MAJORIZATION IN SPACES WITH A CURVED GEOMETRY

CONSTANTIN P. NICULESCU AND IONEL ROVENTA

Abstract. The Hardy-Littlewood-Pólya majorization theorem is extended to the framework of some spaces with a curved geometry (such as the global NPC spaces and the Wasserstein spaces). We also discuss the connection between our concept of majorization and the subject of Schur convexity.

In 1929, G. H. Hardy, J. E. Littlewood and G. Pólya [9], [10] have proved an important characterization of convex functions in terms of a partial ordering of vectors \(x = (x_1, ..., x_n)\) in \(\mathbb{R}^n\). In order to state it we need a preparation. We denote by \(x^\downarrow\) the vector with the same entries as \(x\) but rearranged in decreasing order,

\[ x_1^\downarrow \geq \cdots \geq x_n^\downarrow. \]

Then \(x\) is weakly majorized by \(y\) (abbreviated, \(x \prec^\ast y\)) if

\[ \sum_{i=1}^{k} x_i^\downarrow \leq \sum_{i=1}^{k} y_i^\downarrow \quad \text{for } k = 1, ..., n \]

and \(x\) is majorized by \(y\) (abbreviated, \(x \prec y\)) if in addition

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

Intuitively, \(x \prec y\) means that the components in \(x\) are less spread out than the components in \(y\). As is shown in Theorem 1 below, the concept of majorization admits an order-free characterization based on the notion of doubly stochastic matrix. Recall that a matrix \(A \in \mathcal{M}_n(\mathbb{R})\) is doubly stochastic if it has nonnegative entries and each row and each column sums to unity.

**Theorem 1.** (Hardy, Littlewood and Pólya [9], Theorem 8). Let \(x\) and \(y\) be two vectors in \(\mathbb{R}^n\), whose entries belong to an interval \(I\). Then the following statements are equivalent:

i) \(x \prec y\);

ii) There is a doubly stochastic matrix \(A = (a_{ij})_{1 \leq i,j \leq n}\) such that \(x = Ay\);

iii) The inequality \(\sum_{i=1}^{n} f(x_i) \leq \sum_{i=1}^{n} f(y_i)\), holds for every continuous convex function \(f : I \to \mathbb{R}\).

The proof of this result is also available in the recent monographs [15] and [18].

Date: April 3, 2012.

2000 Mathematics Subject Classification. Primary 52A41; Secondary 46A55, 52A05.

Key words and phrases. generalized convex function, global NPC space, majorization of measures, Wasserstein distance.

Corresponding author: Constantin P. Niculescu.
Remark 1. M. Tomić [25] and H. Weyl [26] have noticed the following characterization of weak majorization: \( x \prec^* y \) if and only if \( \sum_{i=1}^{n} f(x_i) \leq \sum_{i=1}^{n} f(y_i) \) for every continuous nondecreasing convex function \( f \) defined on an interval containing the components of \( x \) and \( y \). The reader will find the details in [15], Proposition B2, p. 157.

Nowadays there are known many important applications of majorization to matrix theory, numerical analysis, probability, combinatorics, quantum mechanics etc. See [3], [15], [18], [21], and [22]. They were made possible by the constant growth of the theory, able to uncover the most diverse situations.

In what follows we will be interested in a simple but basic extension of the concept of majorization as was mentioned above: the weighted majorization. Indeed, the entire subject of majorization can be switched from vectors to Borel probability measures by identifying a vector \( x = (x_1, ..., x_n) \) in \( \mathbb{R}^n \) with the discrete measure \( \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \) acting on \( \mathbb{R} \). By definition,

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

if the conditions (1) and (2) above are fulfilled, and Theorem 1 can be equally seen as a characterization of this instance of majorization.

Choquet’s theory made available a very general framework of majorization by allowing the comparison of Borel probability measures whose supports are contained in a compact convex subset of a locally convex separated space. The highlights of this theory are presented in [22] and refer to a concept of majorization based on condition \( iii \) in Theorem 1 above. Of interest to us is the particular case of discrete probability measures on the Euclidean space \( \mathbb{R}^N \), that admits an alternative approach via condition \( ii \) in the same Theorem 1. Indeed, in this case one can introduce a relation of the form

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

by asking the existence of a \( m \times n \)-dimensional matrix \( A = (a_{ij})_{i,j} \) such that

\[
a_{ij} \geq 0, \text{ for all } i, j
\]

(4)

\[
\sum_{j=1}^{n} a_{ij} = 1, \quad i = 1, ..., m
\]

(5)

\[
\mu_j = \sum_{i=1}^{m} a_{ij} \lambda_i, \quad j = 1, ..., n
\]

(6)

and

\[
x_i = \sum_{j=1}^{n} a_{ij} y_j, \quad i = 1, ..., n
\]

(7)

The matrices verifying the conditions (4)\&(5) are called **stochastic on rows**. When \( m = n \) and all weights \( \lambda_i \) and \( \mu_j \) are equal, the condition (6) assures the **stochasticity on columns**, so in that case we deal with doubly stochastic matrices.
The fact that (3) implies
\[ \sum_{i=1}^{m} \lambda_i f(x_i) \prec \sum_{j=1}^{n} \mu_j f(y_j), \]
for every continuous convex function \( f \) defined on a convex set containing all points \( x_i \) and \( y_j \), is covered by a general result due to S. Sherman [23]. See also the paper of J. Borcea [5] for a nice proof and important applications.

It is worth noticing that the extended definition of majorization given by (3) is related, via equality (7), to an optimization problem as follows:
\[ x_i = \arg \min_{z \in \mathbb{R}^N} \frac{1}{2} \sum_{j=1}^{n} a_{ij} \| z - y_j \|^2, \quad \text{for } i = 1, \ldots, m. \]

The aim of the present paper is to discuss the analogue of the relation of majorization (3) within certain classes of spaces with curved geometry. We will start with the spaces with global nonpositive curvature (abbreviated, global NPC spaces). The subject of majorization in these spaces was touched in [17] via a different concept of majorization. Central to us here is the generalization of Theorem 1.

**Definition 1.** A global NPC space is a complete metric space \( M = (M, d) \) for which the following inequality holds true: for each pair of points \( x_0, x_1 \in M \) there exists a point \( y \in M \) such that for all points \( z \in M \),
\[ d^2(z, y) \leq \frac{1}{2} d^2(z, x_0) + \frac{1}{2} d^2(z, x_1) - \frac{1}{4} d^2(x_0, x_1). \]

These spaces are also known as the Cat 0 spaces. See [3]. 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.

In a global NPC space, the geodesics play the role of segments. The point \( y \) that appears in Definition 1 is the midpoint of \( x_0 \) and \( x_1 \) and has the property
\[ d(x_0, y) = d(y, x_1) = \frac{1}{2} d(x_0, x_1). \]

Every Hilbert space is a global NPC space. Its geodesics are the line segments. The upper half-plane \( \mathbb{H} = \{ z \in \mathbb{C} : \text{Im } z > 0 \} \), endowed with the Poincaré metric,
\[ ds^2 = \frac{dx^2 + dy^2}{y^2}, \]
constitutes another example of a global NPC space. In this case the geodesics are the semicircles in \( \mathbb{H} \) perpendicular to the real axis and the straight vertical lines ending on the real axis.

A Riemannian manifold \( (M, g) \) 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 [6]. More information on global NPC spaces is available in [2], [12], and [24]. See also our papers [17] and [20].

**Definition 2.** A set \( C \subset M \) is called convex if \( \gamma([0, 1]) \subset C \) for each geodesic \( \gamma : [0, 1] \to M \) joining \( \gamma(0), \gamma(1) \in C \).
A function $\varphi : C \to \mathbb{R}$ 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,

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

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

The function $\varphi$ is called concave if $-\varphi$ is convex.

The distance function on a global NPC space $M = (M, d)$ verifies not only the inequality (8), but also the following stronger version of it,

$$d^2(z, x_t) \leq (1 - t)d^2(z, x_0) + td^2(z, x_1) - t(1 - t)d^2(x_0, x_1);$$

here $z$ is any point in $C$ and $x_t$ is any point on the geodesic $\gamma$ joining $x_0, x_1 \in C$. In terms of Definition 2, this shows that all the functions $d^2(\cdot, z)$ are uniformly convex. In particular, they are convex and the balls are convex sets.

In a global NPC space $M = (M, d)$ the distance function $d$ is convex on $M \times M$ and also convex are the functions $d(\cdot, z)$. See [24], Corollary 2.5, for details.

Recall that the direct product of metric spaces $M_i = (M_i, d_i)$ ($i = 1, \ldots, n$) is the metric space $M = (M, d_M)$ defined by $M = \prod_{i=1}^n M_i$ and

$$d_M(x, y) = \left(\sum_{i=1}^n d_i(x_i, y_i)^2\right)^{1/2}.$$ 

It is a global NPC space if all factors are global NPC spaces.

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

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

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

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

and

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

The existence and uniqueness of the problems of optimization (11) is assured by the fact that the objective functions are uniformly convex and positive. See [12], Section 3.1, or [24], Proposition 1.7, p. 3.

Notice that the above definition agrees with the usual one in the Euclidean case. It is also related to the definition of the barycenter of a Borel probability measure $\mu$ defined on a global NPC space $M$. Precisely, if $\mu \in \mathcal{P}_2(M)$ (the set of those probability measures under which all functions $d^2(\cdot, z)$ are integrable), then its barycenter is defined by the formula

$$\text{bar}(\mu) = \arg \min_{z \in M} \frac{1}{2} \int_M d^2(z, x) d\mu(x).$$
This definition, due to E. Cartan [7], was inspired by Gauss’ Least Squares Method. A larger approach of the notion of barycenter is offered by the recent paper of Sturm [24].

The particular case of discrete probability measures $\lambda = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ is of special interest because the barycenter of $\lambda$ can be seen as a good analogue for the convex combination (or weighted mean) $\lambda_1 x_1 + \cdots + \lambda_n x_n$. Indeed,

$$\text{bar}(\lambda) = \arg\min_{z \in M} \frac{1}{2} \sum_{i=1}^{n} \lambda_i d^2(z, x_i),$$

and the way $\text{bar}(\lambda)$ provides a mean with nice features was recently clarified by Lawson and Lim [14]. As an immediate consequence one obtains the relation $\delta_{\text{bar}(\lambda)} \prec \lambda$.

A word of caution when denoting $\text{bar}(\lambda)$ as $\lambda_1 x_1 + \cdots + \lambda_n x_n$. Probably a notation like $\lambda_1 x_1 \oplus \cdots \oplus \lambda_n x_n$ suits better because $\text{bar}(\lambda)$ can be far from the usual the arithmetic mean. Consider for example the case where $M$ is the space $\text{Sym}^{++}(n, \mathbb{R})$ (of all positively definite matrices with real coefficients), endowed with the trace metric,

$$d_{\text{trace}}(A, B) = \left( \sum_{k=1}^{n} \log^2 \lambda_k \right)^{1/2},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $AB^{-1}$. In this case

$$\frac{1}{2} A \oplus \frac{1}{2} B = A^{1/2} (A^{-1/2} BA^{-1/2})^{1/2} A^{1/2},$$

that is, it coincides with the geometric mean of $A$ and $B$. See [4], Section 6.3, or [13], for details.

Since the convex combinations within a global NPC space lack in general the property of associativity,

$$\sum_{i=1}^{n+1} \lambda_i x_i = (1 - \lambda_{n+1}) \left( \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_{n+1}} x_i \right) + \lambda_{n+1} x_{n+1},$$

the proof of Jensen’s inequality is not trivial even in the discrete case. This explains why this inequality was first stated in this context only in 2001 by J. Jost [11]. We recall it here in the formulation of Eells and Fuglede [8], Proposition 12.3, p. 242:

**Theorem 2.** (Jensen’s Inequality). For any lower semicontinuous convex function $f : M \to \mathbb{R}$ and any Borel probability measure $\mu \in \mathcal{P}_2(M)$ we have the inequality

$$f(\text{bar}(\mu)) \leq \int_M f(x) d\mu(x),$$

provided the right hand side is well-defined.

The proof of Eells and Fuglede is based on the following remark concerning barycenters: If a probability measure $\mu$ is supported by a convex closed set $K$, then its barycenter $\text{bar}(q)$ lies in $K$. A probabilistic approach of Theorem 2 is due to Sturm [24].
An immediate consequence of Theorem 2 is the following couple of inequalities that work for any points \( z, x_1, \ldots, x_n, y_1, \ldots, y_n \) in a global NPC space:
\[ d^2 \left( \frac{1}{n} x_1 \oplus \cdots \oplus \frac{1}{n} x_n, z \right) \leq \frac{d^2(x_1, z) + \cdots + d^2(x_n, z)}{n} \]
and
\[ d \left( \frac{1}{n} x_1 \oplus \cdots \oplus \frac{1}{n} x_n, \frac{1}{n} y_1 \oplus \cdots \oplus \frac{1}{n} y_n \right) \leq \frac{d(x_1, y_1) + \cdots + d(x_n, y_n)}{n}. \]

The next theorem offers a partial extension of Hardy-Littlewood-Pólya Theorem to the context of global NPC spaces.

**Theorem 3.** If
\[ \sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{n} \mu_j \delta_{y_j}, \]
in the global NPC space \( M \), then, for every continuous convex function \( f \) defined on a convex subset \( U \subseteq M \) containing all points \( x_i \) and \( y_j \) we have
\[ \sum_{i=1}^{m} \lambda_i f(x_i) \leq \sum_{j=1}^{n} \mu_j f(y_j). \]

**Proof.** By our hypothesis, there is an \( m \times n \)-dimensional matrix \( A = (a_{ij})_{i,j} \) that is stochastic on rows and verifies the conditions (10) and (11). The last condition, shows that each point \( x_i \) is the barycenter of the probability measure \( \sum_{j=1}^{n} a_{ij} \delta_{y_j} \), so by Jensen’s inequality we infer that
\[ f(x_i) \leq \sum_{j=1}^{n} a_{ij} f(y_j). \]

Multiplying each side by \( \lambda_i \) and then summing up over \( i \) from 1 to \( m \), we conclude that
\[ \sum_{i=1}^{m} \lambda_i f(x_i) \leq \sum_{i=1}^{m} \left( \lambda_i \sum_{j=1}^{n} a_{ij} f(y_j) \right) \]
\[ = \sum_{j=1}^{n} \left( \sum_{i=1}^{m} a_{ij} \lambda_i \right) f(y_j) \]
\[ = \sum_{j=1}^{n} \mu_j f(y_j). \]
\[ \square \]

In a global NPC space the distance function from a convex set is a convex function. See [24], Corollary 2.5. Combining this fact with Theorem 3 we infer the following result.

**Corollary 1.** If
\[ \sum_{i=1}^{m} \lambda_i \delta_{x_i} \prec \sum_{j=1}^{n} \mu_j \delta_{y_j}, \]
and all coefficients \( \lambda_i \) are positive, then \( \{x_1, \ldots, x_m\} \) is contained in the convex hull of \( \{y_1, \ldots, y_n\} \).
In particular, the points \( x_i \) spread out less than the points \( y_j \).

Another application of Theorem 3 yields a new set of inequalities verified by the functions \( d(\cdot, z) \) in a global NPC space \( M \). These functions are convex and the same is true for the functions \( f(d(\cdot, z)) \) whenever \( f \) is a continuous nondecreasing convex function defined on \( \mathbb{R}_+ \). According to Theorem 3 if \( \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} < \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i} \), then \( \sum_{i=1}^{n} f(\delta_{x_i}) \leq \sum_{i=1}^{n} f(\delta_{y_i}) \). Taking into account Remark 1 we arrive at the following result:

**Corollary 2.** If \( \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} < \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i} \), in \( M = (M, d) \), then for all \( z \in M \),

\[
(d(x_1, z), ..., d(x_n, z)) \prec_a (d(y_1, z), ..., d(y_n, z))
\]

According to a result due to Ando (see [15], Theorem B.3a, p. 158), the converse of Corollary 2 works when \( M = \mathbb{R} \).

The entropy function, \( H(t) = -t \log t \), is concave and decreasing for \( t \in [1/e, \infty) \), so by Corollary 2 we infer that

\[
\prod_{i=1}^{n} d(x_i, z)^{d(x_i, z)} \geq \prod_{i=1}^{n} d(y_i, z)^{d(y_i, z)},
\]

whenever \( \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} < \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i} \) and all the points \( x_i \) and \( y_i \) are at a distance \( \geq 1/e \) from \( z \).

Many other inequalities involving distances in a global NPC space can be derived from Corollary 2 and the following result due to Fan and Mirsky: if \( x, y \in \mathbb{R}_+^n \), then \( x \prec_a y \) if and only if

\[
\Phi(x) \leq \Phi(y)
\]

for all functions \( \Phi : \mathbb{R}^n \to \mathbb{R} \) such that:

1. \( \Phi(x) > 0 \) when \( x \neq 0 \);
2. \( \Phi(\alpha x) = |\alpha| \Phi(x) \) for all real \( \alpha \);
3. \( \Phi(x + y) \leq \Phi(x) + \Phi(y) \);
4. \( \Phi(x_1, ..., x_n) = \Phi(\varepsilon_1 x_{\pi(1)}, ..., \varepsilon_n x_{\pi(n)}) \) whenever each \( \varepsilon_i \) belongs to \( \{ -1, 1 \} \) and \( \pi \) is any permutation of \( \{ 1, ..., n \} \).

For details, see [15], Proposition B6, p. 160.

It is worth noticing the connection between our definition of majorization and the subject of Schur convexity (as presented in [15]):

**Theorem 4.** Suppose that \( \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} < \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i} \) in the global NPC space \( M = (M, d) \), and \( f : M^n \to \mathbb{R} \) is a continuous convex function invariant under the permutation of coordinates. Then

\[
f(x_1, ..., x_n) \leq f(y_1, ..., y_n).
\]

**Proof.** For the sake of simplicity we will restrict here to the case where \( n = 3 \).

According to the definition of majorization, if \( \frac{1}{3} \sum_{i=1}^{3} \delta_{x_i} < \frac{1}{3} \sum_{i=1}^{3} \delta_{y_i} \), then there exists a doubly stochastic matrix \( A = (a_{ij})_{3 \times 3} \) such that

\[
x_i = \text{bar}(\sum_{j=1}^{n} a_{ij} \delta_{y_j}) \quad \text{for } i = 1, ..., n.
\]
As $A$ can be uniquely represented under the form

$$A = \begin{pmatrix}
\lambda_1 + \lambda_2 & \lambda_3 + \lambda_5 & \lambda_4 + \lambda_6 \\
\lambda_3 + \lambda_4 & \lambda_1 + \lambda_6 & \lambda_2 + \lambda_5 \\
\lambda_5 + \lambda_6 & \lambda_2 + \lambda_4 & \lambda_1 + \lambda_3
\end{pmatrix},$$

where all $\lambda_k$ are nonnegative and $\sum_{k=1}^{6} \lambda_k = 1$ (a simple matter of linear algebra) we can represent the elements $x_j$ as

$$x_1 = \text{bar}((\lambda_1 + \lambda_2)\delta_{y_1} + (\lambda_3 + \lambda_4)\delta_{y_2} + (\lambda_5 + \lambda_6)\delta_{y_3}),$$
$$x_2 = \text{bar}((\lambda_3 + \lambda_5)\delta_{y_1} + (\lambda_1 + \lambda_6)\delta_{y_2} + (\lambda_2 + \lambda_4)\delta_{y_3}),$$
$$x_3 = \text{bar}((\lambda_4 + \lambda_6)\delta_{y_1} + (\lambda_2 + \lambda_5)\delta_{y_2} + (\lambda_1 + \lambda_3)\delta_{y_3}).$$

It is easy to see that $(x_1, x_2, x_3)$ is the barycenter bar of

$$\mu = \lambda_1 \delta_{(y_1, y_2, y_3)} + \lambda_2 \delta_{(y_1, y_3, y_2)} + \lambda_3 \delta_{(y_2, y_1, y_3)} + \lambda_4 \delta_{(y_2, y_3, y_1)} + \lambda_5 \delta_{(y_3, y_1, y_2)} + \lambda_6 \delta_{(y_3, y_2, y_1)},$$

so by Jensen’s inequality and the symmetry of $f$ we get

$$f(x_1, \ldots, x_n) \leq \lambda_1 f(y_1, y_2, y_3) + \lambda_2 f(y_1, y_3, y_2) + \lambda_3 f(y_2, y_1, y_3) + \lambda_4 f(y_2, y_3, y_1) + \lambda_5 f(y_3, y_1, y_2) + \lambda_6 f(y_3, y_2, y_1) = (\lambda_1 + \cdots + \lambda_6) f(y_1, y_2, y_3) = f(y_1, y_2, y_3).$$

□

The following consequence of Theorem 4 relates the majorization of measures to the dispersion of their supports.

**Corollary 3.** If $\frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \prec \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i}$ in the global NPC space $M = (M, d)$, then

$$\sum_{1 \leq i < j \leq n} d^\alpha(x_i, x_j) \leq \sum_{1 \leq i < j \leq n} d^\alpha(y_i, y_j)$$

for every $\alpha \geq 1$.

Alert readers have probably already noticed that essential for the theory of majorization presented above is the occurrence of the following two facts:

1. the existence of a unique minimizer for the functionals of the form

$$J(x) = \frac{1}{2} \sum_{i=1}^{m} \lambda_i d^2(x, x_i)$$

(thought of as the barycenter bar($\lambda$) of the discrete probability measure $\lambda = \sum_{i=1}^{m} \lambda_i \delta_{x_i}$);

2. the Jensen type inequality,

$$f(\text{bar($\lambda$)}) \leq \int f d\lambda = \sum_{i=1}^{m} \lambda_i f(x_i),$$

for $f$ in our class of generalized convex functions.

The recent paper of Agueh and Carlier \[1\] shows that such a framework is available also in the case of certain Borel probability measures, equipped with the Wasserstein metric. More precisely they consider the space $\mathcal{P}_2(\mathbb{R}^N)$ (of all
Borel probability measures on $\mathbb{R}^N$ having finite second moments) endowed with the Wasserstein metric,

$$W_2(\mu, \nu) = \inf \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \|x - y\|^2 \, d\gamma(x, y) \right)^{1/2},$$

where the infimum is taken over all Borel probability measures $\gamma$ on $\mathbb{R}^N \times \mathbb{R}^N$ with marginals $\mu$ and $\nu$.

The barycenter $\text{bar} \left( \sum_{i=1}^m \lambda_i \delta_{\nu_i} \right)$, of a discrete probability measure $\sum_{i=1}^m \lambda_i \delta_{\nu_i}$, is defined as the minimizer of the functional

$$J(\nu) = \frac{1}{2} \sum_{i=1}^m \lambda_i W_2^2(\nu_i, \nu).$$

This minimizer is unique when at least one of the measures $\nu_i$ vanishes on every Borel set of Hausdorff dimension $N - 1$. See [1], Proposition 2.2 and Proposition 3.5.

The natural class of convex function on the Wasserstein space is that of functions convex along barycenters. According to [1], Definition 7.1, a function $F : \mathcal{P}_2(\mathbb{R}^N) \to \mathbb{R}$ is said to be convex along barycenters if for any discrete probability measure $\sum_{i=1}^m \lambda_i \delta_{\nu_i}$ on $\mathcal{P}_2(\mathbb{R}^N)$ we have

$$F(\text{bar} \left( \sum_{i=1}^m \lambda_i \delta_{\nu_i} \right)) \leq \sum_{i=1}^m \lambda_i F(\nu_i).$$

This notion of convexity coincides with the notion of displacement convexity introduced by McCann [16] if $N = 1$, and is stronger than this in the general case. However, the main examples of displacement convex functions (such as the internal energy, the potential energy and the interaction energy) are also examples of functions convex along barycenters. See [1], Proposition 7.7.

**Theorem 5.** The concept of majorization and all results noticed in the case of global NPC spaces (in particular, Theorem 3 and Theorem 4) remain valid in the context discrete probability measures on $\mathcal{P}_2(\mathbb{R}^N)$ having unique barycenters and the functions $F : \mathcal{P}_2(\mathbb{R}^N) \to \mathbb{R}$ convex along barycenters.

We end our paper with an open problem that arises in connection to Rado’s geometric characterization of majorization in $\mathbb{R}^n : (x_1, \ldots, x_n) \prec (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ if and only if $(x_1, \ldots, x_n)$ lies in the convex hull of the $n!$ permutations of $(y_1, \ldots, y_n)$. See [15], Corollary B.3, p. 34. A relation of majorization of this kind can be introduced in the power space $M^n$ (of any global NPC space $M = (M, d)$ as well as of $\mathcal{P}_2(\mathbb{R}^N)$) by putting

$$(x_1, \ldots, x_n) \prec (y_1, \ldots, y_n)$$

if $\frac{1}{n} \sum_{i=1}^n \delta_{x_i} \prec \frac{1}{n} \sum_{i=1}^n \delta_{y_i}$. The proof of Theorem 5 yields immediately the necessity part of Rado’s characterization: if $(x_1, \ldots, x_n) \prec (y_1, \ldots, y_n)$ in $M^n$, then $(x_1, \ldots, x_n)$ lies in the convex hull of the $n!$ permutations of $(y_1, \ldots, y_n)$. Do the converse work? We know that the answer is positive if $M$ is a Hilbert space but the general case remains open.

**Acknowledgement.** This paper is supported by a grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-II-ID-PCE-2011-3-0257.
References

[1] M. Agueh and G. Carlier, Barycenters in the Wasserstein space, *SIAM J. Math. Anal.* 43 (2011), No. 2, 904-924.
[2] W. Ballmann, *Lectures on spaces with nonpositive curvature*, DMV Seminar Band 25, Birkhäuser, Basel, 2005.
[3] R. Bhatia, *Matrix analysis*, Springer-Verlag, 1997.
[4] R. Bhatia, *Positive definite matrices*, Princeton University Press, 2007.
[5] 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.
[6] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften vol. 319, Springer-Verlag, 1999.
[7] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Gauthiers-Villars, Paris, 1928; 2nd ed., 1946.
[8] J. Eells and B. Fuglede, *Harmonic maps between Riemannian polyhedra*, Cambridge University Press, 2001.
[9] G. H. Hardy, J. E. Littlewood and G. Pólya, Some simple inequalities satisfied by convex functions, *Messenger Math.* 58 (1929), 145-152.
[10] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge Mathematical Library, 2nd Edition, 1952, Reprinted 1988.
[11] J. Jost, Equilibrium maps between metric spaces, *Calc. Var.* 2 (1994), 173-204.
[12] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997.
[13] J. D. Lawson and Y. Lim, *The Geometric Mean, Matrices, Metrics, and More*, The American Mathematical Monthly 108 (2001), No. 9, 797-812.
[14] J. D. Lawson and Y. Lim, Monotonic properties of the least squares mean, *Math. Ann.* 351 (2011), 267-279.
[15] A. W. Marshall, I. Olkin and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, 2nd Edition, Springer-Verlag, 2011.
[16] R. J. McCann, A convexity principle for interacting gases, *Advances in Mathematics* 128 (1997), 153-179.
[17] C. P. Niculescu, The Hermite-Hadamard inequality for convex functions on a global NPC space, *J. Math. Anal. Appl.* 356 (2009), no. 1, 295–301.
[18] 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.
[19] C. P. Niculescu and I. Roventa, Fan’s inequality in geodesic spaces, *Appl. Math. Letters* 22 (2009), 1529-1533.
[20] C. P. Niculescu and I. Roventa, Schauder fixed point theorem in metric spaces with non-positive curvature, *Fixed Point Theory and Applications*, volume 2009 , article ID 906727, 8 pages, doi:10.1155/2009/906727.
[21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
[22] R. R. Phelps, *Lectures on Choquet’s Theorem*, 2nd Edition, Lecture Notes in Math. No. 1757, Springer-Verlag, Berlin, 2001.
[23] S. Sherman, On a theorem of Hardy, Littlewood, Pólya, and Blackwell, *Proc. Nat. Acad. Sci. USA* 37 (1951), 826–831; Errata, ibid. 38 (1952), 382.
[24] 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.
[25] M. Tomić, Théorème de Gauss relatif au centre de gravité et son application, *Bull. Soc. Math. Phys. Serbia* 1 (1949), 31-40.
[26] H. Weyl, Inequalities between two kinds of eigenvalues of a linear transformation, *Proc. Nat. Acad. Sci. U. S. A.* 35 (1949), 408-411.
University of Craiova, Department of Mathematics, Street A. I. Cuza 13, Craiova
200585, Romania
E-mail address: cniculescu47@yahoo.com

University of Craiova, Department of Mathematics, Street A. I. Cuza 13, Craiova
200585, Romania
E-mail address: roventaionel@yahoo.com