View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

SciVerse ScienceDirect

ELSEVIER

Journal of Functional Analysis 262 (2012) 1498–1514

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Matrix power means and the Karcher mean

Yongdo Lim [a,*], Miklós Pálfia [b]

[a] *Department of Mathematics, Kyungpook National University, Taegu 702-701, Republic of Korea*

[b] *Department of Automation and Applied Informatics, Budapest University of Technology and Economics, 1117 Budapest, Magyar Tudósok krt. 2, Qép., Hungary*

Received 17 September 2011; accepted 9 November 2011

Available online 23 November 2011

Communicated by Alain Connes

Abstract

We define a new family of matrix means $\{P_t(\omega; \mathbb{A})\}_{t\in[-1,1]}$, where ω and $\mathbb{A}$ vary over all positive probability vectors in $\mathbb{R}^n$ and n-tuples of positive definite matrices resp. Each of these means except $t \neq 0$ arises as a unique positive definite solution of a non-linear matrix equation, satisfies all desirable properties of power means of positive real numbers and interpolates between the weighted harmonic and arithmetic means. The main result is that the Karcher mean coincides with the limit of power means as $t \to 0$. This provides not only a sequence of matrix means converging to the Karcher mean, but also a simple proof of the monotonicity of the Karcher mean, conjectured by Bhatia and Holbrook, and other new properties, which have recently been established by Lawson and Lim and also Bhatia and Karandikar using probabilistic methods on the metric structure of positive definite matrices equipped with the trace metric.
© 2011 Elsevier Inc. All rights reserved.

Keywords: Positive definite matrix; Geometric mean; Monotonicity; Riemannian trace metric; Metric nonpositive curvature; Thompson metric; Power mean; Riemannian barycenter

1. Introduction

The Riemannian trace metric on the convex cone $\mathbb{P} = \mathbb{P}_m$ of $m \times m$ positive definite Hermitian matrices plays an important role in many applied areas involving matrix interpolation, filtering, estimation, optimization and averaging, where it has been increasingly recognized that the Euclidean distance is often not the most suitable for the set $\mathbb{P}$ and that working with the appropriate

* Corresponding author.
E-mail addresses: ylim@knu.ac.kr (Y. Lim), palfia.miklos@aut.bme.hu (M. Pálfia).

0022-1236/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2011.11.012

geometry does matter in computational problems. (Recall the trace metric distance between two positive definite matrices is given by $\delta(A,B) = (\sum_{i=1}^{k} \log^2 \lambda_i(A^{-1}B))^{\frac{1}{2}}$, where $\lambda_i(X)$ denotes the i-th eigenvalue of X in ascending order.) It turns out that the Riemannian geometry plays a key role particularly in the study of inversion invariant data averaging procedures in image processing, in radar detection and in brain-computer interfacing [5,4,29]. An attractive candidate of data averaging procedures is the least squares mean [24] of positive definite matrices. This mean has appeared under a variety of other designations: *Frechet mean*, *Cartan mean*, *Riemannian center of mass* [18], *Riemannian geometric mean* [29], or frequently, *Karcher mean* [14], the terminology we adopt. The *Karcher mean* of n positive definite matrices $A_1, \ldots, A_n$ is defined as the unique minimizer (provided it exists) of the sum of squares of the Riemannian trace metric distances to each of the A_i, i.e.,

$$\Lambda(A_1, \ldots, A_n) = \operatorname*{arg\,min}_{X \in \mathbb{P}} \sum_{i=1}^{n} \delta^2(X, A_i). \tag{1.1}$$

This idea had been anticipated by Élie Cartan (see, for example, Section 6.1.5 of [6]), who showed among other things such a unique minimizer exists if the points all lie in a convex ball in a Riemannian manifold; see also Karcher's paper [18]. Using Karcher's formula for the gradient of the objective function (Theorem 2.1 of [18]) or computing appropriate derivatives as in [10,28] yields that the Karcher mean coincides with the unique positive definite solution of the *Karcher equation*

$$\sum_{i=1}^{n} \log\bigl(X^{1/2} A_i^{-1} X^{1/2}\bigr) = 0. \tag{1.2}$$

Various numerical methods for the solution of (1.1) or (1.2) have been introduced in the literature: fixed point methods, optimization algorithms like Newton's method or a gradient descent method, and iterative methods; see [14] and references therein. Unfortunately neither an explicit expression nor an explicit sequence of matrix means converging directly to the Karcher mean is known. Nevertheless the monotonicity of the Karcher mean, conjectured by Bhatia and Holbrook [11] and one of key axiomatic properties of matrix geometric means, was recently established by Lawson and Lim [24] via a probabilistic convergence of approximations and by Bhatia and Karandikar [12] via some probabilistic counting arguments, both arguments depending heavily on basic inequalities for the Riemannian metric. In this paper we provide a more direct, non-probabilistic proof of the monotonicity of the Karcher mean that depends on finding a sequence of matrix means satisfying monotonicity that converge directly to the Karcher mean. The principal goal of this paper is to construct a particular family of matrix means, each with numerous desirable properties such as monotonicity, that converges to the Karcher mean and show that these properties are preserved in the limit.

The basic family of means we consider are the power means. The power mean $(\frac{a_1^t + \cdots + a_n^t}{n})^{\frac{1}{t}}$ of n positive real numbers $a_1, \ldots, a_n$ arises as the unique positive solution of the elementary equation $x = \frac{1}{n} \sum_{i=1}^{n} x^{1-t} a_i^t$, and converges to the geometric mean of $a_1, \ldots, a_n$ as $t \to 0$. In this paper we consider a matrix analogue of $x = \frac{1}{n} \sum_{i=1}^{n} x^{1-t} a_i^t$, namely

$$X = \frac{1}{n} \sum_{i=1}^{n} X \#_t A_i, \tag{1.3}$$

where $A \#_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$, the t-weighted geometric mean of A and B. We prove that for each $t \in (0,1]$, Eq. (1.3) has a unique positive definite solution, denoted by $P_t(A_1,\ldots,A_n)$, and show that each of these matrix means (called a power mean) arises as a unique fixed point of a strict contraction for the Thompson metric. We show these power means vary continuously with t and satisfy analogues of basic properties of power means of positive real numbers (e.g., monotonicity and joint concavity). We then establish that the Karcher mean is the limit of power means as $t \to 0$. This gives, in particular, a simple and non-probabilistic proof of monotonicity, joint concavity and other new properties of the Karcher mean recently established by Bhatia and Karandikar [12], and a globally convergent method for obtaining the Karcher mean by taking the limit of $X_k = P_{\frac{1}{k}}(A_1,\ldots,A_n)$. Moreover, together with $P_{-t}(A_1,\ldots,A_n) := P_t(A_1^{-1},\ldots,A_n^{-1})^{-1}$, this provides a complete extension of the power means of positive reals to positive definite matrices in the sense that the family of matrix means $\{P_t(A_1,\ldots,A_n)\}_{t\in[-1,1]}$ interpolates continuously between the harmonic ($t=-1$) and arithmetic ($t=1$) means with the Karcher mean appearing at $t=0$.

2. Riemannian and Thompson metrics

Let $\mathbb{H}$ be the space of Hermitian matrices of a fixed size m, and $\mathbb{P}$ the convex cone of positive definite Hermitian matrices. For $X, Y \in \mathbb{H}$, we write that $X \leqslant Y$ if $Y - X$ is positive semidefinite, and $X < Y$ if $Y - X$ is positive definite. The Frobenius norm $\|\cdot\|_2$ gives rise to the Riemannian structure on $\mathbb{P}$: $\langle X, Y\rangle_A = \mathrm{Tr}(A^{-1}XA^{-1}Y)$, where $A \in \mathbb{P}$, $X, Y \in T_A(\mathbb{P}) = \mathbb{H}$. The Riemannian metric distance is given by $\delta(A,B) = [\sum_{i=1}^m \log^2 \lambda_i(A^{-1}B)]^{\frac{1}{2}}$, where the $\lambda_i(X)$ denote the eigenvalues of X. Then $\mathbb{P}$ becomes a Cartan–Hadamard manifold, a simply connected complete Riemannian manifold with non-positive sectional curvature [20]. For $A, B \in \mathbb{P}$ and $t \in \mathbb{R}$, the t-weighted geometric mean of A and B is defined by $A \#_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$. The curve $t \mapsto A \#_t B$ yields the unique geodesic from A to B for the Riemannian metric and $A \# B = A \#_{1/2} B$ is the unique midpoint between A and B. The following properties for the weighted geometric mean are well known [19,24,22].

Lemma 2.1. *Let $A, B, C, D \in \mathbb{P}$ and let $t \in \mathbb{R}$. Then*

(i) $A \#_t B = A^{1-t}B^t$ *for* $AB = BA$, *and* $(aA) \#_t (bB) = a^{1-t}b^t(A \#_t B)$ *for* $a, b > 0$;
(ii) *(Löwner–Heinz inequality)* $A \#_t B \leqslant C \#_t D$ *for* $A \leqslant C$, $B \leqslant D$ *and* $t \in [0,1]$;
(iii) $M(A \#_t B)M^* = (MAM^*) \#_t (MBM^*)$ *for any non-singular* M;
(iv) $A \#_t B = B \#_{1-t} A$, $(A \#_t B)^{-1} = A^{-1} \#_t B^{-1}$;
(v) $(\lambda A + (1-\lambda)B) \#_t (\lambda C + (1-\lambda)D) \geqslant \lambda(A \#_t C) + (1-\lambda)(B \#_t D)$ *for* $\lambda, t \in [0,1]$;
(vi) $\det(A \#_t B) = \det(A)^{1-t}\det(B)^t$; *and*
(vii) $((1-t)A^{-1} + tB^{-1})^{-1} \leqslant A \#_t B \leqslant (1-t)A + tB$ *for* $t \in [0,1]$.

The Thompson metric on $\mathbb{P}$ is defined by $d_\infty(A,B) = \|\log(A^{-1/2}BA^{-1/2})\|_\infty$, where $\|X\|_\infty$ denotes the spectral norm of X. It is known that d_∞ is a complete metric on $\mathbb{P}$ and $d_\infty(A,B) = \max\{\log M(B/A), \log M(A/B)\}$, where $M(B/A) = \inf\{\alpha > 0\colon B \leqslant \alpha A\} = \lambda_1(A^{-1/2}BA^{-1/2})$, the largest eigenvalue of $A^{-1/2}BA^{-1/2}$. See [32,16].

Lemma 2.2. *(See [9,16].) We have*

(i) $d_\infty(A,B) = d_\infty(A^{-1},B^{-1}) = d_\infty(MAM^*, MBM^*)$ *for any* $M \in \mathrm{GL}(m,\mathbb{C})$;

(ii) $d_\infty(A \# B, A) = d_\infty(A \# B, B) = \frac{1}{2} d_\infty(A, B)$;
(iii) $d_\infty(A \#_t B, C \#_t D) \leqslant (1-t) d_\infty(A, C) + t d_\infty(B, D)$, $t \in [0, 1]$.

The following non-expansive property of addition for the Thompson metric will be useful for our purpose.

Lemma 2.3. *(See [23].) Let $A_i, B_i \in \mathbb{P}$ and let $t_i > 0$, $i = 1, 2, \ldots, n$. Then*

$$d_\infty\left(\sum_{i=1}^n t_i A_i, \sum_{i=1}^n t_i B_i\right) \leqslant \max_{1 \leqslant i \leqslant n} \bigl\{d_\infty(A_i, B_i)\bigr\}.$$

3. Matrix power means

We denote by Δ_n the simplex of positive probability vectors in $\mathbb{R}^n$ convexly spanned by the unit coordinate vectors.

Theorem 3.1. *Let $A_1, \ldots, A_n \in \mathbb{P}$ and let $\omega = (w_1, \ldots, w_n) \in \Delta_n$. Then for each $\mathbf{t} = (t_1, \ldots, t_n) \in (0, 1]^n$, the following equation has a unique positive definite solution*:

$$X = \sum_{i=1}^n w_i (X \#_{t_i} A_i). \tag{3.4}$$

Furthermore, the solution varies continuously over $\mathbf{t} \in (0, 1]^n$.

Proof. We will show that the map $f : \mathbb{P} \to \mathbb{P}$ defined by $f(X) = \sum_{i=1}^n w_i (X \#_{t_i} A_i)$ is a strict contraction with respect to the Thompson metric. Let $X, Y > 0$. By Lemma 2.2 and Lemma 2.3,

$$\begin{aligned} d_\infty\bigl(f(X), f(Y)\bigr) &\leqslant \max_{1 \leqslant i \leqslant n} \bigl\{d_\infty\bigl(w_i (X \#_{t_i} A_i), w_i (Y \#_{t_i} A_i)\bigr)\bigr\} \\ &\leqslant \max_{1 \leqslant i \leqslant n} \bigl\{d_\infty(X \#_{t_i} A_i, Y \#_{t_i} A_i)\bigr\} \\ &\leqslant \max_{1 \leqslant i \leqslant n} \bigl\{(1-t_i) d_\infty(X, Y)\bigr\} = \max_{1 \leqslant i \leqslant n} \bigl\{(1-t_i)\bigr\} d_\infty(X, Y). \end{aligned}$$

Since $\max_{1 \leqslant i \leqslant n}\{1 - t_i\} \in [0, 1)$, f is a strict contraction.

By the continuity of fixed points of strict contractions (see e.g. [30]), the solution of (3.4) varies continuously over $\mathbf{t} \in (0, 1]^n$. □

Definition 3.2 *(Matrix power means).* Let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$ and $\omega \in \Delta_n$. For $t \in (0, 1]$, we denote by $P_t(\omega; \mathbb{A})$ the unique solution of

$$X = \sum_{i=1}^n w_i (X \#_t A_i). \tag{3.5}$$

For $t \in [-1, 0)$, we define $P_t(\omega; \mathbb{A}) = P_{-t}(\omega; \mathbb{A}^{-1})^{-1}$, where $\mathbb{A}^{-1} = (A_1^{-1}, \ldots, A_n^{-1})$. We call $P_t(\omega; \mathbb{A})$ the ω-weighted power mean of order t of $A_1, \ldots, A_n$. To simplify the notation we write $P_t(\mathbb{A}) = P_t(1/n, \ldots, 1/n; \mathbb{A})$.

Remark 3.3. We note that $P_1(\omega;\mathbb{A})=\sum_{i=1}^n w_iA_i$ and $P_{-1}(\omega;\mathbb{A})=(\sum_{i=1}^n w_iA_i^{-1})^{-1}$, the ω-weighted arithmetic and harmonic means of $A_1,\dots,A_n$, respectively. For $t\in[-1,0)$, $P_t(\omega;\mathbb{A})$ is the unique positive definite solution of

$$X=\left[\sum_{i=1}^n w_i(X\#_{-t}A_i)^{-1}\right]^{-1}. \tag{3.6}$$

Indeed, $X^{-1}=\sum_{i=1}^n w_i(X^{-1}\#_{-t}A_i^{-1})$ if and only if $X^{-1}=P_{-t}(\omega;\mathbb{A}^{-1})$.

Remark 3.4. Let $f:\mathbb{P}\to\mathbb{P}$ defined by $f(X)=\sum_{i=1}^n w_i(X\#_t A_i)$, $t\in(0,1]$. Then by the Löwner–Heinz inequality, f is monotone: $X\leqslant Y$ implies that $f(X)\leqslant f(Y)$. By Theorem 3.1, f is a strict contraction for the Thompson metric with the least contraction coefficient less than or equal to $1-t$.

For $\mathbb{A}=(A_1,\dots,A_n)\in\mathbb{P}^n$, $M\in\mathrm{GL}(m,\mathbb{C})$, $\mathbf{a}=(a_1,\dots,a_n)\in(0,\infty)^n$, $\omega=(w_1,\dots,w_n)\in\Delta_n$, and for a permutation σ on n-letters, we set

$$M\mathbb{A}M^*=\big(MA_1M^*,\dots,MA_nM^*\big),\qquad \mathbb{A}_\sigma=(A_{\sigma(1)},\dots,A_{\sigma_n}),$$

$$\mathbb{A}^{(k)}=(\underbrace{\mathbb{A},\dots,\mathbb{A}}_{k})\in\mathbb{P}^{nk},\qquad \omega^{(k)}=\frac{1}{k}(\underbrace{\omega,\dots,\omega}_{k})\in\Delta_{nk},$$

$$\mathbf{a}^t=\big(a_1^t,\dots,a_n^t\big),\qquad \omega\odot\mathbf{a}=\frac{1}{\sum_{i=1}^n w_ia_i}(w_1a_1,\dots,w_na_n)\in\Delta_n,$$

$$\hat{\omega}=\frac{1}{1-w_n}(w_1,\dots,w_{n-1})\in\Delta_{n-1},\qquad \mathbf{a}\cdot\mathbb{A}=(a_1A_1,\dots,a_nA_n).$$

We list some basic properties of $P_t(\omega;\mathbb{A})$.

Proposition 3.5. *Let* $\mathbb{A}=(A_1,\dots,A_n)$, $\mathbb{B}=(B_1,\dots,B_n)\in\mathbb{P}^n$, $\omega\in\Delta_n$, $\mathbf{a}=(a_1,\dots,a_n)\in(0,\infty)^n$ *and let* $s,t\in[-1,1]\setminus\{0\}$.

(1) $P_t(\omega;\mathbb{A})=(\sum_{i=1}^n w_iA_i^t)^{\frac{1}{t}}$ *if the* A_i*'s commute*;
(2) $P_t(\omega;\mathbf{a}\cdot\mathbb{A})=(\sum_{i=1}^n w_ia_i^t)^{\frac{1}{t}}P_t(\omega\odot\mathbf{a}^t;\mathbb{A})$;
(3) $P_t(\omega_\sigma;\mathbb{A}_\sigma)=P_t(\omega;\mathbb{A})$ *for any permutation* σ;
(4) $P_t(\omega;\mathbb{A})\leqslant P_t(\omega;\mathbb{B})$ *if* $A_i\leqslant B_i$ *for all* $i=1,2,\dots,n$;
(5) $d_\infty(P_t(\omega;\mathbb{A}),P_t(\omega;\mathbb{B}))\leqslant\max_{1\leqslant i\leqslant n}\{d_\infty(A_i,B_i)\}$;
(6) $(1-u)P_{|t|}(\omega;\mathbb{A})+uP_{|t|}(\omega;\mathbb{B})\leqslant P_{|t|}(\omega;(1-u)\mathbb{A}+u\mathbb{B})$ *for any* $u\in[0,1]$;
(7) $P_t(\omega;M\mathbb{A}M^*)=MP_t(\omega;\mathbb{A})M^*$ *for any invertible matrix* M;
(8) $P_t(\omega;\mathbb{A}^{-1})^{-1}=P_{-t}(\omega;\mathbb{A})$;
(9) $\mathrm{Det}(P_{-|t|}(\omega;\mathbb{A}))\leqslant\prod_{i=1}^n\mathrm{Det}(A_i)^{w_i}\leqslant\mathrm{Det}(P_{|t|}(\omega;\mathbb{A}))$;
(10) $(\sum_{i=1}^n w_iA_i^{-1})^{-1}\leqslant P_t(\omega;\mathbb{A})\leqslant\sum_{i=1}^n w_iA_i$;
(11) $P_t(\omega^{(k)};\mathbb{A}^{(k)})=P_t(\omega;\mathbb{A})$ *for any* $k\in\mathbb{N}$;
(12) $P_t(\omega;A_1,\dots,A_{n-1},X)=X$ *if and only if* $X=P_t(\hat{\omega};A_1,\dots,A_{n-1})$. *In particular,* $P_t(A_1,\dots,A_n,X)=X$ *if and only if* $X=P_t(A_1,\dots,A_n)$;

(13) *For* $s \in (0,1]$, $P_t(\omega; X \#_s A_1, \ldots, X \#_s A_n) = X$ *if and only if* $X = P_{st}(\omega; \mathbb{A})$;
(14) *If* $t \in (0,1]$, *then* $\Phi(P_t(\omega; \mathbb{A})) \leqslant P_t(\omega; \Phi(\mathbb{A}))$ *for any positive unital linear map* Φ, *where* $\Phi(\mathbb{A}) = (\Phi(A_1), \ldots, \Phi(A_n))$. *If* $t \in [-1, 0)$, *then* $\Phi(P_t(\omega; \mathbb{A})) \geqslant P_t(\omega; \Phi(\mathbb{A}))$ *for any strictly positive unital linear map* Φ; *and*
(15) *For any unitarily invariant norm* $|||\cdot|||$ *and* $t \in (0,1]$,

$$\big|\big|\big|P_t(\omega; \mathbb{A})\big|\big|\big| \leqslant \left[\sum_{i=1}^{n} w_i |||A_i|||^t\right]^{\frac{1}{t}} \quad and \quad \big|\big|\big|P_{-t}(\omega; \mathbb{A})\big|\big|\big| \geqslant \left[\sum_{i=1}^{n} w_i \big|\big|\big|A_i^{-1}\big|\big|\big|^t\right]^{-\frac{1}{t}}.$$

Proof. (1) Suppose that the A_i's commute. Let $t \in (0,1]$ and $X = (\sum_{i=1}^n w_i A_i^t)^{1/t}$. Then $X \#_t A_i = X^{1-t} A_i^t$ and $\sum_{i=1}^n w_i (X \#_t A_i) = \sum_{j=1}^n w_i X^{1-t} A_i^t = X^{1-t} \sum_{i=1}^n A_i^t = X^{1-t} X^t = X$. By uniqueness, $(\sum_{i=1}^n w_i A_i^t)^{1/t} = X = P_t(\omega; \mathbb{A})$. Furthermore, $P_{-t}(\omega; \mathbb{A}) = P_t(\omega; \mathbb{A}^{-1})^{-1} = (\sum_{i=1}^n w_i A_i^{-t})^{-1/t}$.

(2) Let $t \in (0,1]$. Set $\beta = (\sum_{i=1}^n w_i a_i^t)^{\frac{1}{t}}$, $\zeta = \omega \odot \mathbf{a}^t \in \Delta_n$ and $X = P_t(\omega \odot \mathbf{a}^t; \mathbb{A})$. Then $\zeta_i = (\omega \odot \mathbf{a}^t)_i = \frac{1}{\sum_{i=1}^n w_i a_i^t} w_i a_i^t$ and $X = \sum_{i=1}^n \zeta_i (X \#_t A_i)$. Therefore,

$$\begin{aligned}\sum_{i=1}^{n} w_i\big((\beta X) \#_t (a_i A_i)\big) &= \sum_{i=1}^{n} \beta^{1-t} w_i a_i^t (X \#_t A_i) = \sum_{i=1}^{n} \beta^{1-t} \zeta_i \beta^t (X \#_t A_i)\\ &= \beta \sum_{i=1}^{n} \zeta_i (X \#_t A_i) = \beta X.\end{aligned}$$

By uniqueness, $(\sum_{i=1}^n w_i a_i^t)^{\frac{1}{t}} P_t(\omega \odot \mathbf{a}^t; \mathbb{A}) = \beta X = P_t(\omega; \mathbf{a} \cdot \mathbb{A})$.

For $t \in [-1, 0)$, we have

$$\begin{aligned}P_t(\omega; \mathbf{a} \cdot \mathbb{A}) &= P_{-t}\big(\omega; (\mathbf{a} \cdot \mathbb{A})^{-1}\big)^{-1} = \left[\left(\sum_{i=1}^{n} w_i a_i^t\right)^{-\frac{1}{t}} P_{-t}\big(\omega \odot \mathbf{a}^t; \mathbb{A}^{-1}\big)\right]^{-1}\\ &= \left(\sum_{i=1}^{n} w_i a_i^t\right)^{\frac{1}{t}} P_{-t}\big(\omega \odot \mathbf{a}^t; \mathbb{A}^{-1}\big)^{-1} = \left(\sum_{i=1}^{n} w_i a_i^t\right)^{\frac{1}{t}} P_t\big(\omega \odot \mathbf{a}^t; \mathbb{A}\big).\end{aligned}$$

(3) Follows from the defining Eqs. (3.5) and (3.6).

(4) Suppose that $A_i \leqslant B_i$ for all $i = 1, 2, \ldots, n$. Let $t \in (0,1]$. Define

$$f(X) = \sum_{i=1}^{n} w_i (X \#_t A_i) \quad \text{and} \quad g(X) = \sum_{i=1}^{n} w_i (X \#_t B_i).$$

Then $P_t(\omega; \mathbb{A}) = \lim_{k\to\infty} f^k(X)$ and $P_t(\omega; \mathbb{B}) = \lim_{k\to\infty} g^k(X)$ for any $X \in \mathbb{P}$, by the Banach fixed point theorem. By the Löwner–Heinz inequality, $f(X) \leqslant g(X)$ for all $X \in \mathbb{P}$, and $f(X) \leqslant f(Y)$, $g(X) \leqslant g(Y)$ whenever $X \leqslant Y$. Let $X_0 > 0$. Then $f(X_0) \leqslant g(X_0)$ and $f^2(X_0) = f(f(X_0)) \leqslant g(f(X_0)) \leqslant g^2(X_0)$. Inductively, we have $f^k(X_0) \leqslant g^k(X_0)$ for all $k \in \mathbb{N}$. Therefore, $P_t(\omega; \mathbb{A}) = \lim_{k\to\infty} f^k(X_0) \leqslant \lim_{k\to\infty} g^k(X_0) = P_t(\omega; \mathbb{B})$.

Let $t \in [-1, 0)$. Then $\mathbb{A}^{-1} \geqslant \mathbb{B}^{-1}$ and thus $P_{-t}(\omega; \mathbb{A}^{-1}) \geqslant P_{-t}(\omega; \mathbb{B}^{-1})$. Therefore, $P_t(\omega; \mathbb{A}) = P_{-t}(\omega; \mathbb{A}^{-1})^{-1} \leqslant P_{-t}(\omega; \mathbb{B}^{-1})^{-1} = P_t(\omega; \mathbb{B})$.

(5) Let $t \in (0, 1]$. Let $X = P_t(\omega; \mathbb{A})$ and $Y = P_t(\omega; \mathbb{B})$. Then by Lemma 2.2 and Lemma 2.3,

$$\begin{aligned} d_\infty(X, Y) &= d_\infty\left(\sum_{i=1}^n w_i(X \#_t A_i), \sum_{i=1}^n w_i(Y \#_t B_i)\right) \\ &\leqslant \max_{1\leqslant i\leqslant n}\left\{d_\infty(X \#_t A_i, Y \#_t B_i)\right\} \leqslant \max_{1\leqslant i\leqslant n}\left\{(1-t)d_\infty(X, Y) + t d_\infty(A_i, B_i)\right\} \\ &= (1-t)d_\infty(X, Y) + t \max_{1\leqslant i\leqslant n}\left\{d_\infty(A_i, B_i)\right\}, \end{aligned}$$

which implies that $d_\infty(X, Y) \leqslant \max_{1\leqslant i\leqslant n}\{d_\infty(A_i, B_i)\}$. Since d_∞ is invariant under inversion, we also have

$$\begin{aligned} d_\infty\big(P_{-t}(\omega; \mathbb{A}), P_{-t}(\omega; \mathbb{B})\big) &= d_\infty\big(P_t\big(\omega; \mathbb{A}^{-1}\big)^{-1}, P_t\big(\omega; \mathbb{B}^{-1}\big)^{-1}\big) \\ &= d_\infty\big(P_t\big(\omega; \mathbb{A}^{-1}\big), P_t\big(\omega; \mathbb{B}^{-1}\big)\big) \\ &\leqslant \max_{1\leqslant i\leqslant n}\left\{d_\infty\big(A_i^{-1}, B_i^{-1}\big)\right\} = \max_{1\leqslant i\leqslant n}\left\{d_\infty(A_i, B_i)\right\}. \end{aligned}$$

(6) Let $t \in (0, 1]$. Let $X = P_t(\omega; \mathbb{A})$ and $Y = P_t(\omega; \mathbb{B})$. For $u \in [0, 1]$, we set $Z_u = (1-u)X + uY$. Let $f(Z) = \sum_{i=1}^n w_i(Z \#_t ((1-u)A_i + uB_i))$. Then by the joint concavity of the two-variable geometric mean

$$\begin{aligned} Z_u = (1-u)X + uY &= \sum_{i=1}^n w_i\big[(1-u)(X \#_t A_i) + u(Y \#_t B_i)\big] \\ &\leqslant \sum_{i=1}^n w_i\big(\big((1-u)X + uY\big) \#_t \big((1-u)A_i + uB_i\big)\big) = f(Z_u). \end{aligned}$$

Inductively, $Z_u \leqslant f^k(Z_u)$ for all $k \in \mathbb{N}$. Therefore, $(1-u)P_t(\omega; \mathbb{A}) + uP_t(\omega; \mathbb{B}) = Z_u \leqslant P_t(\omega; (1-u)\mathbb{A} + u\mathbb{B})$.

(7) Follows from the defining equation of $P_t(\omega; \mathbb{A})$ and the uniqueness of the positive definite solution.

(8) True by definition.

(9) Let $t \in (0, 1]$. Let $X = P_t(\omega; \mathbb{A})$. Then $X = \sum_{i=1}^n w_i(X \#_t A_i)$ implies that

$$\begin{aligned} \mathrm{Det}(X) &= \mathrm{Det}\left(\sum_{i=1}^n w_i(X \#_t A_i)\right) \geqslant \prod_{i=1}^n \mathrm{Det}(X \#_t A_i)^{w_i} \\ &= \prod_{i=1}^n \mathrm{Det}(X)^{(1-t)w_i}\mathrm{Det}(A_i)^{tw_i} = \mathrm{Det}(X)^{(1-t)}\left[\prod_{i=1}^n \mathrm{Det}(A_i)^{w_i}\right]^t, \end{aligned}$$

where the inequality follows by Corollary 7.6.9 of [17] for $n = 2$ and by an appropriate symmetrization method for $n > 2$, and thus $\mathrm{Det}(P_t(\omega; \mathbb{A})) \geqslant \prod_{i=1}^n \mathrm{Det}(A_i)^{w_i}$. From this, we also have

$$\operatorname{Det}\big(P_{-t}(\omega;\mathbb{A})\big)=\operatorname{Det}\big(P_t\big(\omega;\mathbb{A}^{-1}\big)^{-1}\big)=\big[\operatorname{Det}\big(P_{-t}\big(\omega;\mathbb{A}^{-1}\big)\big)\big]^{-1}$$
$$\leqslant\prod_{i=1}^{n}\big[\operatorname{Det}\big(A_i^{-1}\big)\big]^{-w_i}=\prod_{i=1}^{n}\operatorname{Det}(A_i)^{w_i}.$$

(10) Let $t\in(0,1]$. Let $X=P_t(\omega;\mathbb{A})$. By using the two-variable weighted arithmetic-geometric mean inequality, we obtain

$$X=\sum_{i=1}^{n}w_i(X\#_t A_i)\leqslant\sum_{i=1}^{n}w_i\big((1-t)X+tA_i\big)=(1-t)X+t\sum_{i=1}^{n}w_iA_i,$$

which implies that $X\leqslant\sum_{i=1}^{n}w_iA_i$. Similarly,

$$X=\sum_{i=1}^{n}w_i(X\#_t A_i)\geqslant\left[\sum_{i=1}^{n}w_i(X\#_t A_i)^{-1}\right]^{-1}=\left[\sum_{i=1}^{n}w_i\big(X^{-1}\#_t A_i^{-1}\big)\right]^{-1}.$$

Taking inverses of both sides leads to

$$X^{-1}\leqslant\sum_{i=1}^{n}w_i\big(X^{-1}\#_t A_i^{-1}\big)\leqslant\sum_{i=1}^{n}w_i\big[(1-t)X^{-1}+tA_i^{-1}\big]=(1-t)X^{-1}+t\sum_{i=1}^{n}w_iA_i^{-1},$$

which implies that $X\geqslant(\sum_{i=1}^{n}w_iA_i^{-1})^{-1}$.

The case $t\in[-1,0)$ holds by duality.

(11) Let $t\in(0,1]$ and let $X=P_t(\omega;\mathbb{A})$. Then

$$X=\sum_{i=1}^{n}w_i(X\#_t A_i)=\frac{1}{k}\Bigg(\underbrace{\sum_{i=1}^{n}w_i(X\#_t A_i)+\cdots+\sum_{i=1}^{n}w_i(X\#_t A_i)}_{k}\Bigg)$$

and therefore $X=P_t(\omega^{(k)};\mathbb{A}^{(k)})$. The case $t\in[-1,0)$ is similar.

(12) Let $t\in(0,1]$. Then $P_t(\omega;A_1,\dots,A_{n-1},X)=X$ if and only if $X=\sum_{i=1}^{n-1}w_i(X\#_t A_i)+w_nX$ if and only if $X=\frac{1}{1-w_n}\sum_{i=1}^{n-1}w_i(X\#_t A_i)$ if and only if $X=P_t(\hat{\omega};A_1,\dots,A_{n-1})$. By duality, (12) holds for $t\in[-1,0)$. If $\omega=\frac{1}{n+1}(1,1,\dots,1)\in\Delta_{n+1}$, then $\hat{\omega}=\frac{1}{n}(1,\dots,1)\in\Delta_n$, and thus, $P_t(A_1,\dots,A_n,X)=X$ if and only if $X=P_t(A_1,\dots,A_n)$.

(13) Note that $X\#_t(X\#_s A_i)=X\#_{st}A_i$. Let $s\in(0,1]$. Suppose that $t\in(0,1]$. Then $X=P_t(\omega;X\#_s A_1,\dots,X\#_s A_n)$ if and only if $X=\sum_{i=1}^{n}w_i(X\#_{st}A_i)$ if and only if $X=P_{st}(\omega;\mathbb{A})$. If $t\in[-1,0)$, then $X=P_t(\omega;X\#_s A_1,\dots,X\#_s A_n)$ if and only if $X^{-1}=P_{-t}(\omega;X^{-1}\#_s A_1^{-1},\dots,X^{-1}\#_s A_n^{-1})$ if and only if $X^{-1}=P_{-st}(\omega;\mathbb{A}^{-1})$ if and only if $X=P_{st}(\omega;\mathbb{A})$, since $st\in(0,1]$.

(14) Note that $\Phi(A\#_t B)\leqslant\Phi(A)\#_t\Phi(B)$ for any $A,B>0$ and $t\in[0,1]$ (cf. Theorem 4.1.5 of [10]). Let $t\in(0,1]$ and $X_t=P_t(\omega;\mathbb{A})$. Then

$$\Phi(X_t)=\sum_{i=1}^{n}w_i\Phi(X_t\#_t A_i)\leqslant\sum_{i=1}^{n}w_i\big(\Phi(X_t)\#_t\Phi(A_i)\big).\tag{3.7}$$

Define $f(X) = \sum_{i=1}^{n} w_i(X \#_t \Phi(A_i))$. Then $\lim_{k\to\infty} f^k(X) = P_t(\omega; \Phi(\mathbb{A}))$ for any $X > 0$. By (3.7), $f(\Phi(X_t)) \geqslant \Phi(X_t)$. Since f is monotonic, $f^k(\Phi(X_t)) \geqslant \Phi(X_t)$ for all $k \in \mathbb{N}$. Thus, $P_t(\omega; \Phi(\mathbb{A})) = \lim_{k\to\infty} f^k(\Phi(X_t)) \geqslant \Phi(X_t) = \Phi(P_t(\omega; \mathbb{A}))$.

Let $t \in [-1, 0)$ and let Φ be a strictly positive unital linear map. By Choi's inequality (Theorem 2.3.6 of [10]), $\Phi(A)^{-1} \leqslant \Phi(A^{-1})$ for all $A > 0$. By (4) and the preceding paragraph, $\Phi(P_{-t}(\omega; \mathbb{A}^{-1})) \leqslant P_{-t}(\omega; \Phi(\mathbb{A}^{-1})) \leqslant P_{-t}(\omega; \Phi(\mathbb{A})^{-1})$. This implies that $\Phi(P_t(\omega; \mathbb{A})) = \Phi(P_{-t}(\omega; \mathbb{A}^{-1})^{-1}) \geqslant \Phi(P_{-t}(\omega; \mathbb{A}^{-1}))^{-1} \geqslant P_{-t}(\omega; \Phi(\mathbb{A})^{-1})^{-1} = P_t(\omega; \Phi(\mathbb{A}))$.

(15) Let $t \in (0, 1]$ and $X = P_t(\omega; \mathbb{A})$. Then

$$|||X||| \leqslant \sum_{i=1}^{n} w_i |||X \#_t A_i||| \leqslant \sum_{i=1}^{n} w_i |||X|||^{1-t} |||A_i|||^t = |||X|||^{1-t} \sum_{i=1}^{n} w_i |||A_i|||^t,$$

where the second inequality follows from Theorem 2.10 of [27] and Corollary IX.5.3 [8], and hence $|||P_t(\omega; \mathbb{A})||| \leqslant (\sum_{i=1}^{n} w_i |||A_i|||^t)^{\frac{1}{t}}$. Since $|||A^{-1}||| \geqslant |||A|||^{-1}$ for any $A > 0$, $|||P_{-t}(\omega; \mathbb{A})||| = |||P_t(\omega; \mathbb{A}^{-1})^{-1}||| \geqslant |||P_t(\omega; \mathbb{A}^{-1})|||^{-1} \geqslant [\sum_{i=1}^{n} w_i |||A_i^{-1}|||^t]^{-\frac{1}{t}}$. □

Remark 3.6. From the (AGH) inequalities (Proposition 3.5(10)) we can obtain other inequalities from operator monotone functions on the positive reals. Let $f : (0, \infty) \to (0, \infty)$ be an operator monotone increasing (resp. decreasing) function. Then

$$P_t\big(\omega; f(\mathbb{A})\big) \leqslant \sum_{i=1}^{n} w_i f(A_i) \leqslant f\left(\sum_{i=1}^{n} w_i A_i\right),$$

$$P_t\big(\omega; f(\mathbb{A})\big) \geqslant \left[\sum_{i=1}^{n} w_i f(A_i)^{-1}\right]^{-1} \geqslant f\left(\sum_{i=1}^{n} w_i A_i\right),$$

respectively; these follow from the equivalence between operator monotonicity and operator log-concavity by Ando and Hiai [1].

Property (12) implies in particular that $P_t(\hat{\omega}; A_1, \ldots, A_{n-1})$ is the unique fixed point of the map $f(X) = P_t(\omega; A_1, \ldots, A_{n-1}, X)$. By Proposition 3.5(5), f is a non-expansive map for the Thompson metric.

Corollary 3.7. *Let $t \in [-1, 1] \setminus \{0\}$, $\omega \in \Delta_n$ and let $A_1, \ldots, A_{n-1} \in \mathbb{P}$. Then there exists $X_0 \in \mathbb{P}$ such that $\lim_{k\to\infty} f^k(X_0) = P_t(\hat{\omega}; A_1, \ldots, A_{n-1})$, where $f : \mathbb{P} \to \mathbb{P}$ is defined by $f(X) = P_t(\omega; A_1, \ldots, A_{n-1}, X)$. Furthermore, for $B \in \mathbb{P}$,*

$$P_t(\omega : A_1, \ldots, A_{n-1}, B) \leqslant B \quad \textit{implies} \quad P_t(\hat{\omega}; A_1, \ldots, A_{n-1}) \leqslant B.$$

In particular, $P_t(A_1, \ldots, A_{n-1}, B) \leqslant B$ implies that $P_t(A_1, \ldots, A_{n-1}) \leqslant B$.

Proof. Let $f(X) = P_t(\omega; A_1, \ldots, A_{n-1}, X)$. By Proposition 3.5(4), f is monotonic. Let $B \in \mathbb{P}$. Pick $\alpha, \beta > 0$ such that $B, A_i \in [\beta I, \alpha I] = \{X \in \mathbb{P}\colon \beta I \leqslant X \leqslant \alpha I\}$ for all $i = 1, \ldots, n-1$. Then by Proposition 3.5(10), f maps $[\beta I, \alpha I]$ into itself. Indeed,

$$\beta I = P_t(\omega; \beta I, \ldots, \beta I) \leqslant P_t(\omega; A_1, \ldots, A_{n-1}, X) \leqslant P_t(\omega; \alpha I, \ldots, \alpha I) = \alpha I$$

for any $X \in [\beta I, \alpha I]$. So, $f^k(X_0) \in [\beta I, \alpha I]$ for all $k \in \mathbb{N}$ and for any $X_0 \in [\beta I, \alpha I]$. Let $X_0 \in [\beta I, \alpha I]$ such that $f(X_0) \leqslant X_0$. Then by induction $f^{k+1}(X_0) \leqslant f^k(X_0)$ for all $k \in \mathbb{N}$. That is, $\{f^k(X_0)\}_{k=1}^{\infty}$ is a decreasing sequence bounded below by βI and thus converges to $P_t(\hat{\omega}; A_1, \ldots, A_{n-1})$, which is the unique fixed point of f. In particular for $X_0 = \alpha I$, we have that $f(\alpha I) \leqslant \alpha I$ and $\lim_{k\to\infty} f^k(\alpha I) = P_t(\hat{\omega}; A_1, \ldots, A_{n-1})$.

Suppose that $f(B) \leqslant B$. Then $f^{k+1}(B) \leqslant f^k(B) \leqslant B$ for all $k \in \mathbb{N}$ and hence $P_t(\hat{\omega}; A_1, \ldots, A_{n-1}) = \lim_{k\to\infty} f^k(B) \leqslant f(B) \leqslant B$. □

The problem of finding an explicit form of $P_t(\omega; \mathbb{A})$ is non-trivial, except for $n = 2$.

Proposition 3.8. *For $t \in (0, 1]$, we have*

$$P_t(w_1, w_2; A, B) = A \#_{\frac{1}{t}} \big[w_1 A + w_2(A \#_t B)\big] = A^{1/2}\big(w_1 I + w_2\big(A^{-1/2} B A^{-1/2}\big)^t\big)^{\frac{1}{t}} A^{1/2}.$$

In particular, $P_{\frac{1}{n}}(w_1, w_2; A, B) = \sum_{k=0}^{n} \binom{n}{k} w_1^k w_2^{n-k} (B \#_{\frac{k}{n}} A)$.

Proof. Let $X = P_t(w_1, w_2; A, B)$. Then by definition, $X = w_1(X \#_t A) + w_2(X \#_t B)$. Setting $U = A^{-1/2} X A^{-1/2}$ and $Z = A^{-1/2} B A^{-1/2}$ yields $U = w_1 U^{1-t} + w_2(U \#_t Z)$, which is equivalent to $I = w_1 U^{-t} + w_2(U^{-1/2} Z U^{-1/2})^t$, that is, $Z = (\frac{U^t - w_1 I}{w_2})^{\frac{1}{t}}$. This implies that $U = [w_1 I + w_2 Z^t]^{\frac{1}{t}}$ and

$$\begin{aligned} X = A^{1/2} U A^{1/2} &= A^{1/2}\big[w_1 I + w_2 Z^t\big]^{\frac{1}{t}} A^{1/2} = A^{1/2}\big[w_1 I + w_2\big(A^{-1/2} B A^{-1/2}\big)^t\big]^{\frac{1}{t}} A^{1/2} \\ &= A^{1/2}\big(I \#_{1/t} \big[w_1 I + w_2\big(A^{-1/2} B A^{-1/2}\big)^t\big]\big) A^{1/2} = A \#_{1/t} \big[w_1 A + w_2(A \#_t B)\big]. \end{aligned}$$

If $t = \frac{1}{n}$ for some $n \in \mathbb{N}$, then

$$A^{-1/2} X A^{-1/2} = U = \big[w_1 I + w_2 Z^{\frac{1}{n}}\big]^n = \sum_{i=1}^{n} \binom{n}{k} w_1^k w_2^{n-k} Z^{\frac{n-k}{n}}$$

and hence

$$\begin{aligned} X &= A^{1/2}\left(\sum_{i=1}^{n} \binom{n}{k} w_1^k w_2^{n-k} Z^{\frac{n-k}{n}}\right) A^{1/2} = \sum_{i=1}^{n} \binom{n}{k} w_1^k w_2^{n-k} \big(A^{1/2} Z^{\frac{n-k}{n}} A^{1/2}\big) \\ &= \sum_{i=1}^{n} \binom{n}{k} w_1^k w_2^{n-k} (A \#_{\frac{n-k}{n}} B) = \sum_{i=1}^{n} \binom{n}{k} w_1^k w_2^{n-k} (B \#_{\frac{k}{n}} A). \qquad \square \end{aligned}$$

Remark 3.9. We observe that for any $(w_1, w_2) \in \Delta_2$,

$$\lim_{t\to 0} P_t(w_1, w_2; A, B) = A \#_{w_2} B.$$

Indeed, setting $Z = A^{-1/2}BA^{-1/2}$, we have for $t > 0$

$$A^{-1/2}\big(A \#_{\frac{1}{t}} \big[w_1 A + w_2(A \#_t B)\big]\big)A^{-1/2} = \big[w_1 I + w_2 Z^t\big]^{\frac{1}{t}} \to Z^{w_2}. \tag{3.8}$$

That is, $P_t(w_1, w_2; A, B) \to A^{1/2}Z^{w_2}A^{1/2} = A^{1/2}(A^{-1/2}BA^{-1/2})^{w_2}A^{1/2} = A \#_{w_2} B$. This further implies that

$$\begin{aligned}\lim_{t\to 0^-} P_t(w_1, w_2; A, B) &= \lim_{t\to 0^-} P_{-t}\big(w_1, w_2; A^{-1}, B^{-1}\big)^{-1} \\ &= \big(A^{-1} \#_{w_2} B^{-1}\big)^{-1} = A \#_{w_2} B.\end{aligned}$$

Remark 3.10. We note that the power mean $P_t(A, B)$ coincides with the operator mean arising from $f_t : (0, \infty) \to (0, \infty)$ defined by $f_t(x) = (\frac{x^t+1}{2})^{\frac{1}{t}}$. It is called the quasi-arithmetic (power) mean of order t. Its operator monotonicity (cf. Proposition 3.5), infinite divisibility, and the complete positivity of an associated linear operator have been studied by Bhatia and Kosaki [13] and Besenyei and Petz [7]. It turns out [31] that the power mean $P_t(A, B)$ arises as the midpoint operation of a manifold equipped with an affine connection. One can also see that $P_t(w_1, w_2; A, B)(= A \#_{\frac{1}{t}} [w_1 A + w_2(A \#_t B)]) \neq (w_1 A^t + w_2 B^t)^{\frac{1}{t}}$ for non-commuting A and B. In fact, $\lim_{t\to 0}(\frac{A_1^t+\cdots+A_n^t}{n})^{\frac{1}{t}} = \exp(\frac{\log A_1+\cdots+\log A_n}{n})$ and is known as the Log-Euclidean mean [3]. We note that the Log-Euclidean mean is far from the geometric mean $A \# B$ for $n = 2$.

4. The Karcher mean via power means

Let $\mathbb{A} = (A_1, \ldots, A_n) \in \mathbb{P}^n$, $\omega = (w_1, \ldots, w_n) \in \Delta_n$. The ω-weighted Karcher mean $\Lambda(\omega; \mathbb{A})$ of $\mathbb{A}$ is defined to be the unique positive definite solution of the equation

$$\sum_{i=1}^{n} w_i \log\big(X^{-1/2}A_i X^{-1/2}\big) = 0. \tag{4.9}$$

We note from (4.9) that $\Lambda(\omega; \mathbb{A}^{-1})^{-1} = \Lambda(\omega; \mathbb{A})$, the self-duality of the Karcher mean.

Lemma 4.1. *Let $D \subset \mathbb{R}$ be an open interval and let $\epsilon_0 > 0$. Let $F : D \times (-\epsilon_0, \epsilon_0) \to \mathbb{R}$ be a map satisfying*

(i) *F is an increasing function in the first variable,*
(ii) *there exists a continuous and increasing function $f : D \to \mathbb{R}$ such that for all $a \in D$, $f(a) = \frac{\partial}{\partial t}F(a, t)|_{t=0}$.*

Then for any $a \in D$ and any sequence a_n of D converging to a,

$$\lim_{n\to\infty} n\left(F\left(a_n, \frac{1}{n}\right) - F(a, 0)\right) = f(a).$$

Proof. Let $a \in D$ and let a_n be a sequence converging to a. Let $\epsilon > 0$. Since f is a continuous and increasing function, there exists $\delta > 0$ such that

$$f(a) - \epsilon < f(a - \delta) \leqslant f(a) \leqslant f(a + \delta) < f(a) + \epsilon. \tag{4.10}$$

Since $a_n \to a$, there exists $N_1 > 0$ such that $a - \delta < a_n < a + \delta$ for all $n \geqslant N_1$. Since F is an increasing function in the first variable, $F(a - \delta, 1/n) \leqslant F(a_n, 1/n) \leqslant F(a + \delta, 1/n)$ for all $n \geqslant N_1$. That is, for all $n \geqslant N_1$,

$$n\big(F(a - \delta, 1/n) - F(a, 0)\big) \leqslant n\big(F(a_n, 1/n) - F(a, 0)\big) \leqslant n\big(F(a + \delta, 1/n) - F(a, 0)\big).$$

By (4.10), $f(a) - \epsilon < f(a - \delta)$ and $f(a + \delta) < f(a) + \epsilon$, and by (ii),

$$\lim_{n \to \infty} n\big(F(a \pm \delta, 1/n) - F(a, 0)\big) = f(a \pm \delta)$$

one can find $N_2 > N_1$ such that for all $n \geqslant N_2$, $f(a) - \epsilon < n(F(a \pm \delta, 1/n) - F(a, 0))$ and $f(a) + \epsilon > n(F(a + \delta, 1/n) - F(a, 0))$. This completes the proof. □

Since the map $F(x, t) = x^t$ on $(0, \infty) \times \mathbb{R}$ satisfies the conditions in Lemma 4.1, we have the following result.

Lemma 4.2. *Let $x_0 > 0$ and let x_n be a sequence of positive real numbers converging to x_0. Then* $\lim_{n \to \infty} n(x_n^{1/n} - 1) = \log x_0$.

The main result of this paper is the following.

Theorem 4.3. *We have*

$$\lim_{t \to 0} P_t(\omega; \mathbb{A}) = \Lambda(\omega; \mathbb{A}).$$

Proof. By Proposition 3.5(8) and the self-duality of the Karcher mean, it suffices to show that $\lim_{t \to 0^+} P_t(\omega; \mathbb{A}) = \Lambda(\omega; \mathbb{A})$. Set $X_t = P_t(\omega; \mathbb{A})$. Since $X_t = \sum_{i=1}^n w_i (X_t \#_t A_i)$,

$$I = \sum_{i=1}^n w_i \big(X_t^{-1/2} (X_t \#_t A_i) X_t^{-1/2}\big) = \sum_{i=1}^n w_i \big(X_t^{-1/2} A_i X_t^{-1/2}\big)^t.$$

In particular for all $t \in (0, 1]$,

$$0 = \sum_{i=1}^n w_i \left[\frac{(X_t^{-1/2} A_i X_t^{-1/2})^t - I}{t} \right]. \tag{4.11}$$

Let $\{t_k\}_{k=1}^\infty$ be a sequence in $(0, 1]$ converging to 0. Since X_t lies in the order interval determined by the ω-weighted harmonic mean and arithmetic mean, which is compact, the sequence $\{X_{t_k}\}$ has at least one limit point. Suppose that X_0 is a limit point of $\{X_{t_k}\}$. We will show that $X_0 = \Lambda(\omega; \mathbb{A})$. Passing to a subsequence, we may assume that $X_{t_k} \to X_0$, as

$k\to\infty$. Then $X_{t_k}^{-1/2}A_iX_{t_k}^{-1/2}\to X_0^{-1/2}A_iX_0^{-1/2}$ for all i. Setting $Y_{t_k}=X_{t_k}^{-1/2}A_iX_{t_k}^{-1/2}$ and $Y_0=X_0^{-1/2}A_iX_0^{-1/2}$ yields $Y_{t_k}\to Y_0$. Let U_{t_k} be a unitary matrix such that $U_{t_k}Y_{t_k}U_{t_k}^*:=D_{t_k}$ is a diagonal matrix. Since Y_{t_k} converges to Y_0 and the unitary group is compact, we may assume that $U_{t_k}\to U_0$ and $D_{t_k}\to D_0$ for some unitary matrix U_0 and a diagonal matrix D_0. Indeed, first consider a subsequence of U_{t_k} converging to a unitary matrix U_0, second the corresponding subsequence of Y_{t_k}, which always converges to Y_0, and then finally consider the corresponding subsequence of D_{t_k}, which converges to $D_0:=U_0Y_0U_0^*$. By Lemma 2.3, $\lim_{k\to\infty}\frac{D_{t_k}^{t_k}-I}{t_k}=\log D_0$. This implies that

$$\begin{aligned}\lim_{k\to\infty}\left[\frac{(X_{t_k}^{-1/2}A_iX_{t_k}^{-1/2})^{t_k}-I}{t_k}\right]&=\lim_{k\to\infty}\left[\frac{(Y_{t_k})^{t_k}-I}{t_k}\right]=\lim_{k\to\infty}\left[\frac{U_{t_k}^*(D_{t_k})^{t_k}U_{t_k}-I_m}{t_k}\right]\\&=\lim_{k\to\infty}U_{t_k}^*\left[\frac{D_{t_k}^{t_k}-I}{t_k}\right]U_{t_k}=\log\big(U_0^*D_0U_0\big)\\&=\log Y_0=\log\big(X_0^{-1/2}A_iX_0^{-1/2}\big).\end{aligned}$$

This together with (4.11) yields $0=\sum_{i=1}^n w_i\log(X_0^{-1/2}A_iX_0^{-1/2})$. That is, $X_0=\Lambda(\omega;\mathbb{A})$. □

From Theorem 3.1 and Theorem 4.3 we obtain

Corollary 4.4. *With $P_0(\omega;\mathbb{A})=\Lambda(\omega;\mathbb{A})$, the map $t\mapsto P_t(\omega;\mathbb{A})$ is continuous on* $[-1,1]$.

The basic properties of power means in Proposition 3.5 together with Theorem 4.3 provide simple proofs of some important properties of the Karcher mean.

Corollary 4.5. *(Cf. [24,12].) The Karcher mean satisfies the following properties*:

(P1) *(Consistency with scalars) $\Lambda(\omega;\mathbb{A})=A_1^{w_1}\cdots A_n^{w_n}$ if the A_i's commute*;
(P2) *(Joint homogeneity) $\Lambda(\omega;a_1A_1,\dots,a_nA_n)=a_1^{w_1}\cdots a_n^{w_n}\Lambda(\omega;\mathbb{A})$*;
(P3) *(Permutation invariance) $\Lambda(\omega_\sigma;\mathbb{A}_\sigma)=\Lambda(\omega;\mathbb{A})$, where $\omega_\sigma=(w_{\sigma(1)},\dots,w_{\sigma(n)})$*;
(P4) *(Monotonicity) If $B_i\leqslant A_i$ for all $1\leqslant i\leqslant n$, then $\Lambda(\omega;\mathbb{B})\leqslant\Lambda_n(\omega;\mathbb{A})$*;
(P5) $d_\infty(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B}))\leqslant\max_{1\leqslant i\leqslant n}\{d_\infty(A_i,B_i)\}$;
(P6) *(Invariancy) $\Lambda(\omega;M^*\mathbb{A}M)=M^*\Lambda(\omega;\mathbb{A})M$ for any invertible M*;
(P7) *(Joint concavity) $\Lambda(\omega;(1-u)\mathbb{A}+u\mathbb{B})\geqslant(1-u)\Lambda(\omega;\mathbb{A})+u\Lambda(\omega;\mathbb{B})$ for $0\leqslant u\leqslant 1$*;
(P8) *(Self-duality) $\Lambda(\omega;A_1^{-1},\dots,A_n^{-1})^{-1}=\Lambda(\omega;A_1,\dots,A_n)$*;
(P9) *(Determinant identity)* $\mathrm{Det}\Lambda(\omega;\mathbb{A})=\prod_{i=1}^n(\mathrm{Det}A_i)^{w_i}$; *and*
(P10) *(AGH weighted mean inequalities) $(\sum_{i=1}^n w_iA_i^{-1})^{-1}\leqslant\Lambda(\omega;\mathbb{A})\leqslant\sum_{i=1}^n w_iA_i$*;
(P11) *$\Lambda(\omega^{(k)};\mathbb{A}^{(k)})=\Lambda(\omega;\mathbb{A})$ for any $k\in\mathbb{N}$*;
(P12) *$\Lambda(\omega;A_1,\dots,A_{n-1},X)=X$ if and only if $X=\Lambda(\hat\omega;A_1,\dots,A_{n-1})$. In particular, $\Lambda(A_1,\dots,A_n,X)=X$ if and only if $X=\Lambda(A_1,\dots,A_n)$*;
(P13) *for any $t\in(0,1]$, $X=\Lambda(\omega;X\#_tA_1,\dots,X\#_tA_n)$ if and only if $X=\Lambda(\omega;\mathbb{A})$*;
(P14) *$\Phi(\Lambda(\omega;\mathbb{A}))\leqslant\Lambda(\omega;\Phi(\mathbb{A}))$ for any positive unital linear map Φ. If Φ is strictly positive, then $\Phi(\Lambda(\omega;\mathbb{A}))=\Lambda(\omega;\Phi(\mathbb{A}))$*;
(P15) *$\prod_{i=1}^n|||A_i^{-1}|||^{-w_i}\leqslant|||\Lambda(\omega;\mathbb{A})|||\leqslant\prod_{i=1}^n|||A_i|||^{w_i}$ for any unitarily invariant norm $|||\cdot|||$.*

Proof. By Proposition 3.5, Theorem 4.3 and by the Karcher equation, (P1)–(P13) are immediate. For instance, since each $P_t(\omega;\cdot)$ is monotonic, its limit $\Lambda(\omega;\cdot)$ also is.

(P14) By Proposition 3.5(13), $\Phi(P_t(\omega;\mathbb{A})) \leqslant P_t(\omega;\Phi(\mathbb{A}))$ for all $t \in (0,1]$. As $t \to 0$, we have $\Phi(\Lambda(\omega;\mathbb{A})) \leqslant \Lambda(\omega;\Phi(\mathbb{A}))$. If Φ is strictly positive, then $\Phi(P_t(\omega;\mathbb{A})) \geqslant P_t(\omega;\Phi(\mathbb{A}))$ for all $t \in [-1,1)$ by Proposition 3.5(13). Then $\Phi(\Lambda(\omega;\mathbb{A})) \geqslant \Lambda(\omega;\Phi(\mathbb{A}))$.

(P15) By Proposition 3.5(11), $|||P_t(\omega;\mathbb{A})||| \leqslant (\sum_{i=1}^n w_i |||A_i|||^t)^{\frac{1}{t}}$ for all $t \in (0,1]$. As $t \to 0$, we have $|||\Lambda(\omega;\mathbb{A})||| \leqslant \lim_{t\to 0}(\sum_{i=1}^n w_i |||A_i|||^t)^{\frac{1}{t}} = \prod_{i=1}^n |||A_i|||^{w_i}$, where the equality follows from the fact that weighted power means of positive real numbers converge to the weighted geometric mean. The other inequality follows similarly. □

Ando, Li and Mathias [2] listed the ten properties (P1)–(P10) for the unweighted case $\omega = (1/n,\ldots,1/n)$ as properties that a geometric mean of n positive definite matrices should satisfy, and their mean, called the ALM geometric mean, possesses all of them. The BMP geometric mean of Bini, Meini and Poloni [15] is also a matrix geometric mean in this axiomatic sense. In fact, there are infinitely many matrix geometric means: fixed point means of the ALM and BMP geometric means [26] and their weighted geometric means $\mathrm{ALM}(A_1,\ldots,A_n)\,\#_t\,\mathrm{BMP}(A_1,\ldots,A_n)$, $t \in [0,1]$. The properties (P11)–(P13) are special for the Karcher mean. Some parts of the properties (P14) and (P15) have been established by Bhatia and Karandikar [12]. For the weighted case, there are also infinitely many weighted geometric means of n positive definite matrices: the weighted Karcher mean, the weighted BMP geometric mean [25] and their weighted geometric means. We note that there has been no successful weighted extension of the ALM geometric mean.

Next we investigate some other properties of the power mean that hold for the Karcher mean.

Corollary 4.6. *If $\mathbb{A} \leqslant \mathbb{B}$ and $A_i < B_i$ for some i, then $\Lambda(\omega;\mathbb{A}) < \Lambda(\omega;\mathbb{B})$ and $P_t(\omega;\mathbb{A}) < P_t(\omega;\mathbb{B})$ for any sufficiently small t.*

Proof. Find $0 < \alpha < 1$ such that $A_i \leqslant \alpha B_i$. Then

$$\Lambda(\omega;A_1,\ldots,A_i,\ldots,A_n) \leqslant \Lambda(\omega;B_1,\ldots,\alpha B_i,\ldots,B_n) = \alpha^{w_i}\Lambda(\omega;B_1,\ldots,B_i,\ldots,B_n) < \Lambda(\omega;B_1,\ldots,B_i,\ldots,B_n)$$

where we used the joint homogeneity and monotonicity of the Karcher mean. Finding $0 < \beta < 1$ such that $\Lambda(\omega;\mathbb{A}) < \beta\Lambda(\omega;\mathbb{B}) < \Lambda(\omega;\mathbb{B})$, we have from Theorem 4.3 that

$$\lim_{t\to 0} P_t(\omega;\mathbb{A}) = \Lambda(\omega;\mathbb{A}) < \beta\Lambda(\omega;\mathbb{B}) < \Lambda(\omega;\mathbb{B}) = \lim_{t\to 0} P_t(\omega;\mathbb{B})$$

which implies that $P_t(\omega;\mathbb{A}) < P_t(\omega;\mathbb{B})$ for any sufficiently small t. □

The continuity, indeed Lipschitz continuity, of the Karcher mean (P5) follows from $d_\infty(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B})) \leqslant \max\{d_\infty(A_i,B_i)\}$, which in turn follows from Proposition 3.5(5) and Theorem 4.3. A stronger result is the non-expansiveness of the Karcher mean;

$$\delta\big(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B})\big)\leqslant\sum_{i=1}^{n}w_i\delta(A_i,B_i). \tag{4.12}$$

This nice inequality has been proved by Lawson and Lim [24] and Bhatia and Karandikar [12].

Corollary 4.7. *If $\delta(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B}))\neq\sum_{i=1}^{n}w_i\delta(A_i,B_i)$, then for any sufficiently small t, $\delta(P_t(\omega;\mathbb{A}),P_t(\omega;\mathbb{B}))<\sum_{i=1}^{n}w_i\delta(A_i,B_i)$.*

Proof. Note that

$$\lim_{t\to 0}\delta\big(P_t(\omega;\mathbb{A}),P_t(\omega;\mathbb{B})\big)=\delta\big(\Lambda(\omega;\mathbb{A}),\Lambda(\omega;\mathbb{B})\big)<\sum_{i=1}^{n}w_i\delta(A_i,B_i). \qquad \square$$

Property (P12) of the Karcher mean implies that $\Lambda(\hat{\omega};A_1,\ldots,A_{n-1})$ is the unique fixed point of the map $f(X)=\Lambda(\omega;A_1,\ldots,A_{n-1},X)$. By (4.12), f is a strict contraction on $\mathbb{P}$ with respect to the Riemannian metric. The following (Löwer) order behavior around the fixed point of f is special for the Karcher mean.

Corollary 4.8. *We have*

$$\Lambda(\omega;A_1,\ldots,A_{n-1},A_n)<A_n \quad \textit{implies} \quad \Lambda(\hat{\omega};A_1,\ldots,A_{n-1})<\Lambda(\omega;A_1,\ldots,A_{n-1},A_n),$$

and if $\sum_{i=1}^{n}w_i\log A_i<0$, then $\Lambda(\omega;A_1,\ldots,A_n)<I$ and $P_t(\omega;\mathbb{A})<I$ for any sufficiently small t.

Proof. Suppose that $B:=\Lambda(\omega;A_1,\ldots,A_{n-1},A_n)<A_n$. Define $f:\mathbb{P}\to\mathbb{P}$ by $f(X)=\Lambda(\omega;A_1,\ldots,A_{n-1},X)$. Then by (4.12), it is a strict contraction for δ and hence has a unique positive definite solution. By (P12) and Banach fixed point theorem, $\lim_{k\to\infty}f^k(B)=\Lambda(\hat{\omega};A_1,\ldots,A_{n-1})$. It follows from $f(A_n)=B<A_n$ and strict monotonicity of the Karcher mean (Corollary 4.6) that $f(B)=\Lambda(\omega;A_1,\ldots,A_{n-1},B)<\Lambda(\omega;A_1,\ldots,A_{n-1},A_n)=f(A_n)=B$. By induction, $f^k(B)<\cdots<f(B)<f(A_n)=B$ and therefore $\Lambda(\hat{\omega};A_1,\ldots,A_{n-1})=\lim_{k\to\infty}f^k(B)<B=\Lambda(\omega;A_1,\ldots,A_{n-1},A_n)$.

Next, suppose that $\sum_{i=1}^{n}w_i\log A_i<0$. Set $Y=-\sum_{i=1}^{n}w_i\log A_i$. Then $Y>0$ and hence $B:=\exp(Y)>I$. From

$$0=\sum_{i=1}^{n}w_i\log A_i+Y=\sum_{i=1}^{n}w_i\log A_i+\log B=\frac{1}{2}\left(\sum_{i=1}^{n}w_i\log A_i+\log B\right)$$

and $\omega_1:=\frac{1}{2}(w_1,\ldots,w_n,1)\in\Delta_{n+1}$, we have $\Lambda(\omega_1;A_1,\ldots,A_n,B)=I<B$. By strict monotonicity of the Karcher mean, $\Lambda(\omega_1;A_1,\ldots,A_n,I)<\Lambda(\omega_1;A_1,\ldots,A_n,B)=I$. From the first paragraph and the fact that $\hat{\omega}_1=\omega$, we have $\Lambda(\omega;A_1,\ldots,A_n)\leqslant\Lambda(\omega_1;A_1,\ldots,A_n,I)<I$. Finally $\lim_{t\to 0}P_t(\omega;A_1,\ldots,A_n)=\Lambda(\omega;A_1,\ldots,A_n)<I$ implies that for sufficiently small t, $P_t(\omega;A_1,\ldots,A_n)<I$. $\square$

One can obtain in a way similar to the preceding that $\Lambda(\omega; A_1, \ldots, A_{n-1}, A_n) \leqslant A_n$ implies $\Lambda(\hat{\omega}; A_1, \ldots, A_{n-1}) \leqslant \Lambda(\omega; A_1, \ldots, A_{n-1}, A_n)$ and

$$(Y) \quad \sum_{i=1}^{n} w_i \log A_i \leqslant 0 \quad \text{implies} \quad \Lambda(\omega; A_1, \ldots, A_n) \leqslant I.$$

The property (Y), which was established by Yamazaki [33], is one of characteristic properties of the Karcher mean by the following result.

Theorem 4.9. *The Karcher mean is uniquely determined by congruence invariancy* (P6)*, self-duality* (P8)*, and* (Y).

Proof. Let $g : \Delta_n \times \mathbb{P}^n \to \mathbb{P}$ be a map satisfying (P6), (P8) and (Y). By (P8) and (Y), $\sum_{i=1}^n w_i \log A_i = 0$ implies that $g(\omega; A_1, \ldots, A_n) = I$. Let $X = \Lambda(\omega; \mathbb{A})$. Then $\sum_{i=1}^n w_i \log(X^{-1/2} A_i X^{-1/2}) = 0$ and hence $g(\omega; X^{-1/2} A_1 X^{-1/2}, \ldots, X^{-1/2} A_n X^{-1/2}) = I$. By (P6), $g(A_1, \ldots, A_n) = X = \Lambda_n(A_1, \ldots, A_n)$. $\square$

Our method of deriving the monotonicity of the Karcher mean is free from any probabilistic and Riemannian geometric techniques because we have just started from the Karcher equation (4.9). The Karcher equation can be defined on the convex cone of positive definite operators on an infinite dimensional Hilbert space. But the existence and uniqueness of a positive definite solution have not previously been investigated in any depth. The weighted power means exist since the Thompson metric exists on the cone of positive definite operators, on which Lemma 2.2 and Lemma 2.3 are still valid [21,22]. So if one can show the monotonicity of the power mean function $t \to P_t(\omega; \mathbb{A})$, then the strong limit of the sequence $X_k = P_{\frac{1}{k}}(\omega; \mathbb{A})$ exists and is probably a solution of the Karcher equation. Note that the power mean $P_t(\omega; \mathbb{A})$ is contained in the order interval determined by the weighted harmonic and arithmetic means.

By a numerical simulation, the following result seems to be true. If $0 \leqslant t \leqslant s \leqslant 1$, then $P_t(\omega; \mathbb{A}) \leqslant P_s(\omega; \mathbb{A})$ for all $\omega \in \Delta_n$ and $\mathbb{A} \in \mathbb{P}^n$. By (3.8), it is true for $n = 2$.

Changing the weighted arithmetic mean operation in the defining Eq. (3.4) of the power mean $P_t(\omega; \mathbb{A})$ into any weighted geometric mean $G(\omega; \mathbb{A})$ of n positive definite matrices, which is non-expansive for the Riemannian metric or the Thompson metric, yields other matrix geometric means via the geometric mean equation

$$X = G(\omega; X \#_t A_1, \ldots, X \#_t A_n), \quad t \in (0, 1].$$

For instance, one may take $G = \text{BMP}$ and $G = \text{ALM}$ for the unweighted case (see [25] for their non-expansiveness). One can check by the non-expansive property that a unique positive definite solution exists, denoted by $G_t(\omega; \mathbb{A})$. By the self-duality of G, $G_{-t}(\omega; \mathbb{A}) := G_t(\omega; \mathbb{A}^{-1})^{-1} = G_t(\omega; \mathbb{A})$. Then by using the fixed point approach in Proposition 3.5, one can see that G_t is a weighted matrix mean (satisfies (P1)–(P10)) and is also non-expansive. By (P13), $\Lambda_t = \Lambda$ for all $t \in (0, 1]$. The general convergence of $G_t(\omega; \mathbb{A})$ as $t \to 0$ and the monotonicity of $t \mapsto G_t(\omega; \mathbb{A})$ are non-trivial and suggest interesting future work.

Acknowledgments

The authors thank an anonymous referee for his/her insightful comments and suggestions. This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (No. 2009-0070972).

References

[1] T. Ando, F. Hiai, Operator log-convex functions and operator means, Math. Ann. 350 (2011) 611–630.
[2] T. Ando, C.K. Li, R. Mathias, Geometric means, Linear Algebra Appl. 385 (2004) 305–334.
[3] V. Arsigny, P. Fillard, X. Pennec, N. Ayache, Geometric means in a novel vector space structure on symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 29 (2006) 328–347.
[4] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Riemannian geometry applied to BCI classification, preprint.
[5] F. Barbaresco, Interactions between symmetric cone and information geometries: Bruhat–Tits and Siegel spaces models for higher resolution autoregressive doppler imagery, in: Lecture Notes in Comput. Sci., vol. 5416, 2009, pp. 124–163.
[6] M. Berger, A Panoramic View of Riemannian Geometry, Springer-Verlag, 2003.
[7] A. Besenyei, D. Petz, Completely positive mappings and mean matrices, Linear Algebra Appl. 435 (2011) 984–997.
[8] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997.
[9] R. Bhatia, On the exponential metric increasing property, Linear Algebra Appl. 375 (2003) 211–220.
[10] R. Bhatia, Positive Definite Matrices, Princeton Ser. Appl. Math., Princeton University Press, Princeton, NJ, 2007.
[11] R. Bhatia, J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006) 594–618.
[12] R. Bhatia, R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann., doi:10.1007/s00208-011-0721-9, in press.
[13] R. Bhatia, H. Kosaki, Mean matrices and infinite divisibility, Linear Algebra Appl. 424 (2007) 36–54.
[14] D. Bini, B. Iannazzo, Computing the Karcher mean of symmetric positive definite matrices, preprint.
[15] D. Bini, B. Meini, F. Poloni, An effective matrix geometric mean satisfying the Ando–Li–Mathias properties, Math. Comp. 79 (2010) 437–452.
[16] G. Corach, H. Porta, L. Recht, Convexity of the geodesic distance on spaces of positive operators, Illinois J. Math. 38 (1994) 87–94.
[17] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[18] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. 30 (1977) 509–541.
[19] F. Kubo, T. Ando, Means of positive linear operators, Math. Ann. 246 (1980) 205–224.
[20] S. Lang, Fundamentals of Differential Geometry, Grad. Texts in Math., Springer, 1999.
[21] J. Lawson, Y. Lim, Symmetric spaces with convex metrics, Forum Math. 19 (2007) 571–602.
[22] J. Lawson, Y. Lim, Metric convexity of symmetric cones, Osaka J. Math. 44 (2007) 795–816.
[23] J. Lawson, Y. Lim, A general framework for extending means to higher orders, Colloq. Math. 113 (2008) 191–221.
[24] J. Lawson, Y. Lim, Monotonic properties of the least squares mean, Math. Ann. 351 (2011) 267–279.
[25] H. Lee, Y. Lim, T. Yamazaki, Multi-variable weighted geometric means of positive definite matrices, Linear Algebra Appl. 435 (2011) 307–322.
[26] Y. Lim, On Ando–Li–Mathias geometric mean equations, Linear Algebra Appl. 428 (2008) 1767–1777.
[27] J. Matharu, J. Aujla, Some inequalities for unitarily invariant norms, Linear Algebra Appl., doi:10.1016/j.laa.2010.08.013, in press.
[28] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 26 (2005) 735–747.
[29] M. Moakher, On the averaging of symmetric positive-definite tensors, J. Elasticity 82 (2006) 273–296.
[30] K.-H. Neeb, Compressions of infinite-dimensional bounded symmetric domains, Semigroup Forum 61 (2001) 71–105.
[31] M. Pálfia, Classification of affine matrix means, preprint, 2011.
[32] A.C. Thompson, On certain contraction mappings in a partially ordered vector space, Proc. Amer. Math. Soc. 14 (1963) 438–443.
[33] T. Yamazaki, The Riemannian mean and matrix inequalities related to the Ando–Hiai inequality and Chaotic order, Oper. Matrices, in press.