}{2}\right) \\ &\geq \frac{1}{2}[f_1(x_1) :_p f_2(y_1)] + \frac{1}{2}[f_1(x_2) :_p f_2(y_2)], \end{aligned}$$

which establish the joint concavity of $f_1(x) :_p f_2(x)$. $\square$

Now we are able to present the main result from [164].

Theorem 4.3.3. *Let $1 \leq k \leq n$. Then, the function*

$$\phi_{k,n}(x) := \left[\frac{e_k(x^p)}{e_{k-1}(x^p)}\right]^{1/p}$$

is concave for $x \in \mathbb{R}^n_+$ and $p \in (0,1)$.

Proof. We use induction on k and for $k = 1$, we have

$$\phi_{1,n}(x) = (x_1^p + x_2^p + \cdots + x_n^p)^{1/p}.$$

which is clearly concave. We assume that $\phi_{k-1,n}$ is concave for some $k > 1$. The key step in our proof is the following remarkable observation of Anderson et al [12]:

$$\text{Let } \Psi_{k,n}(x) := \frac{\binom{n}{k-1}e_k(x)}{\binom{n}{k}e_{k-1}(x)}, \text{ then } \Psi_{k,n} = \sum_{j=1}^{n} \frac{1}{n-k+1}x_j : \frac{1}{k-1}\Psi_{k-1,n}(x_{[j]}),$$

where $x_{[j]}$ denotes the vector x with x_j omitted. Using the previous representation we see that for suitable positive scaling factors a_k and b_k, we can write

$$\phi_{k,n}(x) = \left(\sum_{j=1}^{n} a_k^p x_j^p : b_k^p \phi_{k-1,n}^p(x_{[j]})\right)^{1/p} = \left(\sum_{j=1}^{n} [a_k x_j :_p b_k \phi_{k-1,n}(x_{[j]})^p\right)^{1/p}. \tag{4.3.9}$$

From induction hypothesis we know that $\phi_{k-1,n}(\cdot)$ is concave; thus, applying Lemma 4.3.2 we see that

$$g_j(x) := (a_k x_j) :_p (b_k \psi_{k-1,n}(x_{[j]}))$$

is jointly concave in x_j and $x_{[j]}$ and thus in x. Consequently, we can further rewrite (4.3.9) as

$$\phi_{k,n}(x) = \left(\sum_{j=1}^{n} g_j(x)^p\right)^{1/p},$$

which is clearly concave as it is the (vector) composition the coordinate-wise increasing concave function $(\sum_j x_j^p)^{1/p}$ with the vector $(g_1(x), \ldots, g_n(x))$, where each $g_j(x)$ is itself concave. □

Following the idea of Marcus and Lopez [102], who proved concavity of $e_k(x)^{1/k}$, we can now prove (4.3.3), that is, the concavity of $[e_k(x^p)]^{1/p}$, by leveraging the ratio-concavity proved in Theorem 4.3.3.

Theorem 4.3.4. *The function $x \to [e_k(x^p)]^{1/pk}$ is concave for $p \in (0,1)$ and $x \in \mathbb{R}_+^n$.*

In the same vein, we easily obtain a proof to (4.3.5) which generalizes Theorem 4.3.3.

For any $1 \le l \le k \le n$, then

$$\phi_{k,l,n}(x) := \left[\frac{e_k(x^p)}{e_{k-l}(x^p)}\right]^{1/lp},$$

is concave for $x \in \mathbb{R}_+^n$ for $p \in (0,1)$. See [164].

Proposition 4.3.1. *The multivariate p-parallel sum map*

$$x \to [x_1^p : x_2^p : \cdots : x_n^p]^{1/p}$$

is concave on $\mathbb{R}_+^n$ for $p > 0$.

Proof. The proof is by induction on n. Assume thus that the claim holds for $\mathbb{R}_+^k$ for $k = 1, 2, \ldots, n-1$. Consider thus,

$$S_n(x_1, \ldots, x_n) = [x_1^p : x_2^p : \cdots : x_n^p]^{\frac{1}{p}}.$$

From the induction hypothesis, we have that

$$S_{n-1}\left(\frac{1}{2}(x_1+y_1), \ldots, \frac{1}{2}(x_{n-1}+y_{n-1})\right) \ge \frac{1}{2}S_{n-1}(x_1, \ldots, x_{n-1}) + \frac{1}{2}S_{n-1}(y_1, \ldots, y_{n-1}). \tag{4.3.10}$$

Thus, using the monotonicity of the $:_p$ operation and (4.3.10) we get

$$\begin{aligned} S_n\left(\frac{1}{2}(x_1+y_1), \ldots, \frac{1}{2}(x_n+y_n)\right) &= \left\{\left[S_{n-1}\left(\frac{1}{2}(x_1+y_1), \ldots, \frac{1}{2}(x_{n-1}+y_{n-1})\right)\right]^p : \left(\frac{1}{2}x_n + \frac{1}{2}y_n\right)^p\right\}^{1/p} \\ &\ge \left\{\left[\frac{1}{2}S_{n-1}(x_1, \ldots, x_{n-1}) + \frac{1}{2}S_{n-1}(y_1, \ldots, y_{n-1})\right]^p : \left(\frac{1}{2}x_n + \frac{1}{2}y_n\right)^p\right\}^{1/p}. \end{aligned}$$

Now concavity of $:_p$ immediately yields

$$\left[\left(\frac{S_{n-1}(x_1,\ldots,x_{n-1})+S_{n-1}(y_1,\ldots,y_{n-1})}{2}\right)^p : \left(\frac{x_n+y_n}{2}\right)^p\right]^{1/p}$$

$$\geq \frac{1}{2}\left[\left\{S_{n-1}(x_1,\ldots,x_{n-1})\right\}^p : x_n^p\right]^{1/p} + \frac{1}{2}\left[\left\{S_{n-1}(y_1,\ldots,y_{n-1})\right\}^p + y_n^p\right]^{1/p},$$

establishing concavity of S_n. Thus, by induction the said map is concave. □

In a similar way, we can deduce that the function $x \to e_k(x^p)$ is reciprocally concave for $p \in (-1, 0)$ and $x \in \mathbb{R}_+^n$. See [164]. Note also that that for any $\mathbf{X}$ a Hermitian positive definite matrix, the function

$$\mathbf{X} \to \frac{1}{e_k(\lambda(\mathbf{X})^p)}, \quad p \in (-1, 0),$$

is concave, where $\lambda(\cdot)$ denotes the eigenvalue map.

Using the following representation for complete homogeneous symmetric polynomials

$$\frac{1}{k!}\mathbb{E}[(\xi_1 x_1 + \cdots + \xi_n x_n)^k] = \sum_{1\leq i_1\leq i_2\leq\cdots\leq i_k\leq n} \prod_{j=1}^{k} x_{i_j} = h_k(x), \tag{4.3.11}$$

we can deduce the proof of the following convexity result of McLeod [105]:

$$[h_k(x+y)]^{1/k} \leq [h_k(x)]^{1/k} + [h_k(y)]^{1/k}. \tag{4.3.12}$$

More precisely, using the fact that for any $p \geq 1$, $k \geq 1$, $x, y \in \mathbb{R}_+^n$ we have

$$\left[\sum_i \left(\sum_j x_{ij}^p + y_{ij}^p\right)^k\right]^{1/pk} \leq \left(\sum_{ij} x_{ij}^{pk}\right)^{1/pk} + \left(\sum_{ij} y_{ij}^{pk}\right)^{1/pk} \tag{4.3.13}$$

we get the following two inequalities

$$[h_k(x+y)^p]^{1/kp} \leq [h_k(x^p)]^{1/kp} + [h_k(y)^p]^{1/kp},$$

$$\left[\frac{h_k((x+y)^p)}{h_1((x+y)^p)}\right]^{1/p(k-1)} \leq \left[\frac{h_k(x^p)}{h_1(x^p)}\right]^{1/p(k-1)} + \left[\frac{h_k(y^p)}{h_1(y^p)}\right]^{1/p(k-1)}.$$

Our previous estimates obtained in this doctoral thesis can be useful to continue this research subject.

4.4 A finite difference approach of Korn's inequalities

In this section we introduce a completely new topic related to future perspectives of this doctoral thesis. Note that this part is inspired from [61].

Korn's inequalities represent a main tool in linearized elasticity theory, that proves useful not only in connection with the basic theoretical issues such as existence and uniqueness, but also in a large variety of applications.

Korn himself considered only the case of vector fields vanishing on the boundary (see [88]-[90]). His results were subsequently improved and extended by many people including Friedrichs [57], Gobert [62], Duvaut and Lions [54], Ciarlet [41], [46], Kondratiev and Oleinik [86], Oleinik, Shamaev and Yosifan [136], Horgan [70], Desvillettes and Villani [50], Fuchs [58], Neff, Pauly and Witsch [110], to mention here just a few contributions in chronological order. There is by now a huge literature on the subject: a search on the electronic database of Google lists about 19,400 references concerned with Korn's inequalities.

Our aim is to offer new very simple proofs of these inequalities via the finite difference method. We follow the terminology and notation used in the classical book of Ciarlet [41].

Let Ω be a domain in $\mathbb{R}^N$, that is an open, bounded and simply connected subset with sufficiently smooth boundary $\partial\Omega$. We consider a smooth vector field $f : \Omega \to \mathbb{R}^N$ of components f_k having the Jacobian matrix $\nabla f = \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j}$. The symmetric part of ∇f is the matrix $\nabla^{sym} f$ with entries

$$\frac{1}{2}\left(\frac{\partial f_j}{\partial x_i} + \frac{\partial f_i}{\partial x_j}\right).$$

Denoting by $|\nabla f|$ and $|D^{sym} f|$ the corresponding Hilbert-Schmidt matrix norms, the original first inequality of Korn states that, if f belongs to $C_c^1(\Omega, \mathbb{R}^N)$ (the space of continuously differentiable fields with compact support), then

$$\int_\Omega |\nabla f|^2 dx \leq 2 \int_\Omega |\nabla^{sym} f|^2 dx. \tag{$K1$}$$

Friedrichs [57] noticed that this inequality follows from an elementary identity. In fact, when f is $C_c^2(\Omega, \mathbb{R}^N)$, a twice applications of integration by parts yields

$$\int_\Omega \frac{\partial f_i}{\partial x_j}\frac{\partial f_j}{\partial x_i} dx = \int_\Omega \frac{\partial f_i}{\partial x_i}\frac{\partial f_j}{\partial x_j} dx,$$

from which one derives the formula

$$\begin{aligned}
\int_\Omega |\nabla^{sym} f|^2 \, dx &= \frac{1}{4}\sum_{i,j=1}^N \int_\Omega \left(\frac{\partial f_j}{\partial x_i} + \frac{\partial f_i}{\partial x_j}\right)^2 dx \\
&= \frac{1}{2}\sum_{i,j=1}^N \left[\int_\Omega \left(\frac{\partial f_j}{\partial x_i}\right)^2 dx + \int_\Omega \frac{\partial f_i}{\partial x_j}\frac{\partial f_j}{\partial x_i} dx\right] \\
&= \frac{1}{2}\int_\Omega \sum_{i,j=1}^N \left(\frac{\partial f_j}{\partial x_i}\right)^2 dx + \int_\Omega \sum_{i,j=1}^N \frac{\partial f_i}{\partial x_i}\frac{\partial f_j}{\partial x_j} dx \\
&= \frac{1}{2}\int_\Omega |\nabla f|^2 \, dx + \frac{1}{2}\int_\Omega (\operatorname{div} f)^2 \, dx,
\end{aligned}$$

and Korn's inequality $(K1)$ is now obvious. Then a standard approximation procedure yields the general statement: the inequality $(K1)$ holds for all fields in the Sobolev space $H_0^1(\Omega; \mathbb{R}^N)$.

In the next subsection we will present an alternative approach of Korn's first inequality using finite differences. The complexity of the proof is pretty much the same, and represent a bridge for a similar approach of Korn's second inequality (whose all known proofs are nontrivial).

Theorem 4.4.1. *(The second Korn inequality) Let Ω be a domain in $\mathbb{R}^N$ with Lipschitz boundary. There exist of a positive constant $C = C(\Omega)$ such that*

$$\|f\|_{H^1}^2 \leq C\left(\|f\|_{L^2}^2 + \|\nabla^{sym} f\|_{L^2}^2\right) \quad \text{for all } f \in H^1(\Omega;\mathbb{R}^N).$$

Having Lipschitz boundary means that for any point $a \in \partial\Omega$ one can introduce orthogonal coordinates $y = C(x-a)$, where C is an $N \times N$-dimensional constant matrix such that in coordinates $y = (\hat{y}, y_n)$, with $\hat{y} = (y_1, ..., y_{n-1})$, the intersection of $\partial\Omega$ with the cylinder

$$C_{R,L} = \{(\hat{y}, y_n) : \ |\hat{y}| < R, \ -LR < y_n < LR\},$$

is given by the equation $y_n = \varphi(\hat{y})$, where $\varphi(\hat{y})$ statisfies the Lipschitz condition in $\{\hat{y} : |\hat{y}| < R\}$ with Lipschitz constant not larger than L and

$$\Omega \cap \overline{C}_{R,L} = \{y : |\hat{y}| < R, \ -LR \leq y_n \leq LR\}.$$

The numbers R and L are assumed to be the same for any point $a \in \partial\Omega$ and depend only on Ω.

For convenience of the reader, we recall here some properties of functions defined in domains with Lipschitz boundary.

Theorem 4.4.2. *(See [136], Theorem 1.2, p. 4) Let Ω be a domain in $\mathbb{R}^N$ with Lipschitz boundary. Then:*

a) The embedding of $H^1(\Omega)$ in $L^2(\Omega)$ is compact.

b) If $\overline{\Omega} \subset \Omega^0$ and Ω^0 is a domain of $\mathbb{R}^N$, then each $u \in H^1(\Omega)$ can be extended to Ω^0 as a function $\widetilde{u} \in H^1(\Omega^0)$ such that

$$\|\widetilde{u}\|_{H^1(\Omega^0)} \leq C\|u\|_{H^1(\Omega)},$$

where C is a constant depending only on Ω.

c) Each function $u \in H^1(\Omega)$ possesses a trace on $\partial\Omega$ belonging to $L^2(\partial\Omega)$ and such that

$$\|u\|_{L^2(\partial\Omega)} \leq C_1\|u\|_{H^1(\Omega)},$$

where C_1 is a constant depending only on Ω.

4.4.1 A proof of Korn's first inequality

For simplicity, we restrict here to the 2-dimensional case and consider smooth vector fields $f : \Omega \to \mathbb{R}^2$ of components $f_1 = f_1(x,y)$, $f_2 = f_2(x,y)$ defined on an open square $\Omega = (0,a) \times (0,a)$. Since Korn's first inequality is proved under the assumption that f has compact support, this particular shape of Ω makes no loss of generality.

Let us consider the mesh of $\overline{\Omega}$ associated to the equidistant division $0 = a_0 < a_1 < a_2 < \cdots < a_n = a$ of $[0,a]$, given by $a_i = ih$, where $h = \frac{a}{n}$. Put also

$$x_{i,j} = f_1((a_i,a_j)), \quad y_{i,j} = f_2((a_i,a_j)), \qquad \forall\, i,j \in \{0,\ldots,n\}.$$

Then the partial derivatives can be approximated by finite differences according to the formulas

$$\frac{\partial f_1}{\partial x_1}(a_i,a_j) \approx \frac{x_{i+1,j}-x_{i-1,j}}{2h}, \quad \frac{\partial f_1}{\partial x_2}(a_i,a_j) \approx \frac{x_{i,j+1}-x_{i,j-1}}{2h}$$
$$\frac{\partial f_2}{\partial x_1}(a_i,a_j) \approx \frac{y_{i+1,j}-y_{i-1,j}}{2h}, \quad \frac{\partial f_2}{\partial x_2}(a_i,a_j) \approx \frac{y_{i,j+1}-y_{i,j-1}}{2h}.$$

In order to make the discretisation more flexible to our purpose we embed Ω into $\Omega^h = (-h, a+h) \times (-h, a+h)$ and extend f_1, f_2 with zero on $\overline{\Omega_h} \setminus \Omega$. It is worth noticing that for n sufficiently large f_1, f_2 vanish on a rectangular corona $\{(x,y) \in \Omega : d((x,y), \partial\Omega) < h\}$. Use the fact that f_1, f_2 have compact supports.

Accordingly, the mesh of $\overline{\Omega}$ is enlarged to a mesh of $\overline{\Omega^h}$ by considering the division points $a_i = ih$, for $i \in \{-1, 0, \dots, n+1\}$. Thus $x_{i,j} = y_{i,j} = 0$, for $i, j \in \{-1, 0, n, n+1\}$.

Taking into account that

$$\int_\Omega f(x,y)dxdy = \lim_{h\to 0} h^2 \sum_{0\le i,j\le n} f(x_i, y_j), \tag{4.4.1}$$

the proof of first Korn's inequality reduces to the proof of its discrete analogue:

$$\begin{aligned}
\sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 &+ \sum_{0\le 1,j\le n} (x_{i,j+1}-x_{i,j-1})^2 \\
&+ \sum_{0\le i,j\le n} (y_{i+1,j}-y_{i-1,j})^2 + \sum_{0\le 1,j\le n} (y_{i,j+1}-y_{i,j-1})^2 \\
&\le 2\Bigg(\sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 + \sum_{0\le 1,j\le n} (y_{i,j+1}-y_{i,j-1})^2 \\
&\qquad + \frac{1}{2} \sum_{0\le i,j\le n} (x_{i,j+1}-x_{i,j-1}+y_{i+1,j}-y_{i-1,j})^2 \Bigg).
\end{aligned} \tag{4.4.2}$$

Equivalently, we have to prove that

$$\begin{aligned}
\sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 &+ (y_{i,j+1}-y_{i,j-1})^2 \\
&+ 2(x_{i,j+1}y_{i+1,j} + x_{i,j-1}y_{i-1,j} - x_{i,j+1}y_{i-1,j} - x_{i,j-1}y_{i+1,j}) \ge 0.
\end{aligned} \tag{4.4.3}$$

Indeed, taking into account that f_1, f_2 vanish on $\overline{\Omega^h} \setminus \Omega$, we have

$$\begin{aligned}
\sum_{0\le i,j\le n} x_{i,j+1}y_{i+1,j} &= \sum_{0\le i,j\le n} x_{i-1,j}y_{i,j-1}, \\
\sum_{0\le i,j\le n} x_{i,j-1}y_{i-1,j} &= \sum_{0\le i,j\le n} x_{i+1,j}y_{i,j+1}, \\
\sum_{0\le i,j\le n} x_{i,j+1}y_{i-1,j} &= \sum_{0\le i,j\le n} x_{i+1,j}y_{i,j-1}, \\
\sum_{0\le i,j\le n} x_{i,j-1}y_{i+1,j} &= \sum_{0\le i,j\le n} x_{i-1,j}y_{i,j+1},
\end{aligned}$$

which yields

$$\sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 + (y_{i,j+1}-y_{i,j-1})^2 + 2(x_{i,j+1}y_{i+1,j} + x_{i,j-1}y_{i-1,j} - x_{i,j+1}y_{i-1,j} - x_{i,j-1}y_{i+1,j}) = \sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j}+y_{i,j+1}-y_{i,j-1})^2 \ge 0.$$

Hence, the proof of first Korn's inequality is complete.

4.4.2 The proof of second Korn's inequality

As in the precedent subsection we restrict ourselves to the two dimensional case, by considering $\Omega = (0,a)\times(0,a)$. According to Theorem 4.4.2 this particular shape of Ω covers the general case.

Under these assumptions, the second Korn's inequality asserts the existence of a positive constant $C = C(\Omega)$ such that,

$$\int_\Omega f_1^2 + f_2^2 + \left(\frac{\partial f_1}{\partial x}\right)^2 + \left(\frac{\partial f_1}{\partial y}\right)^2 + \left(\frac{\partial f_2}{\partial x}\right)^2 + \left(\frac{\partial f_2}{\partial y}\right)^2 \, dxdy \le C\int_\Omega f_1^2 + f_2^2 + \left(\frac{\partial f_1}{\partial x}\right)^2 + \left(\frac{\partial f_2}{\partial y}\right)^2 + \frac{1}{2}\left(\frac{\partial f_1}{\partial y} + \frac{\partial f_2}{\partial x}\right)^2 \, dxdy, \quad (4.4.4)$$

for all $f = (f_1, f_2) \in H^1(\Omega, \mathbb{R}^2)$.

We consider the same kind of meshes for Ω and respectively Ω^h, imposing the following conditions at the boundary points

$$x_{-1,j} = x_{0,j},\ x_{n+1,j} = x_{n,j},\ y_{i,-1} = y_{i,0},\ y_{i,n+1} = y_{i,n} \qquad i,j = 0,\dots,n.$$

The discrete form of the second Korn's inequality is

$$4h^2 \sum_{0\le i,j\le n} (x_{ij}^2 + y_{ij}^2) + \sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 + \sum_{0\le 1,j\le n} (x_{i,j+1}-x_{i,j-1})^2 + \sum_{0\le i,j\le n} (y_{i+1,j}-y_{i-1,j})^2 + \sum_{0\le 1,j\le n} (y_{i,j+1}-y_{i,j-1})^2$$
$$\le C\left(4h^2 \sum_{0\le i,j\le n} (x_{ij}^2 + y_{ij}^2) + \sum_{0\le i,j\le n} (x_{i+1,j}-x_{i-1,j})^2 + \sum_{0\le 1,j\le n} (y_{i,j+1}-y_{i,j-1})^2 + \frac{1}{2}\sum_{0\le i,j\le n} (x_{i,j+1}-x_{i,j-1}+y_{i+1,j}-y_{i-1,j})^2\right). \quad (4.4.5)$$

Note that we have the following estimates

$$\sum_{0\le i,j\le n} x_{i,j+1}y_{i+1,j} = \sum_{0\le i,j\le n} x_{i-1,j}y_{i,j-1} - \sum_{j=0}^{n} x_{-1,j}y_{0,j-1} - \sum_{i=0}^{n} x_{i-1,0}y_{i,-1} + \sum_{j=1}^{n+1} x_{n,j}y_{n+1,j-1} + \sum_{i=1}^{n+1} x_{i-1,n+1}y_{i,n},$$

$$\sum_{0\le i,j\le n} x_{i,j-1}y_{i-1,j} = \sum_{0\le i,j\le n} x_{i+1,j}y_{i,j+1} - \sum_{i=0}^{n} x_{i+1,n}y_{i,n+1} - \sum_{j=0}^{n} x_{n+1,j}y_{n,j+1} + \sum_{i=-1}^{n-1} x_{i+1,-1}y_{i,0} \sum_{j=-1}^{n-1} x_{0,j}y_{-1,j+1},$$

$$\sum_{0\le i,j\le n} x_{i,j+1}y_{i-1,j} = \sum_{0\le i,j\le n} x_{i+1,j}y_{i,j-1} - \sum_{i=0}^{n} x_{i+1,1}y_{i,-1} - \sum_{j=0}^{n} x_{n+1,j}y_{n,j-1} + \sum_{j=1}^{n+1} x_{0,j}y_{-1,j-1} + \sum_{i=-1}^{n-1} x_{i+1,n+1}y_{i,n},$$

$$\sum_{0\le i,j\le n} x_{i,j-1}y_{i+1,j} = \sum_{0\le i,j\le n} x_{i-1,j}y_{i,j+1} - \sum_{j=0}^{n} x_{-1,j}y_{0,j+1} - \sum_{i=0}^{n} x_{i-1,n}y_{i,n+1} + \sum_{i=1}^{n+1} x_{i-1,-1}y_{i,0} + \sum_{j=-1}^{n-1} x_{n,j}y_{n+1,j+1}.$$

Hence we have that

$$\begin{aligned}
&\sum_{0\le i,j\le n} (x_{i,j+1}y_{i+1,j} + x_{i,j-1}y_{i-1,j} - x_{i,j+1}y_{i-1,j} - x_{i,j-1}y_{i+1,j}) \\
&= \sum_{0\le i,j\le n} x_{i-1,j}y_{i,j-1} + x_{i+1,j}y_{i,j+1} - x_{i+1,j}y_{i,j-1} - x_{i-1,j}y_{i,j+1} \\
&\quad + \sum_{j=1}^{n+1}(x_{n,j} - x_{n,j-2})y_{n+1,j-1} + \sum_{i=1}^{n+1} x_{i-1,n+1}(y_{i,n} - y_{i-2,n}) \\
&\quad - \sum_{i=0}^{n}(x_{i+1,n} - x_{i-1,n})y_{i,n+1} - \sum_{j=0}^{n} x_{n+1,j}(y_{n,j+1} - y_{n,j-1}) \\
&\quad + \sum_{i=0}^{n}(x_{i+1,0} - x_{i-1,0})y_{i,-1} - \sum_{j=1}^{n+1}(x_{0,j} - x_{0,j-2})y_{-1,j-1} \\
&\quad + \sum_{j=0}^{n} x_{-1,j}(y_{0,j+1} - y_{0,j-1}) - \sum_{i=1}^{n+1} x_{i-1,-1}(y_{i,0} - y_{i-2,0}).
\end{aligned}$$

Thus (4.4.5) is equivalent to

$$\begin{aligned}
&4h^2(C-1)\sum_{0\le i,j\le n}(x_{ij}^2+y_{ij}^2)+\frac{C}{2}\sum_{0\le i,j\le n}(x_{i+1,j}-x_{i-1,j}+y_{i,j+1}-y_{i,j-1})^2\\
&+\left(\frac{C}{2}-1\right)\sum_{0\le i,j\le n}\left((x_{i+1,j}-x_{i-1,j})^2+(x_{i,j+1}-x_{i,j-1})^2+(y_{i+1,j}-y_{i-1,j})^2+(y_{i,j+1}-y_{i,j-1})^2\right)\\
&\qquad+C\sum_{j=0}^{n}(x_{n,j+1}-x_{n,j-1})y_{n+1,j}+C\sum_{i=0}^{n}x_{i,n+1}(y_{i+1,n}-y_{i-1,n})\\
&\qquad-C\sum_{i=0}^{n}(x_{i+1,n}-x_{i-1,n})y_{i,n+1}-C\sum_{j=0}^{n}x_{n+1,j}(y_{n,j+1}-y_{n,j-1})\\
&\qquad+C\sum_{i=0}^{n}(x_{i+1,0}-x_{i-1,0})y_{i,-1}-C\sum_{j=0}^{n}(x_{0,j+1}-x_{0,j-1})y_{-1,j}\\
&\qquad\qquad+C\sum_{j=0}^{n}x_{-1,j}(y_{0,j+1}-y_{0,j-1})-C\sum_{i=0}^{n}x_{i,-1}(y_{i+1,0}-y_{i-1,0})\ge 0.
\end{aligned}$$

By using a discrete type Green formula for the last 8 terms (provided by the discretization of a curviligne integral on the boundary of the square) we have to prove that

$$\begin{aligned}
&4h^2(C-1)\sum_{0\le i,j\le n}(x_{ij}^2+y_{ij}^2)+\frac{C}{2}\sum_{0\le i,j\le n}(x_{i+1,j}-x_{i-1,j}+y_{i,j+1}-y_{i,j-1})^2\\
&+\left(\frac{C}{2}-1\right)\sum_{0\le i,j\le n}\left((x_{i+1,j}-x_{i-1,j})^2+(x_{i,j+1}-x_{i,j-1})^2+(y_{i+1,j}-y_{i-1,j})^2+(y_{i,j+1}-y_{i,j-1})^2\right)\\
&\qquad-C\sum_{0\le i,j\le n}((x_{i+1,j}-x_{i-1,j})(y_{i,j+1}-y_{i,j-1})-(x_{i,j+1}-x_{i,j-1})(y_{i+1,j}-y_{i-1,j}))\ge 0.
\end{aligned}$$

Since we have that

$$\begin{aligned}
&C\sum_{0\le i,j\le n}((x_{i+1,j}-x_{i-1,j})(y_{i,j+1}-y_{i,j-1})-(x_{i,j+1}-x_{i,j-1})(y_{i+1,j}-y_{i-1,j}))\\
&\le\frac{C}{2}\sum_{0\le i,j\le n}\left((x_{i+1,j}-x_{i-1,j})^2+(x_{i,j+1}-x_{i,j-1})^2+(y_{i+1,j}-y_{i-1,j})^2+(y_{i,j+1}-y_{i,j-1})^2\right),
\end{aligned}$$

it is sufficiently to prove that

$$\begin{aligned}
&4h^2(C-1)\sum_{0\le i,j\le n}(x_{ij}^2+y_{ij}^2)+\frac{C}{2}\sum_{0\le i,j\le n}(x_{i+1,j}-x_{i-1,j}+y_{i,j+1}-y_{i,j-1})^2\\
&-\sum_{0\le i,j\le n}\left((x_{i+1,j}-x_{i-1,j})^2+(x_{i,j+1}-x_{i,j-1})^2+(y_{i+1,j}-y_{i-1,j})^2+(y_{i,j+1}-y_{i,j-1})^2\right)\ge 0.
\end{aligned}$$

Notice that for any $C\ge 2$ depending on norm of the function and of the norm of the gradient of the function we can deduce that the previous inequality hold. This dependence will be a subject of a future work.

4.5 Weak majorization in global NPC spaces

In this section we introduce the concept of weak majorization in the framework of some spaces with a curved geometry (such as the global NPC spaces). The key point is given by a perturbed barycenter concept, defined by a minimization argument. Then, by using a generalized concept of majorization in a global NPC space (for which we can prove Hardy-Littlewood-Pólya's majorization theorem) we can study the concept of weak majorization in this context. Note that this part is inspired from [120].

In the following we study a perturbed barycenter concept in a space with global nonpositive curvature, by using a minimization procedure. This notion of perturbed barycenter is the main ingredient used to introduce the concept of weak majorization in global NPC spaces.

Let $\alpha > 0$, $y \in M$ and let μ be a probability measure defined on a global NPC space (M, d), such that $\mu \in \mathcal{P}_2(M)$.

We shall make use of the process of augmentation of a discrete probability measure $\mu = \sum_{i=1}^{n} \mu_i \delta_{x_i}$ by adding a new point y in its support. This consists in choosing arbitrarily a positive number $\alpha > 0$ followed by a reallocation of the mass of μ to the point y, in other words, by replacing μ by the probability measure

$$\mu_{y,\alpha} = \sum_{i=1}^{n} \frac{\mu_i}{1+\alpha}\delta_{x_i} + \frac{\alpha}{1+\alpha}\delta_y.$$

We define the barycenter of the *augmented measure* $\mu_{z,\alpha}$ as the point given by the formula

$$\begin{aligned}
\mathrm{bar}(\mu_{y,\alpha}) &= \arg\min_{z\in M} \frac{1}{2}\int_M \left(d^2(z,x) + \alpha d^2(z,y)\right) d\mu(x) \\
&= \arg\min_{z\in M} \frac{1}{2}(\sum_{i=1}^{n} \mu_i (d^2(z,x_i) + \alpha d^2(z,y))
\end{aligned}$$

and Jensen's inequality associated to it asserts that

$$f(\mathrm{bar}(\mu_{z,\alpha})) \le \frac{1}{1+\alpha}\sum_{i=1}^{n} \mu_i f(x_i) + \frac{\alpha}{1+\alpha} f(y) \tag{4.5.1}$$

for any lower semicontinuous convex function $f : M \to \mathbb{R}$.

In the particular case when $M = \mathbb{R}^N$ and $\mu_1 = \cdots = \mu_n = \alpha = 1/n$, we have

$$\mathrm{bar}(\mu) = \frac{\mathbf{x}_1 + \cdots + \mathbf{x}_n}{n} \text{ and } \mathrm{bar}(\mu_{\mathbf{y},\alpha}) = \frac{\mathbf{x}_1 + \cdots + \mathbf{x}_n + \mathbf{y}}{n+1},$$

so that the inequality (4.5.1) reads as

$$f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_n + \mathbf{y}}{n+1}\right) \le \frac{n}{n+1} f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_n}{n}\right) + \frac{1}{n+1} f(\mathbf{y}).$$

The extension of the concept of weak majorization to the context of global NPC spaces is based on the process of augmentation described in the previous sentences.

Definition 4.5.1. *The relation of weak majorization,*

$$\sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec_{NPCw} \sum_{j=1}^{n} \mu_j \delta_{y_j}, \tag{4.5.2}$$

means the existence of a vector $z \in M$, of a string $\alpha_1, ..., \alpha_n$ of nonnegative numbers and also of an $m \times n$-dimensional stochastic on rows matrix $A = (a_{ij})_{i,j}$ that verify the following two conditions:

$$\mu_j = \sum_{i=1}^{m} a_{ij}\lambda_i, \quad j = 1, ..., n \tag{4.5.3}$$

and

$$x_i = \arg\min_{x\in M} \frac{1}{2}\sum_{j=1}^{n} a_{ij}(d^2(x, y_j) + \alpha_i d^2(x, z)), \quad i = 1, ..., m. \tag{4.5.4}$$

When $\alpha_1 = \cdots = \alpha_n = 0$, the definition of weak majorization reduces to that of majorization.

The existence and uniqueness of the solution for problem (4.5.4), when $\alpha = 0$, is assured by the fact that the objective functions are uniformly convex and positive. See [80, Section 3.1], or [163, Proposition 1.7]. We prove now the existence and uniqueness of the perturbed barycenter for $\alpha > 0$.

Proposition 4.5.1. *Let $\alpha > 0$, $y \in M$ and let μ be a probability measure defined on a global NPC space M, such that $\mu \in \mathcal{P}_2(M)$. Let us consider*

$$T_y(z) = \int_M d^2(z, x) + \alpha d^2(z, y) d\mu(x) \qquad (z \in M).$$

Then, T_y has a unique minimizer on M.

Proof. In [163, Proposition 2.3] has been proved that the function $z \to d^2(z, x)$ is uniformly convex. Consequently, denoting by z_t the joining geodesic of the points z_0, z_1 we have that

$$T_y(z_t) = \int_M d^2(z_t, x) + \alpha d^2(z_t, y) d\mu(x)$$

$$\leq (1-t)\int_M d^2(z_0, x) + \alpha d^2(z_0, y) d\mu(x) + t\int_M d^2(z_1, x) + \alpha d^2(z_1, y) d\mu(x)$$
$$-t(1-t)(1+\alpha)d^2(z_0, z_1)$$
$$\leq (1-t)T_y(z_0) + tT_y(z_1) - t(1-t)d^2(z_0, z_1),$$

which precisely means that T_y is uniformly convex.

Moreover, the continuity of the distance function $z \to d^2(z, x)$ implies the continuity of T_y, hence by using in addition the uniform convexity, we deduce that T_y has a unique minimizer. □

The agreement of the weak majorization concept in different settings makes the objective of the following result.

Theorem 4.5.1. *If $\mu = \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$ and $\nu = \frac{1}{N}\sum_{j=1}^{n}\delta_{y_j}$ are two discrete probability measures on $\mathbb{R}$ such that $x_1 \geq x_2 \geq \cdots \geq x_n$ and $y_1 \geq y_2 \geq \cdots \geq y_n$, then*

$$\mu \prec_{NPCw} \nu \text{ if and only if } \mu \prec_{HLPw} \nu.$$

Proof. Suppose that $\mu \prec_{NPCw} \nu$. According to Definition 4.5.1, there exist a real number z, a string $\alpha_1, \ldots, \alpha_n$ of nonnegative numbers and also an $n \times n$-dimensional doubly stochastic matrix $A = (a_{ij})_{i,j}$ such that

$$x_i = \arg\min_{x\in\mathbb{R}} \frac{1}{2}\sum_{j=1}^{n} a_{ij}((x-y_j)^2 + \alpha_i(x-z)^2), \quad \text{for } i = 1, \ldots, n,$$

equivalently,

$$x_i = \frac{\sum_{j=1}^{n} a_{ij}y_j + \alpha_i z}{1+\alpha_i} \qquad \text{for } i = 1, \ldots, n,$$

which implies

$$\varphi(x_i) \leq \frac{1}{1+\alpha_i}\sum_{j=1}^{n} a_{ij}\varphi(y_j) + \frac{\alpha_i}{1+\alpha_i}\varphi(z).$$

Now, by choosing

$$z = \min_i \sum_{j=1}^{n} \min\{y_j, 0\},$$

the following relation holds

$$x_i = \arg\min_{x\in M} \frac{1}{2}\sum_{j=1}^{n} a_{ij}((x-y_j)^2 + \alpha_i(x-z)^2), \quad i = 1, \ldots, m, \tag{4.5.5}$$

where $(\alpha_i)_i$ are some positive real numbers. Now, using the fact that $y \leq 0$, a simple calculus gives that

$$x_i \leq \frac{\sum_{j=1}^{n} a_{ij}y_j}{1+\alpha_i} \leq \sum_{j=1}^{n} a_{ij}y_j \qquad (i = 1, \ldots, n).$$

If we denote by $\overline{x_i} = \sum_{j=1}^{n} a_{ij}y_j$, since $(a_{ij})_{i,j}$ is doubly stochastic it follows that $(\overline{x_1}, \ldots, \overline{x_n}) \prec (y_1, \ldots, y_n)$. Hence the following sequence of inequalities holds

$$x_1 \leq \overline{x_1} \leq y_1,$$

$$x_1 + x_2 \leq \overline{x_1} + \overline{x_2} \leq y_1 + y_2,$$

$$\ldots$$

$$x_1 + \cdots + x_n \leq \overline{x_1} + \cdots + \overline{x_n} = y_1 + \cdots + y_n.$$

Suppose now that $\mu \prec_{HLPw} \nu$. By replacing μ and ν respectively by

$$\bar{\mu} = \frac{1}{n+1}\sum_{i=1}^{n+1}\delta_{x_i} \text{ and } \bar{\nu} = \frac{1}{n+1}\sum_{i=1}^{n+1}\delta_{y_i}$$

where $x_{n+1} = \min\{x_1, \ldots, x_n, y_1, \ldots, y_n\} = y_n$ and $y_{n+1} = \sum_{i=1}^{n+1} x_i - \sum_{i=1}^{n} y_i$ we have

$$\bar{\mu} \prec_{HLP} \bar{\nu}$$

and this yields the existence of a doubly stochastic matrix $A = (a_{ij})_{i,j}$ such that

$$x_i = \sum_{j=1}^{n+1} a_{ij} y_j \qquad \text{for } i = 1, \ldots, n+1.$$

Therefore

$$x_i = \arg\min_{x\in\mathbb{R}} \frac{1}{2}\sum_{j=1}^{n} a_{ij}((x-y_j)^2 + \alpha_{in+1}(x-y_{n+1})^2), \quad \text{for } i = 1, \ldots, n$$

and

$$\sum_{j=1}^{n} a_{ij} \leq 1 \quad \text{for } i = 1, \ldots, n,$$

and the proof is finished. □

The relevance of the study of such perturbed minimizers can be viewed in different ways. One of them, consist in the fact that frequently we need to consider a minimizer for perturbed functionals, and this minimizer is the solution of a partial differential equation. On the other hand, the added norm term can be seen as the distance to a fixed point, which in optimization theory has the meaning to add a fix point which will be taking into account in our minimization problem.

Based on the above definition, we are able to give a nice version of Hardy-Littlewood-Polya's inequality and to extend classical results from majorization theory. For more details, see [126]. On the other hand, the concept of weak majorization is a completely open an very interesting problem in such a general settings. We are now in position to prove an important consequence, which gives a discrete version of Jensen's inequality in the context of global NPC spaces.

Theorem 4.5.2. *(A discrete Jensen's inequality) Let $\varphi : M \to \mathbb{R}$ be a lower semicontinuous convex function and $\mu \in \mathcal{P}_2(M)$, where μ is a discrete probability measure $\mu = \sum_{i=1}^n \mu_i \delta_{x_i}$, with $\sum_{i=1}^n \mu_i = 1$. The following inequality holds*

$$\varphi(\mathrm{bar}(\mu;\alpha)) \leq \frac{1}{1+\alpha}\sum_{i=1}^{n} \mu_i \varphi(x_i) + \frac{\alpha}{1+\alpha}\varphi(y). \tag{4.5.6}$$

Proof. Firstly, note that

$$\begin{aligned}
\mathrm{bar}(\mu;\alpha) &= \arg\min_{z\in M} \frac{1}{2}\sum_{i=1}^{n} \mu_i \left(d^2(z,x_i) + \alpha d^2(z,y)\right) \\
&= \arg\min_{z\in M} \frac{1+\alpha}{2}\sum_{i=1}^{n} \left(\frac{\mu_i}{1+\alpha} d^2(z,x_i) + \frac{\alpha\mu_i}{1+\alpha} d^2(z,y)\right) \\
&= \arg\min_{z\in M} \frac{1}{2}\sum_{i=1}^{2n} \lambda_i d^2(z,y_i),
\end{aligned}$$

where $\lambda_i = \frac{\mu_i}{1+\alpha}$, $\lambda_{i+n} = \frac{\alpha\mu_i}{1+\alpha}$, $y_i = x_i$ and $y_{i+n} = y$, for each $i = 1, \ldots, n$.

Hence, we have proved that

$$\mathrm{bar}(\mu;\alpha) = \mathrm{bar}(\lambda), \text{ where } \lambda = \sum_{i=1}^{2n} \mu_i \delta_{y_i}.$$

By using classical Jensen's inequality from [163, Theorem 6.2], we have that

$$\varphi\left(\mathrm{bar}(\mu;\alpha)\right)=\varphi\left(\mathrm{bar}(\lambda)\right)\leq\sum_{i=1}^{2n}\lambda_i\varphi(y_i)$$

$$=\sum_{i=1}^{n}\frac{\mu_i}{1+\alpha}\varphi(x_i)+\sum_{i=1}^{n}\frac{\alpha\mu_i}{1+\alpha}\varphi(y)=\frac{1}{1+\alpha}\sum_{i=1}^{n}\mu_i\varphi(x_i)+\frac{\alpha}{1+\alpha}\varphi(y),$$

and the proof of (4.5.6) ends. □

Another consequence consist of a a generalization of Sherman's majorization results.

Theorem 4.5.3. *(HLP-Sherman) Let $\varphi : M \to \mathbb{R}$ be a convex function and let*

$$\sum_{i=1}^{m}\lambda_i\delta_{x_i}\prec_\alpha\sum_{j=1}^{n}\mu_j\delta_{y_j},$$

as in (4.5.2)*. Then the following inequality holds*

$$\sum_{i=1}^{m}\lambda_i(1+\alpha_i)\varphi(x_i)\leq\sum_{j=1}^{n}\mu_j\varphi(y_j)+\varphi(y)\sum_{i=1}^{m}\lambda_i\alpha_i. \tag{4.5.7}$$

Proof. Taking into account (4.5.4) and (4.5.6) we have that

$$\varphi(x_i)\leq\frac{1}{1+\alpha_i}\sum_{j=1}^{n}a_{ij}\varphi(y_j)+\frac{\alpha_i}{1+\alpha_i}\varphi(y)\quad(i=1,\ldots,m.)$$

Hence, it follows that

$$\sum_{i=1}^{m}\lambda_i(1+\alpha_i)\varphi(x_i)\leq\sum_{i=1}^{m}\sum_{j=1}^{n}\lambda_i a_{ij}\varphi(y_j)+\sum_{i=1}^{m}\lambda_i\alpha_i\varphi(y)$$

$$=\sum_{j=1}^{n}\mu_j\varphi(y_j)+\varphi(y)\sum_{i=1}^{m}\lambda_i\alpha_i,$$

and the proof is finished. □

Remark 26. *Note that* (4.5.7) *gives in fact that*

$$\sum_{i=1}^{m}\lambda_i\delta_{x_i}\prec\sum_{j=1}^{n}\frac{\mu_j}{1+\alpha}\delta_{y_j}+\frac{\alpha}{1+\alpha}\delta_y.$$

We have now introduced a variational technique for the definition of weak majorization. At least for our knowledge, this is the first time when the weak majorization is written in term of minimizers. Moreover, the point $y=0$ is essential in the theory of weak majorization on $\mathbb{R}$, which remind us about a nice characterization of weak majorization in terms of convex and nondecreasing functions. In fact, the definition of a nondecreasing function is essentially based on the distance to the origin (the same point which appears in the above assertion).

Thus, we are now in position to state an important consequence which prove that our definition is natural and well posed. We present an extension of Tomic-Weyl result for weak majorization in global NPC spaces.

Theorem 4.5.4. *Let us consider that*

$$\frac{1}{n}\sum_{i=1}^{1}\delta_{x_i} \prec_* \frac{1}{n}\sum_{j=1}^{n}\delta_{y_j}. \tag{4.5.8}$$

Let $\varphi : M \to \mathbb{R}$ be a convex function which verifies

$$\varphi(x_i) \geq \varphi(y) \qquad (i = 1, \dots, n),$$

where $y \in M$ is the point appearing in the definition of (4.5.8)*. Then the following inequality holds*

$$\sum_{i=1}^{n}\varphi(x_i) \leq \sum_{i=1}^{n}\varphi(y_i). \tag{4.5.9}$$

Proof. From (4.5.8) we infer the existence of an $y \in M$, $\alpha \in \mathbb{R}_+^n$ and a matrix $(\lambda_{ij})_{i,j}$ such that

$$x_i = \arg\min_{z\in M}\frac{1}{2}\sum_{j=1}^{n}\lambda_{ij}\left(d^2(z,y_j) + \alpha_i d^2(z,y)\right) \quad i = 1,\dots,m.$$

From (4.5.6) it follows that

$$\varphi(x_i) \leq \frac{1}{1+\alpha_i}\sum_{j=1}^{n}\lambda_{ij}\varphi(y_j) + \frac{\alpha_i}{1+\alpha_i}\varphi(y),$$

$$\sum_{i=1}^{n}(1+\alpha_i)\varphi(x_i) \leq \sum_{j=1}^{n}\varphi(y_j) + \varphi(y)\sum_{i=1}^{n}\alpha_i,$$

$$\sum_{i=1}^{n}\varphi(x_i) + \sum_{i=1}^{n}\alpha_i\varphi(x_i) \leq \sum_{j=1}^{n}\varphi(y_j) + \varphi(y)\sum_{i=1}^{n}\alpha_i,$$

hence, since $\varphi(x_i) \geq \varphi(y)$ we obtain that the conclusion. □

Note that, the hypothesis $\varphi(x_i) \geq \varphi(y)$ is nothing else than the nondecreasing property of a function in $\mathbb{R}_+$.

In this context, we can consider the perturbed minimizers, our α-majorization, into the spaces with global nonpositive curvature. The existence and uniqueness of such perturbed minimizers in a global NPC space, is a difficult task but using the above remarks we are able to be succesfully implemented (see a detailed approach of the notion of barycenter Sturm [163, Proposition 1.7]).

The results from this section can be also extended in the framework of nonpositive weights, but this will be the subject of a future work.

Bibliography

[1] S. Abramovich, S. Ivelić, and J. E. Pečarić, *Improvement of Jensen–Steffensen's inequality for superquadratic functions*, Banach J. Math. Anal. **4** (2010), 159–169.

[2] S. Abramovich, *New inequalities related to superquadratic functions*, Aequ. Math. **96** (2022), 201–2019.

[3] S. Abramovich, *Refinements of Jensen's inequality by uniformly convex functions*, Aequ. Math. **97** (2023), 75–88.

[4] K. Aguilar, Á. Chávez, S. Garcia, and J. Volčič, *Norms on complex matrices induced by complete homogeneous symmetric polynomials*, Bull. London Math. Soc. **54** (2022), 2078–2100.

[5] A. A. Ahmadi, Etienne de Klerk, and G. Hall, *Polynomial norms*, https://arxiv.org/abs/1704.07462.

[6] S. B. Alexander, R. L. Bishop, *The Hadamard-Cartan theorem in locally convex metric spaces*, Enseign. Math. (2) **36** (1990), no. 3-4, 309-320.

[7] A. D. Alexandrov, *A theorem on triangle in a metric space and some applications*, Trudy Math. Inst. Steklov **38**, 5-23, 1951. (Russian; translated into German and combined with more material in [8])

[8] A. D. Alexandrov, *Über eine Verallgemeinerung der Riemannschen Geometrie*, Schr. Forschungsinst. Math. Berlin **1** (1957), 33-84.

[9] C. D. Aliprantis, R. Tourky, *Cones and Duality*, Graduate Studies in Mathematics **84**, American Mathematical Society, Providence, RI, 2007.

[10] H. Amann, *Multiple positive fixed points of asymptotically linear maps*, J. Functional Analysis **17** (2) (1974), 174–213.

[11] V. Anantharam, P. Viswanath, *Optimal sequence for CDMA under colored noise: A Schur-saddle function property*, IEEE Trans. Inform. Theory, vol 48 (6) (2002), 1295-1318.

[12] W. N. Anderson, T. D. Morley, and G. E. Trapp. *Symmetric function means of positive operators*, Linear algebra and its applications, **60** (1984), 129–143.

[13] W. Arveson and R. Kadison, *Diagonals of self-adjoint operators*, Contemp. Math. **414**, pp. 247-263. Amer. Math. Soc. Publ. Providence, R. I., 2006.

[14] A. Bahrouni, V. Rădulescu, D. Repovš, *A weighted anisotropic variant of the Caffarelli-Kohn-Nirenberg inequality and applications*, Nonlinearity **31** (2018), 1518–1534.

[15] M. K. Bakula, *Jensen-Steffensen inequality for strongly convex functions*, J. Inequal. Appl. **306** (2018), 306.

[16] W. Ballmann, M. Gromov, V. Schroeder, *Manifolds of nonpositive curvature*, Progress in Mathematics **61**. Birkháuser Boston Inc., Boston, MA, 1985.

[17] W. Ballmann, *Lectures on spaces with nonpositive curvature*, DMV Seminar Band 25, Birkhäuser, Basel, 2005.

[18] A. Barvinok, *Approximating a norm by a polynomial*, Geometric Aspects of Functional Analysis, Springer, 2003, pp. 20–26.

[19] A. Barvinok and E. Veomett, *The computational complexity of convex bodies, Surveys on discrete and computational geometry*, Contemporary Mathematics **453** (2008), 117–137.

[20] A. I. Barvinok, *Low rank approximations of symmetric polynomials and asymptotic counting of contingency tables*, https://arxiv.org/abs/math/0503170.

[21] J. Bebernes, D. Eberly, *Mathematical Problems from Combustion Theory*, Springer-Verlag, New York, 1989.

[22] M. Bencze, C.P. Niculescu and F. Popovici, *Popoviciu's inequality for functions of several variables*, J. Math. Anal. Appl. **365** (1) (2010), 399–409.

[23] G. Bennett, *Summability matrices and random walk*, Houston J. Math **28** (2002), no. 4, 865–898.

[24] R. Bhatia, *Positive definite matrices*, Princeton University Press, 2007.

[25] R. Bhatia, *Matrix analysis*, Springer-Verlag, 1997.

[26] M. Bîrsan, P. Neff, J. Lankeit, *Sum of squared logarithms - an inequality relating positive definite matrices and their matrix logarithm*, J. Inequal. Appl. (2013), no. 168.

[27] J. Borcea, *Equilibrium points of logarithmic potentials induced by positive charge distributions. I. Generalized de Bruijn-Springer relations*, Trans. Amer. Math. Soc. 359 (2007), 3209–3237.

[28] J. Borwein, A. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples (2 ed.)*, Springer. ISBN 978-0-387-29570-1, (2006).

[29] A. Böttcher, S. R. Garcia, M. Omar, and C. O'Neill, *Weighted means of B-splines, positivity of divided differences, and complete homogeneous symmetric polynomials*, Linear Algebra Appl. **608** (2021), 68–83.

[30] S. I. Bradanovic, *Sherman's inequality and its converse for strongly convex functions with applications generalized f-divergences*, Turkish Journal of Mathematics **43** (6) (2019), 2680–2696.

[31] S. I. Bradanović, N. Latif, Pečarić, and J. Pečarić, *Sherman's and related inequalities with applications in information theory*, J Inequal Appl. 2018, (2018) (1), article 98.

[32] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive cuvature*, Grundlehren der Mathematischen Wissenschaften **319**, Springer-Verlag, Berlin, 1999.

[33] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics **33**. American Mathematical Society, Providence, RI, 2001.

[34] L. A. Caffarrelli, A. Friedman, *Blow-up of solutions of nonlinear heat equation*, J. Math. Anal. Appl. **129** (1988) 409–419.

[35] R. S. Cantrell, C. Cosner, *Diffusive logistic equation with indefinite weights: population models in disrupted environments II*, SIAM J. Math. Anal. 22 **12** (1991) 1043–1064.

[36] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Gauthiers-Villars, Paris, 1928; 2nd éd., 1946.

[37] Á. Chávez, S. Garcia, and J. Hurley, *Norms on Complex Matrices Induced by Random Vectors*, Can. Math. Bull. (2022), 1–19.

[38] Y. Chen, *Blow-up for a system of heat equations with nonlocal sources and absorptions*, Comput. Math. Appl. **48** (2004), 361–372.

[39] Y. M. Chu, G. D. Wang and X. M. Zhang, *The Schur multiplicative and harmonic convexities of the complete symmetric function*, Mathematische Nachrichten **284** (2011), no. 5-6, 653–663.

[40] Y. M. Chu, W. F. Xia, T. H. Zhao, *Schur convexity for a class of symmetric functions*, Sci. China Math. **53** (2010), no. 2, 465–474.

[41] P. G. Ciarlet, *Mathematical elasticity*, Vol. I: Three-dimensional elasticity, North-Holland Publishing Co., Amsterdam, 1988.

[42] P. G. Ciarlet, *On Korn's inequality*, Chinese Ann. Math., Ser B **31** (2010), 607–618.

[43] P.G. Ciarlet, P. Ciarlet Jr., *Linearized elasticity and Korn's inequality revisited*, Mathematical Models and Methods in Applied Sciences (2004).

[44] P.G. Ciarlet, P. Ciarlet Jr., *Another approach to linearized elasticity and a new proof of Korn's inequality*, Math. Models Methods Appl. Sci., **15** (2005), pp. 259–271.

[45] N. Cindea, S. Micu, and I. Rovenţa, *Boundary controllability for finite-difference semi-discretizations of a clamped beam equation*, SIAM J Control Optim. **55** (2017), 785–817.

[46] A. Clim, *Some inequalities for convexifiable function with applications*, P. Romanian Acad. A **11** (3) (2010), 218–223.

[47] P. Czinder, Z. Pales, *An extension of the Hermite-Hadamard inequality and an application for Gini and Stolarsky means*, JIPAM **5** (2004), Issue 2, Article no.42.

[48] A. Daniilidis, P. Georgiev. *Approximate convexity and submonotonicity*, J. Math. Anal. Appl. **291** (1) (2004), 292–301.

[49] J. M. Danskin, *Dresher's inequality*, American Mathematical Monthly, **59** (10):687–688, 1952.

[50] L. Desvillettes, C. Villani, *On a variant of Korn's inequality arising in statistical mechanics*. A tribute to J.L. Lions, ESAIM Control Optim. Calc. Var. **8** (2002) 603–619.

[51] S. S. Dragomir and N. M. Ionescu, *Some converse of Jensen's inequality and applications*, Rev. Anal. Numér. Théor. Approx. **23** (1994), no. 1, 71–78.

[52] M. Dragomirescu, C. Ivan, *Convexity-like inequalities for averages in a convex set*, Aequationes Mathematicae **45** (1993), 179–194.

[53] M. Dresher et al. *Moment spaces and inequalities*, Duke mathematical journal, **20** (2) (1953), 261–271.

[54] G. Duvaut, J.L. Lions, *Inequalities in Mechanics and Physics*, Springer, New York, 1976.

[55] J. Eells, B. Fuglede, *Harmonic maps between Riemannian polyhedra*, Cambridge Tracts in Mathematics, **142**. Cambridge University Press, Cambridge, 2001.

[56] A. M. Fink, *A best possible Hadamard inequality*, Math. Inequal. Appl. **1** (1998), no. 2, 223–230.

[57] K. O. Friedrichs, *On the boundary value problems of the theory of elasticity and Korn's inequality*, Ann. of Math. **48** (2) (1947) 441–471.

[58] M. Fuchs, *Korn inequalities in Orlicz spaces*, Irish Math. Soc. Bull. **65** (2010), 5–9.

[59] J. Furter, M. Grinfield, *Local vs. nonlocal interactions in populations dynamics*, J. Math. Biol. **27** (1989), 65—80.

[60] S. R. Garcia, M. Omar, C. O'Neill, and S Yih, *Factorization length distribution for affine semigroups II: asymptotic behavior for numerical semigroups with arbitrarily many generators*, J. Combin. Theory Ser. A **178** (2021), 105358, 34.

[61] I. D. Ghiba, C. P. Niculescu, and I. Rovenţa, *A finite difference approach of Korn's inequalities*, work in progress.

[62] J. Gobert, *Une inégalité fondamentale de la theorie de l'élasticicité*, Bull. Soc. Roy. Sci. Liege **3** (4) (1962), 182–191.

[63] M. Gromov, *Structures métriques pour les variétés Riemanniennes*, Rédigé par J. Lafontaine et P. Pansu, Cedic/Fernand Nathan, 1981.

[64] M. Gromov, *Hyperbolic groups*, Essays in group theory (S.M.Gersten, ed.) MSRI Publ. 8, Springer-Verlag, Berlin-Heidelberg-New York, 75-264, (1987).

[65] G. Gruss, *Über das maximum des absoluten betrages von* $\frac{1}{b-a}\int_a^b f(x)g(x)dx - \frac{1}{b-a}\int_a^b f(x)dx \frac{1}{b-a}\int_a^b g(x)dx$, Math. Z. **39** (1935), no. 1, 215–226.

[66] K. Guan, *Some properties of a class of symmetric functions*, J. Math. Anal. Appl. **336** (2007), 70-80.

[67] S. Hanly, D. Tse, *Power Control and Capacity of Spread Spectrum Wireless Networks*, Automatica **35** (12) (1999), 1987–2012.

[68] G.H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed., Cambridge University Press, 1952. Reprinted 1988.

[69] G. H. Hardy, J. E. Littlewood and G. Pólya, *Some simple inequalities satisfied by convex functions*, Messenger Math. **58** (1929), 145–152.

[70] C. O. Horgan, *Korn's inequalities and their applications in continuum mechanics*, SIAM Review **37** (1995) No. 4, 491–511.

[71] A. Horn, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math. **76** (1954), no. 3, 620–630.

[72] A. Horn, *Eigenvalues of sums of Hermitian matrices*, Pacific J. Math. **12** (1962), 225–241.

[73] O. Hólder, *Über einen Mittelwertsatz*, Nachr. Ges. Wiss. Goettingen, 1889, 38–47.

[74] D. B. Hunter, *The positive-definiteness of the complete symmetric functions of even order*, Math. Proc. Cambridge Philos. Soc. **82** (1977) (2), 255–258.

[75] D. H. Hyers and S. M. Ulam, *Approximately convex functions*, Proc. Amer. Math. Soc. **3** (1952), 821–828.

[76] A. E. Ingham, *Some trigonometrical inequalities with applications to the theory of series*, Math. Z. **41** (1936), 367–369.

[77] S. Jaffard, and S. Micu, *Estimates of the constants in generalized Ingham's inequality and applications to the control of the wave equation,* Asymptot. Anal. **28** (2001), 181–214.

[78] S. Jaffard, M. Tucsnak, and E. Zuazua, *On a theorem of Ingham*, J. Fourier Anal. Appl. **3** (1997), 577–582.

[79] J. Jost, *Equilibrium maps between metric spaces*, Calc. Var. Partial Differential Equationes **2** (1994), 173–204.

[80] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997.

[81] J. P. Kahane, *Pseudo-Périodicité et Séries de Fourier Lacunaires*, Ann. Scient. Ec. Norm. Sup. **37** (1962), 93–95.

[82] M. A. Khan, *Majorization theorem for convexifiable functions*, Math. Commun. **18** (2013), 61–65.

[83] M. A. Khan, M. Niezgoda, and J. E. Pečarić, *On a refinement of the majorisation type inequality*, Demonstr. Math. **44** (2011), 49–57.

[84] A. A. Klyachko, *Stable brundles, representation theory and Hermitian operators*, Selecta Math. (N.S.) **4** (1998), no. 3, 419–445.

[85] A. Knutson and T. Tao, *The honeycomb model of $GL_n(\mathbf{C})$ tensor products. I. Proof of the saturation conjecture*, J. Amer. Math. Soc. **12** (1999), no. 4, 1055–1090.

[86] V.A. Kondratiev and O.A. Oleinik, *On Korn's inequalities*, C. R. Acad. Sci. Paris Sér. I Math. 308 (1989), 483–487.

[87] N. Korevaar, R. Schoen, *Sobolev spaces and harmonic maps for metric space targets*, Comm. Anal. Geom. **1** (1993), 561–569.

[88] A. Korn, *Die Eigenschwingungen eines elastischen Körpers mit ruhender Oberfläche*, Akad. Wiss. München, Math.-Phys. Kl. Ber., vol. 36, 1906, 51–401.

[89] A. Korn, *Solution general du probleme d'equilibre dans la theorie de l'elasticite*, Annales de la Faculte de Sciences de Toulouse, **10** (1908), 705–724.

[90] A. Korn, *Ueber einige Ungleichungen, welche in der Theorie der elastischen und elektrischen Schwingungen eine Rolle spielen*, Bulletin internationale de l'Academie de Sciences de Cracovie, **9** (1909), 705–724.

[91] J. L. Lagrange, *Sur une nouvelle proprieté du centre de gravité*, Nouveaux Mémoirs de l'Académie Royale de Berlin 1783; see Oeuvres de Lagrange, vol. **5**, 535-540, Gauthier-Villars, Paris, 1870.

[92] J. D. Lawson, Y. Lim, *Monotonic properties of the least squares mean*, Math. Ann. **351** (2011), 267–279.

[93] G. M. Lăchescu, M. Malin, and I. Rovenţa: *New Versions of Uniformly Convex Functions via Quadratic Complete Homogeneous Symmetric Polynomials*, Mediterranean Journal of Mathematics **20** 279 (2023), 1–20.

[94] G. M. Lăchescu, M. Malin, and I. Rovenţa, *Convex type inequalities with nonpositive weights*, (2024), submitted.

[95] G. M. Lăchescu, M. Mălin, and I. Rovenţa, *On the barycenter for discrete Steffensen Popoviciu measures on global NPC spaces*, submitted for publication.

[96] G. M. Lăchescu, and I. Rovenţa *The Hardy-Littlewood-Pólya inequality of majorization in the context of ω-$\mathbf{m}$-star-convex functions*, Aequ. Math. **97** (2023), 523–535.

[97] G. M. Lăchescu and V. Uţă, *On the properties of strongly h_d convex functions*, accepted for publication in Annals of the University of Craiova - Mathematics and Computer Science Series **51** (2024) (2), 1–10.

[98] Y. Lim, *Convex geometric means*, J. Math. Anal. Appl. **404** (2013), No. 1, 115–128.

[99] P. Lissy, and I. Rovenţa, *Optimal approximation of internal controls for a wave-type problem with fractional Laplacian using finite-difference method*, Math. Models Methods Appl. Sci. **30** (2020), 439–475.

[100] P. Lissy, and I. Rovenţa, *Optimal filtration for the approximation of boundary controls for the one-dimensional wave equation using finite-difference method*, Math. Comp. **88** (2019), 273–291.

[101] L. Maligranda, J. Pečarić and L.-E. Persson, *Weighted Favard's and Berwald's inequalities*, J. Math. Anal. Appl. **190** (1) (1995), 248–262.

[102] M. Marcus and L. Lopes. *Inequalities for symmetric functions and hermitian matrices.* Canadian Journal of Mathematics, 9:305–312, 1957.

[103] A.W. Marshall, I. Olkin, *Inequalities: theory of majorization and its applications*, Vol. 143. New York, Academic Press, 1979.

[104] A.W. Marshall, I. Olkin and B. Arnold, *Inequalities: Theory of majorization and its applications*, 2nd ed., Springer Series in Statistics, Springer, New York, 2011.

[105] J. McLeod. *On four inequalities in symmetric functions*, Proceedings of the Edinburgh Mathematical Society, **11** (4):211–219, 1959.

[106] P. Meyer-Nieberg, *Banach Lattices*, Springer-Verlag, Berlin, 1991.

[107] D. S. Mitrinović, J. E. Pećarić and A. M. Fink, *Inequality involving functions and their integrals and derivatives*, Kluwer Academic Publishers, Dordrecht, 1991.

[108] P.T. Mocanu, I. Serb, and G. Toader, *Real star-convex functions*, Stud. Univ. Babes-Bolyai, Math **43** (1997), 65–80.

[109] R. F. Muirhead, *Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters*, Proc, Edinburgh Math. Soc. **21** (1903), 144–157.

[110] P. Neff, D. Pauly, K. J. Witsch, *Poincaré meets Korn via Maxwell: Extending Korn's First Inequality to Incompatible Tensor Fields*, Journal of Differential Equations, Volume 258 (4), pp. 1267–1302.

[111] C. P. Niculescu, *The Hermite-Hadamard inequality for convex functions on a global NPC space*, J. Math. Anal. Appl. **356** (1) (2009), 295-301.

[112] C.P. Niculescu, *A new look at the Hardy-Littlewood-Polya inequality of majorization*, J. Math. Anal. Appl. **501** (2021) (2), article 125211.

[113] C. P. Niculescu, and O. Olteanu, *From the Hahn-Banach extension theorem to the isotonicity of convex functions and the majorization theory*, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM **114** (2020), 1–19.

[114] C.P. Niculescu, L.-E. Persson, *Convex Functions and Their Applications. A Contemporary Approach*, 2nd ed., CMS Books in Mathematics vol. 23, Springer-Verlag, New York, 2018.

[115] C. P. Niculescu, and L.-E. Persson, *Convex Functions and their Applications. A Contemporary Approach*, CMS Books in Mathematics vol. 23, Springer-Verlag, New York, 2006.

[116] C. P. Niculescu, and L.-E. Persson, *Convex Functions and Their Applications. A Contemporary Approach*, 3rd edition. CMS Books in Mathematics vol. 23, xvii + 493 pp., Springer International Publishing AG, 2024.

[117] C.P. Niculescu, F. Popovici, *The extension of majorization inequalities within the framework of relative convexity*, Journal of Inequalities in Pure and Applied Mathematics (JIPAM) **7** (2006) Article No. 27, 6 pp. (Electronic only)

[118] C.P. Niculescu, I. Rovenţa, *Relative convexity and its applications*, Aequationes Math. **89** (2015), 1389–1400.

[119] C. P. Niculescu, I. Rovenţa, *Schauder fixed point theorem in metric spaces with non-positive curvature*, Fixed Point Theory and Applications, volume 2009 , article ID 906727, 8 pages, ISSN 1687-1820, doi:10.1155/2009/906727.

[120] C. P. Niculescu, I. Rovenţa, *Weak majorization in Cat*[0] *spaces*, work in progress.

[121] C. P. Niculescu, I. Rovenţa, *The Fan's inequality in metric spaces with non-positive curvature*, Applied Mathematics Letters, **22** (2009), 1529-1533.

[122] C. P. Niculescu, I. Rovenţa, *Generalized convexity and the existence of finite time blow-up solutions for an evolutionary problem*, Nonlinear Analysis: Theory, Methods and Applications, **75**, 270-277, doi:10.1016/j.na.2011.08.031, (2012).

[123] C. P. Niculescu, I. Rovenţa, *Large Solutions for Semilinear Parabolic Equations Involving Some Special Classes of Nonlinearities*, Discrete Dynamics in Nature and Society, Volume 2010, Article ID 491023, 11 pages.

[124] C.P. Niculescu, I. Rovenţa, *Relative Schur convexity on global NPC spaces*, Math. Inequal. Appl. **18** (2015), 1111–1119.

[125] C.P. Niculescu, I. Rovenţa, *Hardy-Littlewood-Pólya theorem of majorization in the framework of generalized convexity*, Carpathian Journal of Mathematics **33** (2017), 87–95.

[126] C. P. Niculescu, and I. Rovenţa, *An approach of majorization in spaces with a curved geometry*, J. Math. Anal. Appl. **411** (2014), 119–129.

[127] C. P. Niculescu, C. Spiridon, *New Jensen-type inequalities*, J. Math. Anal. Appl. **401** (1) (2013), 343–348.

[128] F. Nielsen and R. Bhatia, eds., *Matrix information geometry*, Heidelberg: Springer, 2013.

[129] M. A. Nielsen, *An introduction to majorization and its applications to quantum mechanics Lecture Notes*, Department of Physics, University of Queensland, Australia, October 18, 2002.

[130] M. A. Nielsen and I. L. Chung, *Quanum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2010.

[131] M. Niezgoda, *Linear maps preserving group majorization*, Linear Algebra Appl. **330** (2001), 113–127.

[132] M. Niezgoda, *Majorization and refined Jensen Mercer type inequalities for self-adjoint operators*, Linear Algebra Appl. **467** (2015), 1–14.

[133] M. Niezgoda, *Nonlinear Sherman-type inequalities*, Adv. Nonlinear Anal. **9** (2020), 168–175.

[134] K. Nikodem, T. Rajba, and S. Wasowicz, *Functions genereting strongly Schur-convex sums*, Inequalities and Applications, International Series of Numerical Mathematics, Springer Basel, **161** (2012), 175–182.

[135] K. Nikodem, Zs. Páles, *Characterization of inner product spaces by strongly convex functions*, Banach J. Math. Anal. **5** (2011) (1), 83–87.

[136] O. A. Oleinik, A.S. Shamaev, G.A. Yosifan, *Mathematical Problems in Elasticity and Homogenisation*, Stud. Math. Appl. vol. 26, North-Holland, 1992.

[137] Z. Páles, *On approximately convex functions*, Proceedings of the American Mathematical Society **131** (1) (2003), 243–252.

[138] C. E. M. Pearce, J. Pecaric, *On convexity-like inequalities*, Revue Roumaine Math. Pure Appl. Math. **42** (1-2) (1997), 133-138.

[139] J.E. Pečarić, F. Proschan, Y.C. Tong, *Convex Functions*, Partial Orderings and Statistical Applications. Academic Press Inc, San Diego (1992).

[140] D. Petz, *Quantum information theory and quantum statistics*, Springer Science & Business Media, 2007.

[141] G. Pólya, *Remark on Weyl's note "Inequalities between the two kinds of eigenvalues of a linear transformation"*, Proc. Nat. Acad. Sci. U.S.A. **36** (1950), 49-51.

[142] B. T. Polyak, *Existence theorems and convergence of minimizing sequences in extremum problems with restrictions*, Doki. Akad. Nauk. SSSR **166** (1966), 287–290 [Russian].

[143] T. Popoviciu, *Sur quelques propriétés des fonctiones d'une ou de deux variables réelles, thèse*, Faculté des Sciences de Paris, 1933. See also Mathematica (Cluj) **VIII** (1934), 1–85.

[144] T. Popoviciu, *Sur certains inégalités qui caractérisent les fonctions convexes*, An. Şti. "Al. I. Cuza", Iasi, Sect. **I** a Mat. (NS) **11 B** (1965), 155-164.

[145] T. Popoviciu, *Sur les équations algébriques ayant toutes leurs racines réelles*, Mathematica (Cluj) **9** (1935), 129-145.

[146] R. Rado, *An inequality*, J. London Math. Soc., **27** (1952), 1-6.

[147] Y. G. Reshetnyak, *Inextensible mapping in a space of curvature no greater than K*, Sib. Math. Jour. **9**, 683-689, 1968.

[148] B. Reznick, *Uniform denominators in Hilbert's 17th problem*, Math Z. **220** (1995), no. 1, 75–97.

[149] W. W. Rogosinski, *Moments of non-negative mass*, Proc. Roy. Soc. London Ser. A, vol. 245, 1958, pp. 1–27.

[150] I. Rovenţa, *Hardy–Littlewood–Polya's Inequality and a New Concept of Weak Majorization*, Mediterr. J. Math. **13** (2016), 573–583.

[151] I. Rovenţa, *Schur-convexity of a class of symmetric functions*, Ann. Univ. Craiova Math. Comput. Sci. Ser. **37** (2010) (1), 12–18.

[152] I. Rovenţa, *A note on Schur-concave functions*, J. Inequal. Appl. **159** (2012) , 9 pp.

[153] I. Rovenţa, L. E. Temereancă, *A note on the positivity of the even degree complete homogeneous symmetric polynomials*, Mediterr. J. Math. **16** (2019), 1–16.

[154] I. Rovenţa, L. E. Temereanca, and A. M. Tudor, *A note on weighted Ingham's inequality for families of exponentials with no gap*, 24th ICSTCC (2020), 43–48.

[155] I. Rovenţa, L. E. Temereanca, and A. M. Tudor, *Weighted Ingham's type inequalities via the positivity of quadratic polynomials*, Aequat. Math. **98** (2024), 865–883.

[156] L. Rüschendorf, *Mathematical risk analysis*, Springer Ser. Oper. Res. Financ. Eng. Springer, Heidelberg 2013.

[157] I. Schur, *Über ein flächen kleinsten flächeninhalts betreffendes problem der variationsrechnung*, Acta Societatis Scientiarum Fennicae **XV** (1888), 318-362.

[158] R. Séroul, *Programming for mathematicians*, Universitext, Springer-Verlag, Berlin, 2000, Translated from the 1995 French original by Donal O'Shea.

[159] A. Shapiro, *On duality theory of conic linear problems*, Semi-Infinite Programming: Recent Advances (Miguel Á. Goberna and Marco A. López, eds.), Springer US, Boston, MA, 2001, pp. 135–165.

[160] S. Sherman, *On a theorem of Hardy, Littlewood, Pólya, and Blackwell*, Proc. Natl. Acad. Sci. USA **37** (1951) 826–831; Proc. Natl. Acad. Sci. USA 38 (1952) 382 (Erratum).

[161] B. Simon, *Convexity. An Analytic Viewpoint*, Cambridge University Press, 2011.

[162] J. F. Steffensen, *On certain inequalities and methods of appoximation*, J. Inst. Actuaries, **51** (1919), no. 3, 274-297.

[163] K. T. Sturm, *Probability measures on metric spaces of nonpositive curvature*, In vol.: *Heat kernels and analysis on manifolds, graphs, and metric spaces* (Pascal Auscher et al. editors). Lecture notes from a quarter program on heat kernels, random walks, and analysis on manifolds and graphs, April 16–July 13, 2002, Paris, France. Contemp. Math. **338** (2003), 357-390.

[164] S. Sra, *New concavity and convexity results for symmetric polynomials and their ratios*, Linear Multilinear Algebra **68** (2020), no. 5, 1031–1038.

[165] T. Tao, https://terrytao.wordpress.com/2017/08/06/schur-convexity-and-positive-definiteness-of-the-even-degree-complete-homogeneous-symmetric-polynomials/.

[166] G. Toader, *Some generalizations of the convexity*, Proceedings of the Colloquium on Approximation and Optimization, Univ. Cluj-Napoca, Cluj-Napoca, 1985, 329-338.

[167] M. Tomić, *Théoréme de Gauss reltif au centre de gracité et son application*, Bull. Soc. Math. Phys. Serbie **1** (1949), 31-40.

[168] H. Weyl, *Inequalities between two kinds of eigenvalues of a linear transformation*, Proc. Natl. Acad. Sci. U. S. A. **35** (1949), no. 7, 408-411.

[169] C. Zălinescu, *On uniformly convex functions*, J. Math. Anal. Appl. **95** (1983), 344–374.

[170] S. Zlobec, *Jensen's inequality for nonconvex functions*, Math. Commun. **9** (2004), 119–124.

[171] S. Zlobec, *Convexifiable functions in integral calculus*, Glas. Mat. **40** (60) (2005), 241–247.

[172] E. Zuazua, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Rev. **47** (2005), 197–243.