2-15-2018

# Barycentric maps for compactly supported measures

Sejong Kim
*Chungbuk National University*

Jimmie Lawson
*Louisiana State University*

Yongdo Lim
*Sungkyunkwan University*



## Recommended Citation

Kim, S., Lawson, J., & Lim, Y. (2018). Barycentric maps for compactly supported measures. *Journal of Mathematical Analysis and Applications, 458* (2), 1009-1026. https://doi.org/10.1016/j.jmaa.2017.09.037

J. Math. Anal. Appl. 458 (2018) 1009–1026

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# Barycentric maps for compactly supported measures

Sejong Kim [a], Jimmie Lawson [b,*], Yongdo Lim [c]

[a] Department of Mathematics, Chungbuk National University, Cheongju 362-763, Republic of Korea
[b] Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA
[c] Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea

ARTICLE INFO

*Article history:*
Received 4 April 2017
Available online 21 September 2017
Submitted by B. Bongiorno

*Keywords:*
Barycentric map
Borel probability measure
Power mean
Karcher geometric mean
Wasserstein metric

ABSTRACT

After outlining the way in which an intrinsic mean $G = \{G_n\}$ on a complete metric space gives rise to a contractive barycentric map on some class of Borel probability measures and some basic examples of this process, we show how the resulting barycentric map gives rise to a general theory of integration of measurable functions into the space. We apply this machinery to the cone of positive invertible elements of a $C^*$-algebra equipped with the Thompson metric to derive barycentric maps and their basic properties arising from the power means. Finally we derive basic results for the Karcher barycenter including its approximation by the barycentric maps for power means and its satisfaction of the Karcher equation.



## 1. Introduction

A barycentric map assigns to each measure in some designated set of Borel probability measures on a topological space $X$ a member of $X$. The map yields in an abstract fashion a method of assigning to a measure a barycenter or "center of mass." In this paper we restrict to the case that $X$ is a metric space and consider barycentric maps on two sets of probability Borel measures: (i) the set of probability measures of finite first moment, i.e., those measures $\mu$ such that $\int_X d(x, y)\, d\mu(x) < \infty$ for some (and hence all) $y \in X$, and (ii) the set of probability measures with compact support.

In [7] the last two authors have developed the theory of power means and Karcher means on the open cone of positive operators on a Hilbert space, a theory that directly extends to the cone $\mathbb{P}$ of positive elements on a monotone complete $C^*$-algebra with identity. Our main goal in this paper is to extend the power means and Karcher mean to contractive barycentric maps on the set of Borel probability measures of compact support. In Sections 2 and 3 we develop in the setting of metric spaces a general theory of contractive barycentric maps on spaces of Borel probability measures equipped with Wasserstein metrics, with particular attention to those with compact support equipped with the $d_\infty$-Wasserstein metric. We

* Corresponding author.
*E-mail addresses:* skim@chungbuk.ac.kr (S. Kim), lawson@math.lsu.edu (J. Lawson), ylim@skku.edu (Y. Lim).

https://doi.org/10.1016/j.jmaa.2017.09.037
0022-247X/

also review from [8] the important technique of obtaining barycentric maps from intrinsic means. Section 4 introduces some basic examples of barycentric maps by this method.

Metric spaces equipped with barycentric maps support a related theory of integration over probability measures. In Section 5 we introduce and develop relevant portions of this theory. We use this theory in Section 6 to introduce the barycentric maps for power means and develop their basic theory. In Section 7 we turn to the barycentric map arising from the Karcher geometric mean. We show that the corresponding geometric barycenter satisfies an appropriate Karcher equation and is given as the limit as $t \to 0$ of the power means.

We remark that finding Karcher and related barycenters of probability measures living on the cone of positive matrices or operators has been considered in other works; see e.g. [3] and [14]. What is different in this context is considering contractive barycentric maps with respect to Wasserstein metrics, particularly from the viewpoint of their arising naturally from means.

## 2. Preliminaries

For a metric space $X$, let $\mathcal{B}(X)$ be the algebra of Borel sets, the smallest $\sigma$-algebra containing the open sets. A *Borel measure* $\mu$ is a countably additive (positive) measure defined on $\mathcal{B}(X)$. The support of $\mu$ consists of all points $x$ for which $\mu(U) > 0$ for each open set $U$ containing $x$. The support of $\mu$ is always a closed set. The *finitely supported probability measures* are those of the form $\sum_{i=1}^n r_i \delta_{x_i}$, where for each $i$, $r_i \geq 0$, $\sum_{i=1}^n r_i = 1$, and $\delta_{x_i}$ is the point measure of mass 1 at the point $x_i$.

We recall the *Prohorov metric* $\pi(\mu, \nu)$ defined for two Borel probability measures $\mu, \nu$ on $X$ as the infimum of all $\varepsilon > 0$ such that for all closed sets $A$,

$$\mu(A) \leq \nu(A^\varepsilon) + \varepsilon, \quad \nu(A) \leq \mu(A^\varepsilon) + \varepsilon,$$

where $A^\varepsilon = \{x \in X : d(x, y) < \varepsilon \text{ for some } y \in A\}$. The following result appears in [4].

**Proposition 2.1.** *A Borel probability measure $\mu$ on a metric space $(X, d)$ has separable support. Furthermore, the following are equivalent.*

(1) *There exists a sequence $\{\mu_n\}$ of finitely supported measures (with rational coefficients) that converges to $\mu$ with respect to the Prohorov metric.*
(2) *The support of $\mu$ has measure 1, i.e., $\mu$ is support-concentrated.*

Let $\mathcal{P}(X)$ be the set of all support-concentrated Borel probability measures on $(X, \mathcal{B}(X))$ and $\mathcal{P}_0(X)$ the set of all $\mu \in \mathcal{P}(X)$ of the form $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$ for some $n \in \mathbb{N}$. Members of $\mathcal{P}_0(X)$ are also referred to as uniform probability measures with finite support. For $p \in [1, \infty)$ let $\mathcal{P}^p(X) \subseteq \mathcal{P}(X)$ be the set of probability measures with *finite p-moment*: for some (and hence all) $y \in X$,

$$\int_X d^p(x, y) d\mu(x) < \infty.$$

For $p = \infty$, $\mathcal{P}^\infty(X)$ denotes the set of probability measures with bounded support (with respect to the metric $d$) and $\mathcal{P}_{cp}(X) \subseteq \mathcal{P}^\infty(X)$ denotes those with compact support. The compactly supported measures will be our focus in what follows.

Let $(X, \mathcal{M})$ be a *measure space*, a set $X$ equipped with a $\sigma$-algebra $\mathcal{M}$, and $(Y, d)$ a metric space. A function $f : X \to Y$ is *measurable* if $f^{-1}(A) \in \mathcal{M}$ whenever $A \in \mathcal{B}(Y)$. For $f$ to be measurable, it suffices that $f^{-1}(U) \in \mathcal{M}$ for each open subset $U$ of $Y$. Hence continuous functions are measurable in the case $X$

is a metrizable space and $\mathcal{M} = \mathcal{B}(X)$, the Borel algebra. A measurable map $f : X \to Y$ between metric spaces induces a *push-forward* map $f_* : \mathcal{P}(X) \to \mathcal{P}(Y)$ defined by $f_*(\mu)(B) = \mu(f^{-1}(B))$ for $\mu \in \mathcal{P}(X)$ and $B \in \mathcal{B}(Y)$. Note for $f$ continuous that $\mathrm{supp}(f_*(\mu)) = f(\mathrm{supp}(\mu))^-$, the closure of the image of the support of $\mu$.

For $X$ a metric space, we say that $\omega \in \mathcal{P}(X \times X)$ is a *coupling* for $\mu, \nu \in \mathcal{P}(X)$ and that $\mu, \nu$ are *marginals* for $\omega$ if for all $B \in \mathcal{B}(X)$

$$\omega(B \times X) = \mu(B) \qquad \text{and} \qquad \omega(X \times B) = \nu(B).$$

Equivalently $\mu$ and $\nu$ are the push-forwards of $\omega$ under the projection maps $\pi_1$ and $\pi_2$ resp. We note that one such coupling is the product measure $\mu \times \nu$, and that for any coupling $\omega$ it must be the case that $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$. We denote the set of all couplings for $\mu, \nu \in \mathcal{P}(X)$ by $\Pi(\mu, \nu)$.

For $1 \leq p < \infty$, the $p$-Wasserstein distance $d_p^W$ (alternatively Kantorovich–Rubinstein distance) on $\mathcal{P}^p(X)$ is defined by

$$d_p^W(\mu_1, \mu_2) := \left( \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{X \times X} d^p(x, y) d\pi(x, y) \right)^{1/p}.$$

It is known that $d_p^W$ is a metric on $\mathcal{P}^p(X)$, is complete resp. separable whenever $d$ is complete resp. separable and that $\mathcal{P}_0(X)$ is $d_p^W$-dense in $\mathcal{P}^p(X)$ [1,16]. Furthermore, it follows from the Hölder inequality that $d_p^W \leq d_{p'}^W$ whenever $p \leq p'$. The last observation makes possible the definition of $d_\infty^W(\mu_1, \mu_2) = \lim_{p \to \infty} d_p^W(\mu_1, \mu_2)$ on $\mathcal{P}^\infty(X)$. The limit is finite on the space $\mathcal{P}^\infty(X)$ of measures with bounded support and yields a metric space, complete if $X$ is a complete metric space. The closure of $\mathcal{P}_0(X)$ in $\mathcal{P}^\infty(X)$ is $\mathcal{P}_{cp}(X)$, the set of probability measures with compact support, and this fact leads to our focus on measures having compact support.

**Remark 2.2.** Alternatively the $\infty$-metric is given by

$$d_\infty^W(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \sup\{d(x, y) : (x, y) \in \mathrm{supp}(\pi)\}. \tag{2.1}$$

For the case that $\mu = (1/n) \sum_{i=1}^n \delta_{x_i}$ and $\nu = (1/n) \sum_{i=1}^n \delta_{y_i}$, the equation (2.1) reduces to

$$d_\infty^W(\mu, \nu) = \min_{\sigma \in S^n} \max\{d(x_j, y_{\sigma(j)}) : 1 \leq j \leq n\}, \tag{2.2}$$

where $S^n$ is the permutation group on $\{1, \dots, n\}$.

We work primarily with $d_\infty^W$, which we henceforth write simply as $d_\infty$. An attractive feature of the $d_\infty$-metric on $\mathcal{P}_{cp}(X)$ is that a continuous map at the metric space level induces a continuous map at the $\mathcal{P}_{cp}$-level.

**Proposition 2.3.** *Let $f : X \to Y$ be a continuous map between metric spaces. Then $f_* : \mathcal{P}_{cp}(X) \to \mathcal{P}_{cp}(Y)$ is continuous in the $d_\infty$-topology.*

**Proof.** Let $\mu_n \to \mu$ in $\mathcal{P}_{cp}(X)$. Note that $\mathrm{supp}(f_*(\mu)) = f(\mathrm{supp}(\mu))$ is compact, so $f_*$ carries $\mathcal{P}_{cp}(X)$ into $\mathcal{P}_{cp}(Y)$. Let $\varepsilon > 0$. By standard compactness arguments applied to $f$ and $\mathrm{supp}(\mu)$, there exists $\delta > 0$ such that for $x \in \mathrm{supp}(\mu)$ and $d(x, z) < \delta$, we have $d(f(x), f(z)) < \varepsilon$. There exists $N$ such that $d_\infty(\mu_n, \mu) < \delta$ for $n \geq N$. For $n \geq N$, there exists by (2.1) $\pi \in \Pi(\mu, \mu_n)$ such that $\sup\{d(x, z) : (x, z) \in \mathrm{supp}(\pi)\} < \delta$. Then $(f \times f)_*(\pi) \in \Pi(f_*(\mu), f_*(\mu_n))$ and has support the compact set $(f \times f)(\mathrm{supp}(\pi))$.

For any $(f \times f)(x, z) = (f(x), f(z))$ in this support set, we have $d(f(x), f(z)) < \varepsilon$ since $d(x, z) < \delta$ by choice of $\pi$. It follows that

$$\begin{aligned} d_\infty(f_*(\mu), f_*(\mu_n)) &\leq \sup\{d(u, v) : (u, v) \in \operatorname{supp}((f \times f)_*(\pi))\} \\ &= \sup\{d(f(x), f(z)) : (x, z) \in \operatorname{supp}(\pi)\} < \varepsilon. \end{aligned}$$

We conclude that $f_*(\mu_n) \to f_*(\mu)$ in $\mathcal{P}_{cp}(Y)$ and thus that $f_*$ is continuous. □

## 3. Means and barycenters

We begin this section by recalling (Definition 3.1 through Proposition 3.5) several needed notions and results from Section 3 of [8].

**Definition 3.1.**

(1) An *n-mean* $G_n$ on a set $X$ for $n \geq 2$ is a function $G_n : X^n \to X$ that is idempotent in the sense that $G_n(x, \ldots, x) = x$ for all $x \in X$.
(2) An $n$-mean $G_n$ is *symmetric* or *permutation invariant* if for each permutation $\sigma$ of $\{1, \ldots, n\}$, $G_n(\mathbf{x}_\sigma) = G_n(\mathbf{x})$, where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{x}_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. A (symmetric) *mean* $G$ on $X$ is a sequence of means $\{G_n\}$, one (symmetric) mean for each $n \geq 2$.
(3) A *barycentric map* or *barycenter* on the set of finitely supported uniform measures $\mathcal{P}_0(X)$ is a map $\beta : \mathcal{P}_0(X) \to X$ satisfying $\beta(\delta_x) = x$ for each $x \in X$.

For $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$, we let

$$\mathbf{x}^k = (x_1, \ldots, x_n, x_1, \ldots, x_n, \ldots, x_1, \ldots, x_n) \in X^{nk}, \tag{3.3}$$

where the number of $\mathbf{x}$-blocks is $k$. We define the *carrier* $S(\mathbf{x})$ of $\mathbf{x}$ to be the set of entries in $\mathbf{x}$, i.e., the smallest finite subset $F$ such that $\mathbf{x} \in F^n$. We set $[\mathbf{x}]$ equal to the equivalence class of all $n$-tuples obtained by permuting the coordinates of $\mathbf{x} = (x_1, \ldots, x_n)$. Note that the operation $[\mathbf{x}]^k = [\mathbf{x}^k]$ is well-defined and that all members of $[\mathbf{x}]$ all have the same carrier set $S(\mathbf{x})$.

A tuple $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$ *induces* on $S(\mathbf{x})$ a uniform probability measure $\mu$ with finite support by $\mu = \sum_{i=1}^n (1/n)\delta_{x_i}$, where $\delta_{x_i}$ is the point measure of mass 1 at $x_i$. Since the tuple may contain repetitions of some of its entries, each singleton set $\{x\}$ for $x \in \{x_1, \ldots, x_n\}$ will have measure $k/n$, where $k$ is the number of times that it appears in the listing $x_1, \ldots, x_n$. Note that every member of $[\mathbf{x}]$ induces the same finitely supported probability measure.

**Lemma 3.2.** *For each probability measure $\mu$ on $X$ with finite support $F$ for which $\mu(x) (= \mu(\{x\}))$ is rational for each $x \in F$, there exists a unique $[\mathbf{x}]$ inducing $\mu$ such that any $[\mathbf{y}]$ inducing $\mu$ is equal to $[\mathbf{x}]^k$ for some $k \geq 1$.*

**Definition 3.3.** A mean $G = \{G_n\}$ on $X$ is said to be *intrinsic* if it is symmetric and for all $n, k \geq 2$ and all $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$,

$$G_n(\mathbf{x}) = G_{nk}(\mathbf{x}^k).$$

We have the following corollary to Lemma 3.2.

**Corollary 3.4.** *Let $G$ be an intrinsic mean. Then for any finitely supported probability measure $\mu$ with support $F$ and taking on rational values, we may define $\beta_G(\mu) = G_n(\mathbf{x})$, for any $\mathbf{x} \in F^n$ that induces $\mu$.*

Corollary 3.4 provides the basis for the following equivalence.

**Proposition 3.5.** *There is a one-to-one correspondence between the intrinsic means and the barycentric maps on $\mathcal{P}_0(X)$ given in one direction by assigning to an intrinsic mean $G$ the barycentric map $\beta_G$ and in the reverse direction assigning to a barycentric map $\beta$ the mean $G_n(x_1, \ldots, x_n) = \beta(\frac{1}{n}\sum_{i=1}^n \delta_{x_i})$.*

We specialize to means and barycenters in metric spaces.

**Definition 3.6.** An $n$-mean $G_n : X^n \to X$ is said to be *subadditive* if for all $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in X^n$,

$$d(G_n(\mathbf{x}), G_n(\mathbf{y})) \leq \frac{1}{n}\sum_{j=1}^n d(x_j, y_j).$$

An $n$-mean is said to be *submaxitive* if

$$d(G_n(\mathbf{x}), G_n(\mathbf{y})) \leq \max\{d(x_j, y_j) : 1 \leq j \leq n\}$$

A mean $G = \{G_n\}$ is said to be subadditive resp. submaxitive if each $G_n$ is.

In [16] Sturm considered the notion of a contractive barycentric map for the Wasserstein metric $d_1^W$ on the set of probability measures of finite first moment on a complete metric space. This notion readily generalizes.

**Definition 3.7.** Let $(X, d)$ be a metric space. A *contractive barycentric map* on $\mathcal{P}^1(X)$ is a map $\beta : \mathcal{P}^1(X) \to X$ satisfying $\beta(\delta_x) = x$ for all $x \in X$ and $d(\beta(\mu), \beta(\nu)) \leq d_1^W(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}^1(X)$. A contractive barycentric map on $\mathcal{P}_{cp}(X)$ is one that is contractive for $d_\infty$.

The following is part of Proposition 2.7 of [8].

**Proposition 3.8.** *A subadditive intrinsic mean $G$ on a metric space $X$ uniquely gives rise to a contractive barycentric map on $\mathcal{P}_0(X)$. If $X$ is complete, the barycentric map uniquely extends to a contractive barycentric map $\beta_G : \mathcal{P}^1(X) \to X$ and from $\mathcal{P}_{cp}(X)$ to $X$ for the case that $G$ is submaxitive.*

**Remark 3.9.** Note that a $d_1^W$-contractive barycentric map restricts to a $d_\infty$-contractive barycentric map on $\mathcal{P}_{cp}(X)$ since $d_1^W \leq d_\infty$.

## 4. Some basic examples

We recall the following basic example, which appears as Example 2.9 of [8].

**Example 4.1.** We consider the arithmetic mean $A_n(x_1, \ldots, x_n)$ for $n \geq 2$ on a Banach space $E$ equipped with the norm metric. This mean is intrinsic and easily seen to be subadditive, and hence uniquely extends to a contractive barycentric map $\mathcal{A} : \mathcal{P}^1(E) \to E$. We note from [9,16] that $\mathcal{P}^1(E)$ is the set of Radon measures $\mu$ on $E$ satisfying $\int_E ||x|| d\mu(x) < \infty$. For each $\mu \in \mathcal{P}^1(E)$, the identity map on $E$ is Bochner $\mu$-integrable and $\mathcal{A}(\mu) = \int_E x \, d\mu(x)$.

Let $C$ be an open cone in a Banach space $E$ such that its closure $\overline{C}$ is a closed normal cone in $E$. The cone $\overline{C}$ defines a partial order on $E$ given by $x \leq y$ if and only if $y - x \in \overline{C}$. The Thompson metric on $C$ is given by $d(x, y) = r$ if and only if $e^r = \min\{t \geq 0 : x \leq ty,\ y \leq tx\}$. It is a standard and basic result that the Thompson metric is a complete metric and that the Thompson metric topology agrees with the relative norm topology [17,12,13].

**Example 4.2.** The hypotheses of Proposition 3.8 hold for the special case that $p = \infty$ and $G$ is the arithmetic mean restricted to an open cone of a Banach space equipped with the Thompson metric for which the closure is a normal cone; see [7, Proposition 2.4]. (Note in particular that the arithmetic mean is submaxitive, but not subadditive with respect to the Thompson metric.) Hence there exists an $d_\infty$-contractive map $\mathcal{A}_\infty : \mathcal{P}_{cp}(C) \to C$. Since the Thompson metric topology agrees with the relative norm topology the inclusion $(C, d) \hookrightarrow (E, \|\cdot\|)$ is continuous, hence from Proposition 2.3 the inclusion $\mathcal{P}_{cp}(C) \hookrightarrow \mathcal{P}_{cp}(E)$ is continuous, and thus the composition

$$\mathcal{P}_{cp}(C) \hookrightarrow \mathcal{P}_{cp}(E) \hookrightarrow \mathcal{P}^1(E) \xrightarrow{\mathcal{A}} E$$

is continuous, the second arrow being continuous since $d_1^W \leq d_\infty$. The composition $\mathcal{P}_{cp}(C) \xrightarrow{\mathcal{A}_\infty} C \hookrightarrow E$ is continuous since $\mathcal{A}_\infty$ is continuous and the Thompson metric topology and relative norm topology agree on $C$. Since both compositions are continuous and are both the usual arithmetic mean on the dense subset $\mathcal{P}_0(C)$, we see that they are equal on all of $\mathcal{P}_{cp}(C)$. Note that the image of the second composition is contained in $C$, so the same is true of the first composition.

The following, which essentially appears as Proposition 4.10 of [14], gives a convenient sufficient condition for a mean to be submaxitive.

**Proposition 4.3.** *Suppose $C$ is an open cone in a Banach space $E$ such that its closure $\overline{C}$ is a closed normal cone in $E$. A mean $G_n : C^n \to C$ is submaxitive if it is monotonic and subhomogeneous, i.e., satisfies for $c \geq 1$, $G_n(cx_1, \ldots, cx_n) \leq cG_n(x_1, \ldots, x_n)$.*

**Proof.** Let $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in C^n$. Let $r_i = d(x_i, y_i)$, the metric $d$ being the Thompson metric (the definition is recalled just below). Then $x_i \leq e^{r_i} y_i$ and $y_i \leq e^{r_i} x_i$ for all $i$. Set $M = \max_{1\leq i\leq n}\{e^{r_i}\}$. Using monotonicity and subhomogeneity, we obtain

$$G_n(\mathbf{x}) \leq G_n(My_1, \ldots, My_n) \leq MG_n(\mathbf{y}),$$

and similarly $G_n(\mathbf{y}) \leq MG_n(\mathbf{x})$, so $d(G_n(\mathbf{x}), G_n(\mathbf{y})) \leq \log M = \max_{1\leq i\leq n}\{r_i\} = \max_{1\leq i\leq n}\{d(x_i, y_i)\}$. Hence $G_n$ is submaxitive. □

Let $\mathbb{A}$ be a $C^*$-algebra with identity, $\mathcal{S}(\mathbb{A})$ the closed subspace of self-adjoint elements, and let $\mathbb{P} = \mathbb{P}(\mathbb{A}) \subseteq \mathcal{S}(\mathbb{A})$ be the open convex cone of strictly positive elements. The group $\mathrm{GL}(\mathbb{A})$ of invertible elements acts on $\mathbb{P}$ via congruence transformations: $\Gamma_c(x) = cxc^*$. For $x, y \in \mathcal{S}(\mathbb{A})$, we write $x \leq y$ if $y - x \in \overline{C}$, and $x < y$ if $y - x \in C$.

For $a, b \in \mathbb{P}$ and $t \in \mathbb{R}$, the $t$-weighted geometric mean of $a$ and $b$ is defined by

$$a\#_t b = a^{1/2}(a^{-1/2}ba^{-1/2})^t a^{1/2}. \tag{4.4}$$

Some basic properties of the $t$-weighted mean are

(i) (Loewner–Heinz inequality) $a\#_t b \leq c\#_t d$ for $a \leq c, b \leq d$ and $t \in [0, 1]$;
(ii) $m(a\#_t b)m^* = (mam^*)\#_t(mbm^*)$ for $m \in \mathrm{GL}(\mathbb{A})$;
(iii) $a\#_t b \leq (1 - t)a + tb$ for $t \in [0, 1]$.

For $t = 1/2$, $a\#_{1/2}b = b\#_{1/2}a$ is called simply the geometric mean of $a$ and $b$ and denoted $a\#b$.

The Thompson metric on $\mathbb{P}$ is defined by $d(a, b) = \max\{\log M(b/a), \log M(a/b)\}$, where $M(b/a) = \inf\{\alpha > 0 : b \leq \alpha a\}$ and coincides with $d(a, b) = ||\log(a^{-1/2}ba^{-1/2})||$. Furthermore, $a\#b$ is a midpoint of $a$ and $b$ in the Thompson metric and $t \mapsto a\#_t b$, $0 \leq t \leq 1$, is a metric geodesic from $a$ to $b$.

The logarithm map $\log : \mathbb{P} \to E := \mathcal{S}(\mathbb{A})$ is differentiable and is contractive from the exponential metric increasing (EMI) property ([5,11])

$$||\log x - \log y|| \leq d(x, y), \quad x, y \in \mathbb{P}. \tag{4.5}$$

This property reflects the seminegative curvature of the Thompson metric, which can be realized as a Banach–Finsler metric arising from the Banach space norm on $\mathcal{S}(\mathbb{A})$: for $a \in \mathbb{P}$, the Finsler norm of $v \in T_a\mathbb{P} = E$ is given by $\|v\| = \|a^{-1/2}va^{-1/2}\|$ and the exponential and logarithm maps are

$$\exp_a(v) = a^{1/2}\exp(a^{-1/2}va^{-1/2})a^{1/2}, \tag{4.6}$$

$$\log_a(x) = a^{1/2}\log(a^{-1/2}xa^{-1/2})a^{1/2}. \tag{4.7}$$

**Example 4.4.** The *Karcher mean* $\Lambda = \{\Lambda_n\}$ on $\mathbb{P}$ is defined as the unique solution in $\mathbb{P}$ of the Karcher equation

$$x = \Lambda_n(a_1, \ldots a_n) \Longleftrightarrow \sum_{i=1}^{n} \log(x^{-1/2}a_i x^{-1/2}) = 0.$$

It has been shown in [7] for $C^*$-algebras that are monotone complete that this equation does indeed have a unique solution in $\mathbb{P}$ and that the resulting mean $\Lambda_n$ for $n \geq 2$ has the following properties:

(i) $\Lambda_n$ is symmetric and idempotent;
(ii) (Monotonicity) If $b_i \leq a_i$ for all $1 \leq i \leq n$, then $\Lambda_n(b_1, \ldots, b_n) \leq \Lambda_n(a_1, \ldots, a_n)$;
(iii) (Subadditivity) $d(\Lambda_n(a_1, \ldots, a_n), \Lambda_n(b_1, \ldots, b_n)) \leq (1/n)\sum_{i=1}^{n} d(a_i, b_i)$, where $d$ is the Thompson metric.

We note also that the Karcher mean $\Lambda$ is intrinsic since the left hand side of the Karcher equation for $(a_1, \ldots, a_n)^k$ is just $k$ times that for $(a_1, \ldots, a_n)$, and hence still equal to 0 for the same $x$. We thus have the following.

**Proposition 4.5.** *Proposition 3.8 yields a uniquely determined contractive barycentric map* $\beta_\Lambda : \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ *satisfying* $\beta((1/n)\sum_{i=1}^{n} \delta_{a_i}) = \Lambda_n(a_1, \ldots, a_n)$.

## 5. Integrals

Each barycentric map gives rise to an associated theory of integration over probability measures. We consider some elementary properties of this integration for barycentric maps for the previously considered cases that $\beta$ is defined on $\mathcal{P}^1(X)$ or on $\mathcal{P}_{cp}(X)$.

**Definition 5.1.** Let $X$ be a metric space and let $\beta$ be a contractive barycentric map. Let $(M, \mathcal{M}, P)$ be a measure space equipped with a probability $P$. For $f : M \to X$ measurable, we define

$$\int_M f\,dP = \int_M f(x)dP(x) := \beta(f_*(P)),$$

provided the push-forward $f_*(P)$ is in the domain of $\beta$. In the latter case we call $f$ *integrable*.

We have the following general "change of variables" formula.

**Proposition 5.2.** *Let $(M, \mathcal{M}, P)$ be a probability measure space, $X$ and $Y$ metric spaces equipped with barycentric maps, and $f : M \to X$, $g : X \to Y$ Borel measurable maps.*

(i) $g_*(f_*(P)) = (g \circ f)_*(P)$;
(ii) $\int_X g\, df_*(P) = \int_M g \circ f\, dP$, *provided either integral exists.*

**Proof.** Item (i) follows directly from the definition of the push-forward map. For item (ii) we observe

$$\int\limits_X g\, df_*(P) = \beta(g_*(f_*(P))) = \beta((g \circ f)_*(P)) = \int\limits_M g \circ f\, dP.$$

The outer equalities hold by definition, the inner one by (i). Since by (i) the two probabilities to which $\beta$ is applied are equal, the last assertion of (ii) follows. □

We have the following general variant of $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$.

**Lemma 5.3.** *Let $\beta : \mathcal{P}^1(X) \to X$ be a contractive barycentric map, where $\mathcal{P}^1(X)$ is equipped with the Wasserstein metric $d_1^W$, let $(M, \mathcal{M}, P)$ be a probability measure space, and let $f, g : M \to X$ be integrable maps. Then*

$$d\left(\int\limits_M f\, dP, \int\limits_M g\, dP\right) \leq \int\limits_M d(f(x), g(x))\, dP(x)$$

*and for $p = \infty$,*

$$d\left(\int\limits_M f\, dP, \int\limits_M g\, dP\right) \leq \sup\{d(f(x), g(x)) : x \in M\}$$

**Proof.** We observe that

$$\begin{aligned} d\left(\int\limits_M f\, dP, \int\limits_M g\, dP\right) &= d(\beta(f_*(P)), \beta(g_*(P))) \;\leq\; d_1^W(f_*(P), g_*(P)) \\ &\leq \int\limits_{X \times X} d(x, y)\, d(f \times g)_*(P)(x, y) \\ &= \int\limits_M d(f(x), g(x))\, dP(x), \end{aligned}$$

where the first inequality follows from the contractivity of $\beta$, the second from the fact $(f \times g)_*(P)$ is a coupling for $f_*(P)$ and $g_*(P)$, and the last equality follows from change of variables. For the case $p = \infty$, we obtain by similar reasoning

$$d\left(\int\limits_M f\, dP, \int\limits_M g\, dP\right) = d(\beta(f_*(P)), \beta(g_*(P))) \;\leq\; d_\infty(f_*(P), g_*(P))$$

$$\leq \sup\{d(y,z) : (y,z) \in \operatorname{supp}((f \times g)_*(P))\}$$
$$\leq \sup\{d(f(x), g(x)) : x \in M\},$$

where the last inequality follows from $\operatorname{supp}(f \times g)_*(P) \subseteq \overline{(f \times g)(M)}$ and the last supremum is taken over a dense subset of $\overline{(f \times g)(M)}$. □

For the following result, which gives a general condition for integrability, see Lemma 3.2 of [8].

**Lemma 5.4.** *Let $f : X \to Y$ be a Lipschitz map with Lipschitz constant $C$. Then $f_* : \mathcal{P}^p(X) \to \mathcal{P}^p(Y)$ is Lipschitz with Lipschitz constant $C$ for $1 \leq p \leq \infty$.*

The following useful result for the $\mathcal{P}_{cp}$-case follows from Proposition 2.3.

**Proposition 5.5.** *Let $f : X \to Y$ be a continuous function between metric spaces, where $Y$ is equipped with a contractive barycentric map $\beta_Y$ on $\mathcal{P}_{cp}(Y)$. Then $\int_X f\,d\mu$ exists for any $\mu \in \mathcal{P}_{cp}(X)$. Furthermore, if $\mu_n$ converges to $\mu$ in the $d_\infty$-topology of $\mathcal{P}_{cp}(X)$, then $\int_X f\,d\mu_n \to \int_X f\,d\mu$ in $Y$.*

**Proof.** By Proposition 2.3, $f_*$ carries $\mathcal{P}_{cp}(X)$ continuously into $\mathcal{P}_{cp}(Y)$. In particular, $\beta_Y(f_*(\mu)) = \int_X f\,d\mu$ exists. By continuity of $f_*$ and $\beta_Y$,

$$\int_X f\,d\mu_n = \beta_Y(f_*(\mu_n)) \to \beta_Y(f_*(\mu)) = \int_X f\,d\mu. \quad \square$$

We also need the following variant of the preceding result.

**Lemma 5.6.** *Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces, let $\beta : \mathcal{P}_{cp}(Y) \to Y$ be a contractive barycentric map, and let $F : \mathcal{P}_{cp}(X) \times X \to Y$ be continuous. Let $\mu_n \to \mu$ in $\mathcal{P}_{cp}(X)$ and suppose there exists a compact set $K \subseteq X$ such that supp$(\mu_n) \subseteq K$ for each $n$. If $x_n \to x$ in $X$, then $\int_X F(x_n, y)d\mu_n(y) \to \int_X F(x,y)d\mu(y)$ in $Y$.*

**Proof.** Without loss of generality we may assume that $\operatorname{supp}(\mu) \subseteq K$ by taking its union with $K$ if necessary. The convergent sequence $\{\mu_n\}$ together with its limit $\mu$ form a compact subset of $\mathcal{P}_{cp}(X)$; call it $J$. Then $J \times K$ is a compact subset of $\mathcal{P}_{cp}(X) \times X$, and hence $F$ restricted to $J \times K$ is uniformly continuous, where the metric $d$ chosen on $\mathcal{P}_{cp}(X) \times X$ is the sum of the coordinate metrics. For $\varepsilon > 0$, choose $\delta > 0$ such that for $(\nu_1, x_1), (\nu_2, x_2) \in J \times K$, $d((\nu_1, x_1), (\nu_2, x_2)) < \delta$ implies $d_Y(F(\nu_1, x_1), F(\nu_2, x_2)) < \varepsilon$. There exists $N$ such that $d_\infty(\mu_n, \mu) < \delta$ for $n \geq N$, and hence $d((\mu_n, x), (\mu, x)) < \delta$ for each $x \in X$. We thus have $d_Y(F(\mu_n, x), F(\mu, x)) < \varepsilon$ for each $x \in K$ and $n \geq N$. We note also by our assumption on the supports that integrals over $X$ for each $\mu_n$ and $\mu$ can be reduced to integrals over $K$. It follows from Lemma 5.3 (taking $f = F(\mu_n, \cdot)$ and $g = F(\mu, \cdot)$) for $n \geq N$ that

$$d_Y\left(\int_X F(\mu_n, x)d\mu_n(x), \int_X F(\mu, x)d\mu_n(x)\right)$$
$$= d_Y\left(\int_K F(\mu_n, x)d\mu_n(x), \int_K F(\mu, x)d\mu_n(x)\right) \leq \varepsilon.$$

Also by Proposition 5.5, $\int_X F(\mu,x)d\mu_n(x) \to \int_X F(\mu,x)d\mu(x)$, so for large $n$,

$$d_Y\left(\int_X F(\mu,x)d\mu_n(x), \int_X F(\mu,x)d\mu(x)\right) \leq \varepsilon.$$

Combining the previous two displays and applying the triangle inequality, we conclude for large enough $n$ that

$$d_Y\left(\int_X F(\mu_n,x)d\mu_n(x), \int_X F(\mu,x)d\mu(x)\right) \leq 2\varepsilon. \quad \square$$

**Remark 5.7.** For a curve $\gamma : [0,1] \to X$ on a metric space $X$ equipped with a contractive barycentric map $\beta$ on $\mathcal{P}_{cp}(X)$, the integral

$$\int_0^1 \gamma(t)\ d\mu(t) := \beta(\gamma_*(\mu))$$

exists for $\mu \in \mathcal{P}^1([0,1])$. For Lebesgue measure $m$, we simply write $\int_0^1 \gamma(t)\ dt$.

For $x, y \in \mathbb{P}$, the cone of positive invertible elements of a $C^*$-algebra, $\gamma(t) := x\#_t y$ is a minimal geodesic for the Thompson metric, that is,

$$d(\gamma(t), \gamma(s)) = d(x,y)|t-s|$$

and hence $\gamma$ is Lipschitz with the Lipschitz constant $d(x,y)$. By Lemma 5.4 and Proposition 5.5, we have a continuous map from $\mathcal{P}^1([0,1])$ to $\mathbb{P}$:

$$\mu \mapsto \int_0^1 x\#_t y\ d\mu(t) := \beta_\Lambda(\gamma_*(\mu)).$$

The preceding gives rise to a (separately) continuous map from Lemma 5.3

$$\mathbb{P}^2 \times \mathcal{P}^1([0,1]) \to \mathbb{P}, \quad (x,y,\mu) \mapsto \int_0^1 x\#_t y\ d\mu(t).$$

## 6. The power mean

In this section we let $\mathbb{P}$ denote the open convex cone of positive definite $n \times n$-matrices, or more generally the open cone (open in the space of self-adjoint elements) of positive invertible elements of a $C^*$-algebra $\mathbb{A}$ with identity $e$ equipped with the Thompson metric. We have seen in Example 4.2 that there is a $\infty$-contractive arithmetic barycentric map $\beta_\mathcal{A} : \mathcal{P}^\infty(\mathbb{P}) \to \mathbb{P}$, where $\mathcal{P}^\infty(\mathbb{P})$ is endowed with the $d_\infty$-metric arising from the Thompson metric.

We can use the integration of the previous section to extend the power means on $\mathbb{P}$ to the Borel measures $\mathcal{P}_{cp}(\mathbb{P})$ of compact support. Note that in the case of the cone of positive definite matrices these agree with the measures of bounded support.

First we define $F(x, y) = x\#_t y$, the weighted geometric mean, and set for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$

$$g(x) = \int_{\mathbb{P}} F(x, y)\, d\mu(y) = \int_{\mathbb{P}} x\#_t y\, d\mu(y).$$

For fixed $x$ and $0 < t < 1$, it is well-known that the maps $y \mapsto x\#_t y$ and $x \mapsto x\#_t y$ are strictly contractive with respect to the Thompson metric $d$; indeed $d(x\#_t y, x\#_t z) \leq td(y, z)$ and $d(w\#_t y, x\#_t y) \leq (1-t)d(w, x)$. In particular, both maps are continuous. By Proposition 5.5,

$$\beta_{\mathcal{A}}(F(x, \cdot)_*(\mu)) = \int_{\mathbb{P}} F(x, y)\, d\mu(y) = \int_{\mathbb{P}} x\#_t y\, d\mu(y)$$

exists. We next define the power mean for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$.

**Definition 6.1.** For the positive cone $\mathbb{P}$ and $0 < t \leq 1$, we define the *power mean* $P_t : \mathcal{P}_{cp}(\mathbb{P}) \to \mathbb{P}$ by $P_t(\mu) = x$, where $x$ is the unique fixed point of the equation

$$x = \int_{\mathbb{P}} x\#_t y\, d\mu(y). \tag{6.8}$$

**Remark 6.2.** For the case $\mu = (1/n)\sum_{i=1}^{n} \delta_{x_i}$, the equation (6.8) reduces to

$$\begin{aligned} x = \int_{\mathbb{P}} x\#_t y\, d\mu(y) &= \beta_{\mathcal{A}}\left(F(x, \cdot)_*\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}\right)\right) \\ &= \beta_{\mathcal{A}}\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{x\#_t x_i}\right) = \frac{1}{n}\sum_{i=1}^{n} x\#_t x_i. \end{aligned}$$

This calculation shows that for the case of means, equivalently uniform probability measures with finite support, this definition collapses to the one appearing in [7, Section 3]. We note that if $x_i$'s mutually commute, then the equation $x = \frac{1}{n}\sum_{i=1}^{n} x\#_t x_i$ has the unique solution in $\mathbb{P}$ given by $x = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^t\right)^{1/t}$.

To establish existence and uniqueness of the power mean in the Borel measure setting, we need to establish existence and uniqueness of the solution to equation (6.8).

**Lemma 6.3.** *The map $f(x) = \int_{\mathbb{P}} x\#_t y\, d\mu(y)$ for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ is a strict contraction, and hence has a unique fixed point.*

**Proof.** By Theorem 3.1 of [7] and its proof (and the calculation of the preceding remark) the map $f$ satisfies $d(f(w), f(x)) \leq (1 - t)d(w, x)$ for all $x, w \in \mathbb{P}$, provided $\mu \in \mathcal{P}_0(\mathbb{P})$. For general $\mu$ by density of $\mathcal{P}_0(\mathbb{P})$ in $\mathcal{P}_{cp}(\mathbb{P})$, we may find a sequence $\{\mu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ that converges to $\mu$ with respect to $d_\infty$. By Proposition 5.5

$$\begin{aligned} d(f(w), f(x)) &= d\left(\int_{\mathbb{P}} w\#_t y\, d\mu(y), \int_{\mathbb{P}} x\#_t y\, d\mu(y)\right) \\ &= \lim_n d\left(\int_{\mathbb{P}} w\#_t y\, d\mu_n(y), \int_{\mathbb{P}} x\#_t y\, d\mu_n(y)\right) \leq (1 - t)d(w, x). \end{aligned}$$

Since the Thompson metric on $\mathbb{P}$ is complete, the lemma follows from the Banach fixed point theorem. □

Since the Thompson metric on $\mathbb{P}$ satisfies $d(y^{-1}, z^{-1}) = d(y, z)$ for any $y, z \in \mathbb{P}$, the map $g(x) = \left[\int_{\mathbb{P}} (x\#_t y)^{-1}\, d\mu(y)\right]^{-1}$ for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in (0, 1]$ is a strict contraction, and hence has a unique fixed point by the Banach fixed point theorem.

**Definition 6.4.** For $-1 \le t < 0$, we define the *power mean* $P_t : \mathcal{P}_{cp}(\mathbb{P}) \to \mathbb{P}$ by $P_t(\mu) = x$, where $x$ is the unique fixed point of the equation

$$x = \left[\int_{\mathbb{P}} (x\#_{-t} y)^{-1}\, d\mu(y)\right]^{-1}. \tag{6.9}$$

We introduce $t$-th powers on $\mathcal{P}(\mathbb{P})$ for $t \in \mathbb{R} \setminus \{0\}$. For $\mathcal{O} \in \mathcal{B}(\mathbb{P})$ we let

$$\mu^t(\mathcal{O}) := \mu(\mathcal{O}^{\frac{1}{t}}), \tag{6.10}$$

where $\mathcal{O}^t := \{a^t : a \in \mathcal{O}\}$. Note that $\mu^t \in \mathcal{P}_{cp}(\mathbb{P})$ whenever $\mu \in \mathcal{P}_{cp}(\mathbb{P})$. In terms of push-forward measures, $\mu^t = f_*(\mu)$, where $f(x) = x^t$. For $t \in [-1, 1] \setminus \{0\}$ the map $f : \mathbb{P} \to \mathbb{P}$ given by $f(x) = x^t$ is a Lipschitz function with Lipschitz constant $|t|$ with respect to the Thompson metric. By Lemma 5.4 the push-forward measure $\mu^t = f_*(\mu) : \mathcal{P}^p(\mathbb{P}) \to \mathcal{P}^p(\mathbb{P})$ is also the Lipschitz function with Lipschitz constant $|t|$ for $1 \le p \le \infty$.

**Remark 6.5.** Since $(x\#_t y)^{-1} = x^{-1}\#_t y^{-1}$, the equation (6.9) can be written as

$$x^{-1} = \int_{\mathbb{P}} x^{-1}\#_{-t} y^{-1}\, d\mu(y) = \int_{\mathbb{P}} x^{-1}\#_{-t} y\, d\mu^{-1}(y).$$

So we have $P_t(\mu)^{-1} = x^{-1} = P_{-t}(\mu^{-1})$.

We recall the following well-known result; see e.g., [11] and [6, Section 4]. The notation $L(f)$ refers to the Lipschitz constant of $f$.

**Proposition 6.6.** *Let $(X, d)$ be a complete metric space, $0 \le \lambda < 1$, and $\mathcal{C}_\lambda(X) = \{f : X \to X : L(f) \le \lambda\}$. For $f \in \mathcal{C}_\lambda(X)$ let $p(f) \in X$ denote the unique fixed point of $f$. If we endow $\mathcal{C}_\lambda(X)$ with the topology of pointwise convergence, then the fixed point map $p : \mathcal{C}_\lambda(X) \to X$ is continuous.*

**Proposition 6.7.** *The power mean is contractive for the Thompson metric $d$, i.e., $d(P_t(\mu), P_t(\nu)) \le d_\infty(\mu, \nu)$ for $\mu, \nu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in [-1, 1] \setminus \{0\}$.*

**Proof.** By [7, Proposition 3.6] and Remark 6.2 for $\mu = (1/n)\sum_{i=1}^n \delta_{x_i}$ and $\nu = (1/n)\sum_{i=1}^n \delta_{y_i}$, $d(P_t(\mu), P_t(\nu)) \le \max\{d(x_i, y_i) : 1 \le i \le n\}$. Since $P_t$ is symmetric, it follows that $d(P_t(\mu), P_t(\nu)) \le d_\infty(\mu, \nu)$, see Remark 2.2. Since any two members of $\mathcal{P}_0(\mathbb{P})$ can be rewritten as uniform measures with finite support for a common $n$ (by appropriately dividing up the point masses of each), we conclude that $P_t$ is contractive on $\mathcal{P}_0(\mathbb{P})$ equipped with the $d_\infty$ metric. By standard metric space properties, $P_t$ uniquely extends to a contractive map $G_t : \mathcal{P}_{cp}(\mathbb{P}) \to \mathbb{P}$.

It remains to show that $G_t(\mu) = P_t(\mu)$ for all $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in [-1, 1] \setminus \{0\}$. Let $\{\mu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ be a sequence converging to $\mu$ in $\mathcal{P}_{cp}(\mathbb{P})$. For each $\nu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in (0, 1]$, define $F_\nu(x) = \int_{\mathbb{P}} x\#_t y\, d\nu(y)$. Then by Proposition 5.5 for any $x \in \mathbb{P}$,

$$F_{\mu_n}(x) = \int_{\mathbb{P}} x\#_t y\, d\mu_n(y) \to \int_{\mathbb{P}} x\#_t y\, d\mu(y) = F_\mu(x).$$

This shows that $F_{\mu_n}$ converges to $F_\mu$ pointwise. From Lemma 6.3 and Proposition 6.6 we conclude that $P_t(\mu_n)$, the fixed point of $F_{\mu_n}$ converges to $P_t(\mu)$, the fixed point of $F_\mu$. By the previous paragraph $P_t(\mu_n) \to G_t(\mu)$. Hence $G_t(\mu) = P_t(\mu)$ for $t \in (0,1]$. Applying the similar argument to $H_\nu(x) = \left[\int_{\mathbb{P}} (x\#_{-t}y)^{-1}\, d\nu(y)\right]^{-1}$ for $\nu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in (0,1]$ and using continuity of the inverse map $x \in \mathbb{P} \mapsto x^{-1} \in \mathbb{P}$, we conclude that $P_t(\mu_n)$, the fixed point of $H_{\mu_n}$ converges to $P_t(\mu)$, the fixed point of $H_\mu$. Hence $G_t(\mu) = P_t(\mu)$ for $t \in [-1,0)$. □

In the following, a subset $U \subseteq \mathbb{P}$ is called an upper set, if whenever $a \in U$ and $a \leq b$, then $b \in U$. We define partial order on the set of Borel probability measures, sometimes called the *stochastic order*, by $\mu \leq \nu$ for $\mu, \nu \in \mathcal{P}(\mathbb{P})$ if and only if $\mu(U) \leq \nu(U)$ for any upper Borel set $U$. Note that the power mean $P_t$ on $\mathcal{P}_0(\mathbb{P})$ is a contractive, monotonic intrinsic mean [7, Proposition 3.6]. The monotonicity extends to the corresponding barycentric map $P_t$ on $\mathcal{P}_{cp}(\mathbb{P})$. To show this we need the following lemma.

**Lemma 6.8.** *Given $\mu \in \mathcal{P}_{cp}(\mathbb{P})$, there exists a sequence $\{\mu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ such that $\mu_n \to \mu$ with respect to the $d_\infty$-metric, and $\mu \leq \mu_n$ for each $n$. Similarly there exists a sequence in $\mathcal{P}_0(\mathbb{P})$ converging to $\mu$ from below.*

**Proof.** One sees directly from the definition of the Thompson metric that the closed $\varepsilon$-ball around $x \in \mathbb{P}$ has largest element $e^\varepsilon x$, smallest element $e^{-\varepsilon}x$, and is equal to the order interval

$$[e^{-\varepsilon}x, e^\varepsilon x] = \{y \in \mathbb{P} : e^{-\varepsilon}x \leq y \leq e^\varepsilon x\}.$$

By the compactness of $\operatorname{supp}(\mu)$, there exist finitely many elements $x_1, \dots, x_n \in \operatorname{supp}(\mu)$ such that $\operatorname{supp}(\mu) \subseteq \bigcup_{i=1}^n [e^{-\varepsilon}x_i, e^\varepsilon x_i]$. Note that $x_i \leq w := \sum_{i=1}^n x_i$ for each $i$. Define $f : \mathbb{P} \to \mathbb{P}$ by $f(x) = e^\varepsilon x_i$, where $x \in [e^{-\varepsilon}x_i, e^\varepsilon x_i]$, but $x \notin [e^{-\varepsilon}x_j, e^\varepsilon x_j]$ for $j < i$. For all $x \notin \bigcup_{i=1}^n [e^{-\varepsilon}x_i, e^\varepsilon x_i]$, we pick some "trash collection" point $q$ and define $f(x) = q$. Since $x \leq f(x)$ for each $x \in \operatorname{supp}(\mu)$, and hence $B \cap \operatorname{supp}(\mu) \subseteq f^{-1}(B)$ for any upper Borel set $B$, it is easy to verify that $\mu \leq f_*(\mu)$ in the stochastic order.

Define $\lambda : \mathbb{P} \to \mathbb{P} \times \mathbb{P}$ by $\lambda(x) = (x, f(x))$. It follows directly that $\lambda_*(\mu) \in \Pi(\mu, f_*(\mu))$ and that $\operatorname{supp}(\lambda_*(\mu))$ is the closure of $\{(x,y) \in \operatorname{supp}(\mu) \times \mathbb{P} : y = f(x)\}$. Since for each $x \in \operatorname{supp}(\mu)$, $d(x, f(x)) \leq 2\varepsilon$ from the definition of $f$, we conclude that $d_\infty(\mu, f_*(\mu)) \leq 2\varepsilon$. Applying the preceding construction to $2\varepsilon = 1/n$ for each $n$ gives the desired sequence $\{\mu_n\}$. If we modify the definition of $f$ to $f(x) = e^{-\varepsilon}x_i$ for $x \in \bigcup_{i=1}^n [e^{-\varepsilon}x_i, e^\varepsilon x_i]$, we obtain a sequence converging to $\mu$ from below. □

**Theorem 6.9.** *Let $\mu, \nu \in \mathcal{P}_{cp}(\mathbb{P})$ and $t \in (0,1]$. Then $P_t(\mu) \leq P_t(\nu)$ whenever $\mu \leq \nu$.*

**Proof.** By Lemma 6.8 pick sequences $\{\mu_n\}, \{\nu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ such that $\mu_n \to \mu$ from below and $\nu_n \to \nu$ from above. Then $\mu_n \leq \nu_n$ and hence $P_t(\mu_n) \leq P_t(\nu_n)$ for each $n$, since $P_t$ is monotonic as a mean [7, Theorem 3.6(4)] and hence on members of $\mathcal{P}_0(\mathbb{P})$. Since $P_t$ is contractive by Proposition 6.7, hence continuous, it follows from the closedness of the Loewner partial order on $\mathbb{P}$ that $P_t(\mu) \leq P_t(\nu)$. □

We also show the monotonicity of power means in parameter $t \in [-1,1] \setminus \{0\}$.

**Theorem 6.10.** *For $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ and $0 < t \leq s \leq 1$,*

$$P_{-s}(\mu) \leq P_{-t}(\mu) \leq P_t(\mu) \leq P_s(\mu).$$

**Proof.** For general $\mu$ by density of $\mathcal{P}_0(\mathbb{P})$ in $\mathcal{P}_{cp}(\mathbb{P})$, we can find a sequence $\{\mu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ that converges to $\mu$ with respect to $d_\infty$. It has been shown in [7] that $P_{-s}(\mu_n) \leq P_{-t}(\mu_n) \leq P_t(\mu_n) \leq P_s(\mu_n)$. Taking the limit as $n \to \infty$ yields the desired inequalities, since the Loewner order is closed and the power mean is continuous. □

## 7. The Karcher barycenter

For positive definite matrices $a_1, \dots, a_n$ of the same dimension the *Karcher mean*, or the least squares mean, $\Lambda(a_1, \dots, a_n)$ is defined as the unique minimizer of the sum of squares of the Riemannian distances to each of $a_1, \dots, a_n$. That is,

$$\Lambda(a_1, \dots, a_n) = \arg\min_{x \in \mathbb{P}} \sum_{j=1}^{n} d^2(x, a_j),$$

where $d(a, b) = \| \log(a^{-1/2} b a^{-1/2}) \|_2$ denotes the Riemannian distance between $a$ and $b$. Furthermore, it has been shown [2] that the Karcher mean $\Lambda(a_1, \dots, a_n)$ is the unique positive definite solution $x$ of the Karcher equation

$$\sum_{j=1}^{n} \log(x^{-1/2} a_j x^{-1/2}) = 0. \tag{7.11}$$

One has no Riemannian metric on the open cone of positive invertible operators of an infinite-dimensional Hilbert space. Nevertheless, Lawson and Lim [7] have defined the Karcher mean $\Lambda(a_1, \dots, a_n)$ of positive invertible operators as the unique positive solution of the Karcher equation (7.11) and have successfully established a generalization of matrix power means to the setting of positive invertible operators. Indeed the theory extends to the open cone (open in the space of self-adjoint elements) of positive invertible elements of a monotone complete $C^*$-algebra $\mathbb{A}$ with identity equipped with the Thompson metric, our setting for the remainder of this section. Moreover, the Karcher mean is intrinsic and contractive with respect to the Thompson metric, so by Proposition 3.8 there exists a unique contractive barycentric map $\beta_\Lambda : \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ satisfying

$$\beta_\Lambda \left( \frac{1}{n} \sum_{j=1}^{n} \delta_{a_j} \right) = \Lambda(a_1, \dots, a_n).$$

We call $\beta_\Lambda(\mu)$ for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ the *Karcher barycenter* of $\mu$. In this section we study the Karcher barycenter for the probability measures with compact support on $\mathbb{P}$.

We have from (4.5) and Lemma 5.4 that the logarithm map induces the Lipschitz map with Lipschitz constant 1

$$\log_* : \mathcal{P}^p(\mathbb{P}) \to \mathcal{P}^p(\mathcal{S}(\mathbb{A})), 1 \le p \le \infty.$$

For $\mu \in \mathcal{P}^1(\mathbb{P})$ and $x \in \mathbb{P}$, let $\mu_x := g_*(\mu) \in \mathcal{P}^1(\mathbb{P})$, where $g : \mathbb{P} \to \mathbb{P}$ defined by $g(a) = x^{-1/2} a x^{-1/2}$. Then

$$\int_{\mathbb{P}} \log(x^{-1/2} a x^{-1/2}) d\mu(a) = \int_{\mathbb{P}} \log a \; d\mu_x(a) = \beta_{\mathcal{A}}(\log_* \mu_x), \tag{7.12}$$

where the second integral is the Bochner integral from Example 4.1. We define the Karcher equation for $\mu \in \mathcal{P}^1(\mathbb{P})$;

$$\int_{\mathbb{P}} \log(x^{-1/2} a x^{-1/2}) \; d\mu(a) = 0. \tag{7.13}$$

In terms of the Banach–Finsler structure on $\mathbb{P}$, the Karcher equation is equivalent to

$$\int_{\mathbb{P}} \exp_x^{-1}(a)\ d\mu(a) = 0. \tag{7.14}$$

**Theorem 7.1.** *The barycenter $\beta_\Lambda(\mu)$ for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ is a solution of the Karcher equation*

$$\int_{\mathbb{P}} \log(x^{-1/2}ax^{-1/2})d\mu(a) = 0. \tag{7.15}$$

**Proof.** Let $\{\mu_n\}$ be a sequence in $\mathcal{P}_0(\mathbb{P})$ that converges to $\mu$. By working in the compact metric space $A = \mathrm{supp}(\mu)$ with probability measure $\mu|_A$ and using the density of $\mathcal{P}_0(A)$ in $\mathcal{P}_{cp}(A)$, we can assume that each $\mu_n$ has support contained in $A$. For each $\mu_n = (1/n_k)\sum_{i=1}^{n_k}\delta_{a_i}$, we have

$$\begin{aligned}\int_{\mathbb{P}} \log\left((\beta_\Lambda(\mu_n))^{-1/2}a(\beta_\Lambda(\mu_n))^{-1/2}\right)d\mu_n(a) &= \frac{1}{n_k}\sum_{i=1}^{n_k}\log\left((\beta_\Lambda(\mu_n))^{-1/2}a_i(\beta_\Lambda(\mu_n))^{-1/2}\right)\\ &= 0\end{aligned}$$

Let $F : \mathcal{P}_{cp}(\mathbb{P}) \times \mathbb{P} \to \mathcal{S}(\mathbb{A})$ be defined by $F(\mu, a) = \log\left((\beta_\Lambda(\mu))^{-1/2}a(\beta_\Lambda(\mu))^{-1/2}\right)$, a continuous function, since $\beta_\Lambda$ is contractive by Proposition 3.8, hence continuous, and taking powers, products and the logarithm map are continuous. Then by Lemma 5.6

$$\begin{aligned}0 &= \lim_n \int_{\mathbb{P}} \log\left((\beta_\Lambda(\mu_n))^{-1/2}a(\beta_\Lambda(\mu_n))^{-1/2}\right)d\mu_n(a)\\ &= \int_{\mathbb{P}} \log\left((\beta_\Lambda(\mu))^{-1/2}a(\beta_\Lambda(\mu))^{-1/2}\right)d\mu(a). \quad \square\end{aligned}$$

**Remark 7.2.** Since the preparation and submission of this manuscript, Y. Lim and M. Pálfia have shown that the solution of the Karcher equation (7.15) is unique, even in the most general setting of $\mu \in \mathcal{P}^1(\mathbb{P})$ [10].

**Theorem 7.3.** *The Karcher barycenter $\beta_\Lambda(\mu)$ for $\mu \in \mathcal{P}_{cp}(\mathbb{P})$ is invariant under inversion and congruence transformations, that is,*

$$\begin{aligned}\beta_\Lambda(\mu)^{-1} &= \beta_\Lambda(\mu^{-1}),\\ m\beta_\Lambda(\mu)m^* &= \beta_\Lambda(m.\mu), \quad m \in \mathrm{GL}(\mathbb{A}),\end{aligned}$$

*where $\mu^{-1}$ and $m.\mu$ are the push-forward of $\mu$ under inversion and the congruence transformation by $m$; $a \mapsto mam^*$.*

**Proof.** The formula is known for the Karcher mean from [7], hence for measures in $\mathcal{P}_0(\mathbb{P})$. The theorem follows from the density of $\mathcal{P}_0(\mathbb{P})$, the continuity of $\beta_\Lambda$, and Proposition 2.3 applied to the inversion map. A similar argument holds for the congruence transformations. $\square$

**Theorem 7.4.** *For $\mu \in \mathcal{P}_{cp}(\mathbb{P})$,*

$$\lim_{t\to 0} P_t(\mu) = \beta_\Lambda(\mu),$$

*in the strong topology of $\mathbb{P}$.*

**Proof.** We first consider the case $t > 0$. Let $\mu \in \mathcal{P}_{cp}(\mathbb{P})$. By Theorem 6.10 $P_t(\mu) \leq P_s(\mu)$ for $0 < t \leq s$. We claim that $\beta_\Lambda(\mu)$ is the greatest lower bound of $\{P_t(\mu) : t > 0\}$, which will guarantee the strong convergence, since $\mathbb{A}$ is a monotone complete $C^*$-algebra. By Lemma 6.8 there exists a sequence $\{\mu_n\} \subseteq \mathcal{P}_0(\mathbb{P})$ such that $\mu_n \to \mu$ with respect to the $d_\infty$-metric and $\mu \leq \mu_n$ for each $n$. From the continuity of $\beta_\Lambda$ and of $P_t$ (Proposition 6.7), we have $\beta_\Lambda(\mu_n) \to \beta_\Lambda(\mu)$ and $P_t(\mu_n) \to P_t(\mu)$. Since from [7] for $t > 0$, $\beta_\Lambda(\mu_n) \leq P_t(\mu_n)$ for each $n$, we have by the closedness of the Loewner order that

$$\beta_\Lambda(\mu) = \lim_n \beta_\Lambda(\mu_n) \leq \lim_n P_t(\mu_n) = P_t(\mu).$$

Thus $\beta_\Lambda(\mu)$ is a lower bound for $\{P_t(\mu) : 0 < t\}$. Suppose that $\nu$ is another lower bound. Then for any $n$ and $t > 0$, $\nu \leq P_t(\mu) \leq P_t(\mu_n)$, the last inequality by the monotonicity of $P_t$. It follows that $\nu \leq \lim_{t\to 0^+} P_t(\mu_n) = \beta_\Lambda(\mu_n)$, the equality coming from [7], and hence $\nu \leq \lim_n \beta_\Lambda(\mu_n) = \beta_\Lambda(\mu)$. Hence $\beta_\Lambda(\mu)$ is the greatest lower bound. Since $\mathbb{A}$ is monotone complete, the desired result follows for $t > 0$. A similar argument obtains for $t < 0$. □

Next, we introduce a curve $\delta_x \#_t \mu$ on $\mathcal{P}^1(\mathbb{P})$ from the Dirac measure $\delta_x$ to $\mu$ and establish a fixed point theorem associated with the curve $\delta_x \#_t \mu$. Let $x \in \mathbb{P}, t \geq 0$ and $\mu \in \mathcal{P}^1(\mathbb{P})$. Define $x\#_t\mu \in \mathcal{P}^1(\mathbb{P})$ by $x\#_t\mu = f_*(\mu)$, where $f(a) = x\#_t a$. Note that $x\#_0\mu = \delta_x$ and $x\#_1\mu = \mu$. For $t > 0$, since $x\#_t z = a$ if and only if $z = x\#_{1/t}a$,

$$(x\#_t\mu)(\mathcal{O}) := \mu(\{x\#_{1/t}a : a \in \mathcal{O}\}).$$

For example, if $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{a_j} \in \mathcal{P}_0(\mathbb{P})$ then $x\#_t\mu = \frac{1}{n}\sum_{j=1}^n \delta_{x\#_t a_j}$.

The following shows in particular the continuity of $t \mapsto x\#_t\mu$.

**Lemma 7.5.** *For $\mu, \nu \in \mathcal{P}^1(\mathbb{P})$ and $t, s \in [0, 1]$,*

$$d_1^W(x\#_t\mu, y\#_s\nu) \leq (1-t)d(x, y) + t d_1^W(\mu, \nu) + |t - s| d_1^W(\delta_y, \nu).$$

**Proof.** Use $d(a\#_t b, c\#_t b) \leq (1-t)d(a, c) + t d(b, d) + |t - s| d(c, d)$ and for $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$, $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, and $1 \leq p < \infty$

$$d_1^W(\mu, \nu) = \min_{\sigma \in S^n} \left( \frac{1}{n} \sum_{j=1}^n d(x_j, y_{\sigma(j)}) \right). \quad \square$$

For $\mu \in \mathcal{P}^1(\mathbb{P})$ and $t > 0$, we set $\mu_x = x^{-1/2}.\mu (= g_*(\mu))$, where $g(a) = x^{-1/2} a x^{-1/2}$ and $\mu_x^t := (\mu_x)^t$. Note that $\mu_x^t = (h \circ g)_*(\mu)$, where $h(a) = a^t$.

**Lemma 7.6.** *We have $x^{-1/2}\beta_\Lambda(x\#_t\mu)x^{-1/2} = \beta_\Lambda(\mu_x^t)$.*

**Proof.** One can directly see that for $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{a_j} \in \mathcal{P}_0(\mathbb{P})$, $\mu_x^t = \frac{1}{n}\sum_{j=1}^n \delta_{(x^{-1/2}a_j x^{-1/2})^t}$ and thus

$$\begin{aligned} x^{-1/2}\beta_\Lambda(x\#_t\mu)x^{-1/2} &= x^{-1/2}\Lambda(x\#_t a_1, \ldots, x\#_t a_n)x^{-1/2} \\ &= \Lambda((x^{-1/2}a_1 x^{-1/2})^t, \ldots, (x^{-1/2}a_n x^{-1/2})^t) \\ &= \beta_\Lambda\left( \frac{1}{n}\sum_{j=1}^n \delta_{(x^{-1/2}a_j x^{-1/2})^t} \right) = \beta_\Lambda(\mu_x^t). \end{aligned}$$

By continuity of $\beta_\Lambda$ and the preceding lemma, passing to the limit yields that it holds for all $\mu \in \mathcal{P}^1(\mathbb{P})$. □

**Theorem 7.7.** *For $t \in (0,1)$ and $\mu \in \mathcal{P}^1(\mathbb{P})$, $\beta_\Lambda(\mu)$ is the unique solution of*

$$x = \beta_\Lambda(x\#_t\mu) \tag{7.16}$$

*which is equivalent to $e = \beta_\Lambda(\mu_x^t)$.*

**Proof.** Define $F : \mathbb{P} \to \mathbb{P}$ by $F(x) = \beta_\Lambda(x\#_t\mu)$. For $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{a_j}$,

$$\begin{aligned} d(F(x), F(y)) &= d\left(\beta_\Lambda\left(\frac{1}{n}\sum_{j=1}^n \delta_{x\#_t a_j}\right), \beta_\Lambda\left(\frac{1}{n}\sum_{j=1}^n \delta_{y\#_t a_j}\right)\right) \\ &= d(\Lambda(x\#_t a_1, \ldots, x\#_t a_n), \Lambda(y\#_t a_1, \ldots, y\#_t a_n)) \\ &\leq \frac{1}{n}\sum_{j=1}^n d(x\#_t a_j, y\#_t a_j) \leq (1-t)d(x,y) \end{aligned}$$

where the last inequality follows from $d(z\#_t y, x\#_t y) \leq (1-t)d(z,x)$. Moreover from (Theorem 6.3, [7]), $\beta_\Lambda(\mu) = \Lambda(a_1, \ldots, a_n)$ is the unique solution of

$$x = \Lambda(x\#_t a_1, \ldots, x\#_t a_n) = \beta_\Lambda(x\#_t\mu).$$

Pick a sequence $\{\mu_n\} \subset \mathcal{P}_0(\mathbb{P})$ converging to $\mu$ in $\mathcal{P}^1(\mathbb{P})$. Then

$$\begin{aligned} d(F(x), F(y)) &= \lim_{n\to\infty} d(\beta_\Lambda(x\#_t\mu_n), \beta_\Lambda(y\#_t\mu_n)) \\ &\leq (1-t)\lim_{n\to\infty} d(x,y) = (1-t)d(x,y) \end{aligned}$$

which shows that $F$ is a strict contraction for the Thompson metric and hence $x = F(x)$ has a unique solution. Moreover, $F(\beta_\Lambda(\mu)) = \lim_{n\to\infty} F(\beta_\Lambda(\mu_n)) = \lim_{n\to\infty} \beta_\Lambda(\mu_n) = \beta_\Lambda(\mu)$.

By Lemma 7.6 the equation (7.16) is equivalent to $e = x^{-1/2}\beta_\Lambda(x\#_t\mu)x^{-1/2} = \beta_\Lambda(\mu_x^t)$. This completes the proof. □

**Theorem 7.8.** *Suppose that for each $\mu \in \mathcal{P}_{cp}(\mathbb{P})$, there exists sufficiently small $t > 0$ such that*

$$0 = \int_{\mathbb{P}} \log(x^{-1/2}ax^{-1/2})^t \, d\mu(a) = \int_{\mathbb{P}} \log a \, d\mu_x^t(a)$$

*has a unique solution in $\mathbb{P}$. Then the Karcher equation (7.15) has a unique solution.*

**Proof.** Let $\mu \in \mathcal{P}_{cp}(\mathbb{P})$. By Theorem 7.1, $\beta_\Lambda(\mu)$ is a solution of the Karcher equation $0 = \int_{\mathbb{P}} \log(x^{-1/2}ax^{-1/2})^t \, d\mu(a)$. Suppose that $w$ is another solution of the Karcher equation. Then

$$0 = t\int_{\mathbb{P}} \log(w^{-1/2}aw^{-1/2}) \, d\mu(a) = \int_{\mathbb{P}} \log(w^{-1/2}aw^{-1/2})^t \, d\mu(a) = \int_{\mathbb{P}} \log a \, d\mu_w^t(a)$$

for any $t > 0$, where the last equality follows from the change of variables. From $e^{-1/2}ae^{-1/2} = eae = a$ and hypothesis, $e = \beta_\Lambda(\mu_w^t)$ for a sufficiently small $t > 0$. By Theorem 7.7, $w = \beta_\Lambda(\mu)$. □

**Remark 7.9.** The hypothesis is valid for $\mu \in \mathcal{P}_0(\mathbb{P})$ by using Implicit Function Theorem ([7]). Note that $\mathrm{supp}(\mu_w^t) = \left(w^{-1/2}\mathrm{supp}(\mu)w^{-1/2}\right)^t$. By compactness of the support of $\mu$, $d(\mathrm{supp}(\mu_w^t), e) \to 0$ as $t \to 0^+$.

**Remark 7.10.** In light of recent work of M. Pálfia [15], the restriction in this section to monotone $C^*$-algebras can be dropped, i.e., the results remain valid for general unital $C^*$-algebras.

## Acknowledgments

The work of S. Kim was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2015R1C1A1A02036407). The work of Y. Lim was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) No.NRF-2015R1A3A2031159.

## References

[1] F. Bolley, Separability and completeness for the Wasserstein distance, in: Séminaire de Probabilités XLI, in: Lecture Notes in Mathematics, vol. 1934, 2008, pp. 371–377.
[2] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. 30 (1977) 509–541.
[3] S. Kim, H. Lee, The power mean and the least squares mean of probability measures on the space of positive definite matrices, Linear Algebra Appl. 465 (2015) 325–346.
[4] J. Lawson, Ordered probability spaces, J. Math. Anal. Appl. 455 (2017) 167–179.
[5] J. Lawson, Y. Lim, Symmetric spaces with convex metrics, Forum Math. 19 (2007) 571–602.
[6] J. Lawson, Y. Lim, A Lipschitz constant formula for vector addition in cones with applications to Stein-like equations, Positivity 16 (2012) 81–95.
[7] J. Lawson, Y. Lim, Karcher means and Karcher equations of positive operators, Trans. Amer. Math. Soc. Ser. B 1 (2014) 1–22.
[8] J. Lawson, Y. Lim, Contractive barycentric maps, J. Operator Theory 77 (2017) 87–107.
[9] M. Ledoux, M. Talagrand, Probability in Banach Spaces. Isoperimetry and Process, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 23, Springer, Berlin, 1991.
[10] Y. Lim, M. Pálfia, Existence and uniqueness of the L1-Karcher mean, arXiv:1703.04292.
[11] K.-H. Neeb, A Cartan–Hadamard theorem for Banach–Finsler manifolds, Geom. Dedicata 95 (2002) 115–156.
[12] R.D. Nussbaum, Hilbert's Projective Metric and Iterated Nonlinear Maps, Memoirs of Amer. Math. Soc., vol. 391, 1988.
[13] R.D. Nussbaum, Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations, Differential Integral Equations 7 (1994) 1649–1707.
[14] M. Pálfia, Operator means of probability measures and generalized Karcher equations, Adv. Math. 289 (2016) 951–1007.
[15] M. Pálfia, Loewner's theorem in several variables, arXiv:1405.5076.
[16] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, Paris, 2002, in: Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2003, pp. 357–390.
[17] A.C. Thompson, On certain contraction mappings in a partially ordered vector space, Proc. Amer. Math. Soc. 14 (1963) 438–443.