THE EXPANDING UNIVERSE OF THE GEOMETRIC MEAN

JIMMIE D. LAWSON AND YONGDO LIM

Abstract. In this paper the authors seek to trace in an accessible fashion the rapid recent development of the theory of the matrix geometric mean in the cone of positive definite matrices up through the closely related operator geometric mean in the positive cone of a unital $C^*$-algebra. The story begins with the two-variable matrix geometric mean, moves to the breakthrough developments in the multivariable matrix setting, the main focus of the paper, and then on to the extension to the positive cone of the $C^*$-algebra of operators on a Hilbert space, even to general unital $C^*$-algebras, and finally to the consideration of barycentric maps that grow out of the geometric mean on the space of integrable probability measures on the positive cone. Besides expected tools from linear algebra and operator theory, one observes a surprisingly substantial interplay with geometrical notions in metric spaces, particularly the notion of nonpositive curvature. Added features include a glance at the probabilistic theory of random variables with values in a metric space of nonpositive curvature, and the appearance of related means such as the inductive and power means.

1. Forward

This manuscript is an expanded and more detailed ArXiv version of the authors’ earlier rather similar article Following the Trail of the Operator Geometric Mean [20]. Some twenty years ago the authors published an article [15] in the MONTHLY that treated in some depth the two-variable matrix geometric mean and that has been rather widely read and cited. In this paper we seek to describe the significant advancement and broad generalization of the theory that has taken place in these twenty years.

Date: October 27, 2021.

Key words and phrases. geometric mean, operator mean, positive definite matrices, Hadamard spaces, contractive barycentric map.
2. Introduction

The problem of “squaring a rectangle” is the problem of constructing the side of a square that has the same area as a given rectangle. Such a construction is given by Euclid in Book II of the *Elements*. If the sides of the rectangle are \(a\) and \(b\), then the side of the square has length \(\sqrt{ab}\), the geometric mean of \(a\) and \(b\). Because of their interest in proportions and musical ratios, the Greeks defined some 2500 years ago at least eleven different means, with the best known ones being the arithmetic, geometric, harmonic, and golden. The subject of (binary) means for positive numbers or line segments has a rich mathematical lineage dating back into antiquity. The study of various means on the positive reals and their properties has continued off and on throughout the history of mathematics up to the present day.

The appropriate definition of the geometric mean for two positive definite matrices of the same size seems to have first appeared in 1975 in a paper of Pusz and Woronowicz [27]. Ando [2] provided the first systematic development of many of its basic properties and gave equivalent characterizations and also applications to matrix inequalities that are otherwise difficult to prove.

In an article [15] appearing some twenty years ago in the *Monthly* the authors presented eight characterizations or prominent properties of the classical geometric mean and showed how each extended to the matrix geometric mean setting, providing convincing documentation that the name “matrix geometric mean” was most appropriate. As hinted in the Forward, that theory has now advanced to a multivariable setting for both positive matrices and operators and beyond, as we will trace out.

Positive definite matrices have become fundamental computational objects in many areas of engineering, statistics, quantum information, applied mathematics, and elsewhere. They appear as “data points” in a diverse variety of settings: covariance matrices in statistics, elements of the search space in convex and semidefinite programming, kernels in machine learning, observations in radar imaging, and diffusion tensors in medical imaging, to cite only a few. A variety of computational algorithms have arisen for approximation, interpolation, filtering, estimation, and averaging. Our interest focuses on the last named, the process of finding an average or mean, which is again positive definite. In recent years it has been increasingly recognized that the Euclidean distance is often not the most suitable for the space \(\mathbb{P}\) of positive definite
matrices and that working with the appropriate geometry does matter in computational problems; see e.g. [4], [25]. The matrix geometric mean grows out of the geometric structure of $\mathbb{P}$, which makes it a particularly suitable averaging tool in a variety of settings.

3. Positive Definite Matrices

Let $\mathcal{M}_m(\mathbb{C})$, or simply $\mathcal{M}_m$, denote the set of $m \times m$ complex matrices. We may identify $\mathcal{M}_m$ with the set of linear operators on $\mathbb{C}^m$, where we consider $\mathbb{C}^m$ to be a complex Hilbert space of column vectors with the usual hermitian inner product. Denoting the conjugate transpose of $A \in \mathcal{M}_m$ by $A^*$, we recall that $A$ is hermitian if $A = A^*$ and unitary if $A^* = A^{-1}$. The hermitian matrix $A$ is positive definite if $\forall u \neq 0, \langle u, Au \rangle > 0$. These notions readily generalize to $\mathcal{B}(H)$, the algebra of operators on an arbitrary Hilbert space.

The following are well-known equivalences for a hermitian matrix $A$ to be positive definite (with the definition appearing first):

1. $\langle Ax, x \rangle > 0$ for all $0 \neq x$, where $\langle \cdot, \cdot \rangle$ is the Hilbert space inner product on $\mathbb{C}^m$.
2. $A = BB^*$ for some invertible $B$.
3. $A$ has all positive eigenvalues.
4. $A = \exp B = \sum_{k=0}^{\infty} B^k/k!$ for some (unique) hermitian $B$.
5. $A = UDU^*$ for some unitary $U$ and diagonal $D$ with positive diagonal entries.

The positive definite $m \times m$-matrices form an open cone in $\mathbb{H}_m$, the $m \times m$ hermitian matrices, with closure the positive semidefinite matrices (equivalently, $\langle Ax, x \rangle \geq 0$ for all $x$). We denote the open cone of positive definite matrices by $\mathbb{P}$ (or $\mathbb{P}_m$ if we need to distinguish the dimension). The exponential map $\exp : \mathbb{H} \to \mathbb{P}$ is an analytic diffeomorphism with inverse analytic diffeomorphism $\log : \mathbb{P} \to \mathbb{H}$.

Every positive definite (resp. hermitian) matrix operator has a unique spectral decomposition

$$A = \sum_{i=1}^{k} \lambda_i E_i,$$
where the $\lambda_i > 0$ (resp. $\lambda_i \in \mathbb{R}$) range over the distinct eigenvalues of $A$ and $E_i$ is the orthogonal projection onto the eigenspace of $\lambda_i$. One then has

$$A^p = \sum_{i=1}^{k} \lambda_i^p E_i,$$

from which one can easily deduce that every positive definite matrix has a unique positive definite $p^{th}$-root. We also note that the exponential map is given alternatively by

$$\exp A = \exp \left( \sum_{i=1}^{k} \lambda_i E_i \right) = \sum_{i=1}^{k} e^{\lambda_i} E_i,$$

from which we can quickly deduce the equivalence of items (3) and (4) in the previous list of equivalent characterizations of positive definite matrices.

We define a partial order (sometimes called the Loewner order) on the vector space $\mathbb{H}_m$ of hermitian matrices by $A \leq B$ if $B - A$ is positive semidefinite. We note $0 \leq A$ iff $A$ is positive semidefinite and write $0 < A$ if $A \in \mathbb{P}$ iff $A$ is positive definite. The matrix $A$ is sometimes called strictly positive in this setting.

For any invertible $M \in \mathcal{M}_m(\mathbb{C})$, the congruence transformation $\Gamma_M(X) = MXM^*$ is an invertible linear map on $\mathcal{M}_m$ that carries each of $\mathbb{H}$, $\mathbb{P}$, and $\mathbb{P}$, the convex cone of positive semidefinite matrices, onto itself. It follows that congruence transformations preserve the Loewner order on $\mathbb{H}$: $A \leq B$ implies $MAM^* \leq MBM^*$ for $M$ invertible. We note, in particular, for each $M \in \mathbb{P}$, $\Gamma_M(X) = MXM$ is a congruence transformation on $\mathbb{P}$. Matrix inversion $A \mapsto A^{-1}$ maps $\mathbb{P}$ onto itself and reverses the Loewner order, as we shall see later.

The geometry of $\mathbb{P}$ will be crucial in what follows. One important approach to geometry is that of Felix Klein’s Erlangen Program, which emphasized the importance of the group of transformations or “symmetries” that preserved basic geometric properties. For the study of $\mathbb{P}$ this group, denoted $G(\mathbb{P})$, is the one generated by the congruence transformations and the inversion map, which acts as the point reflection through the identity matrix.

**Remark 3.1.** For $A, B \in \mathbb{P}$, by the previous item (5), there exists a unitary $U$ such that $U(A^{-1/2}BA^{-1/2})U^* = D$, a diagonal matrix, and hence $\Gamma_{U,A^{-1/2}}$ carries $A$ to $I$. 
and $B$ to $D$. This observation allows various results about $A, B \in \mathbb{P}$ to be reduced to the case $A = I$ and $B$ is a diagonal matrix.

The arithmetic and harmonic means readily extend from $\mathbb{R}^{>0} := (0, \infty)$ to the set $\mathbb{P}$ of positive definite matrices:

$$A(A, B) = \frac{1}{2}(A + B); \quad \mathcal{H}(A, B) = 2(A^{-1} + B^{-1})^{-1}.$$  

The geometric mean is not so obvious (e.g., $\sqrt{AB}$, the square root of $AB$ with positive eigenvalues, need not be positive definite for $A, B$ positive definite). One approach is to rewrite the equation $x^2 = ab$ (which has positive solution the geometric mean of $a$ and $b$) in its appropriate form in the noncommutative setting and solve for $X$:

$$XA^{-1}X = B$$

$$A^{-1/2}XA^{-1/2}A^{-1/2}XA^{-1/2} = A^{-1/2}BA^{-1/2}$$

$$A^{-1/2}XA^{-1/2} = (A^{-1/2}BA^{-1/2})^{1/2}$$

$$X = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$  

**Definition 3.2.** The matrix geometric mean $A \# B$ of $A, B \in \mathbb{P}$ is given by

$$A \# B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$  

Alternatively it can be characterized as the unique positive definite solution $X$ of the elementary Riccati equation $XA^{-1}X = B$.

By inverting both sides $XA^{-1}X = B$ and multiplying through by $X$ on the right and left, one obtains $XB^{-1}X = A$. Since the second equation is equivalent to the first, we see $A \# B = B \# A$. Similarly one can use the Riccati equation to show that the matrix geometric mean is invariant under congruence mappings and inversion, i.e., $M(A \# B)M^* = (MA^*)^\#(MBM^*)$ for $M$ invertible and $(A \# B)^{-1} = A^{-1} \# B^{-1}$. By the Riccati equation $(A \# B)(A^{-1} - B^{-1})(A \# B) = B - A$, and this along with order invariance under congruence maps shows that the inversion map is order reversing. We collect these properties together with other basic properties that can be deduced by elementary arguments; see [15].

**Proposition 3.3.** The following hold in $\mathbb{P}$:

(i) (Commutativity) $A \# B = B \# A$. 

(ii) \textit{(Congruence Invariance)} \( M(A\#B)M^* = MAM^*\#MBM^* \) for \( M \) invertible.

(iii) \textit{(Inversion Invariance)} \((A\#B)^{-1} = A^{-1}\#B^{-1}\).

(iv) \textit{(Monotonicity)} If \( C \leq A \) and \( D \leq B \), then \( C\#D \leq A\#B \).

(v) For \( AB = BA \), \( A\#B = A^{1/2}B^{1/2} = (AB)^{1/2} \); in particular \( A\#I = A^{1/2} \).

(vi) \textit{(AGM Inequality)} \( 2(A^{-1} + B^{-1})^{-1} \leq A\#B \leq (1/2)(A + B) \).

By the Loewner-Heinz inequality \( A\#I = A^{1/2} \leq B^{1/2} = B\#I \) for \( 0 < A \leq B \).

From this, using congruence invariance, one obtains the important property (iv).

As mentioned in the introduction, other formulations and connections between the matrix geometric mean and the one for positive real numbers may be found in [15].

4. The Riemannian Metric

The tools needed for extending the binary geometric mean on \( \mathbb{P} \) to a multivariable one have relied on the metric and geometric structure of \( \mathbb{P} \). Such considerations had already begun in the binary setting; see [15] Section 4. We briefly overview the necessary tools in this section; see [15] Section 4 and [5] Chapter 6 for details.

We equip the space \( \mathbb{H} \) of hermitian matrices of some fixed dimension \( m \) with the \textit{Frobenius} inner product \( \langle A, B \rangle = \text{Tr}(AB) \), the trace of \( AB \), which makes \( \mathbb{H} \) a Hilbert space. The corresponding norm \( \|A\|_2 = \langle A, A \rangle^{1/2} \) is called the \textit{Frobenius} or \textit{Hilbert-Schmidt} norm. We can write \( A = UDU^* \) for some unitary \( U \), where \( D \) is a diagonal matrix with entries the eigenvalues of \( A \). We now observe that

\[
\|A\|_2 = (\text{Tr}(UDU^*)^{1/2}) = (\text{Tr}D^{1/2}) = \left( \sum_{i=1}^{m} \lambda_i^2(A) \right)^{1/2},
\]

where \( \{\lambda_i\}_{i=1}^{m} \) is the set of eigenvalues of \( A \).

We define the \textit{Riemannian metric} \( \delta \) on \( \mathbb{P} \) by

\[
\delta(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_2 = \left( \sum_{i=1}^{m} \log^2 \lambda_i(A^{-1/2}BA^{-1/2}) \right)^{1/2},
\]

where \( \{\lambda_i\} \) are the eigenvalues of \( A^{-1/2}BA^{-1/2} \). In the last expression we may replace \( A^{-1/2}BA^{-1/2} \) by \( BA^{-1} \), since the two are similar and hence have the same eigenvalues.

We list basic properties of the Riemannian metric \( \delta \).
Proposition 4.1. The Riemannian metric $\delta$ is a metric making $\mathbb{P}$ a complete metric space exhibiting the following properties:

1. For $M \in \text{GL}_m(\mathbb{C})$, the group of $m \times m$ invertible matrices, the congruence transformation $\Gamma_M : (\mathbb{P}, \delta) \rightarrow (\mathbb{P}, \delta)$ defined by $\Gamma_M(X) = MXM^*$ is an isometry. The inversion map $A \mapsto A^{-1}$ is also an isometry on $\mathbb{P}$.

2. The exponential map $\exp : \mathbb{H} \rightarrow \mathbb{P}$ is expansive, that is, $\delta(e^A, e^B) \geq \|B - A\|_2$ for all $A, B \in \mathbb{H}$.

3. The exponential map restricted to any one-dimensional subspace of $\mathbb{H}$ is an isometry. Furthermore, any metric on $\mathbb{P}$ that has this property and is invariant under congruence transformations must agree with $\delta$.

4. For $A, B \in \mathbb{P}$, $A \# B$ is the unique metric midpoint between $A$ and $B$.

Proof. To give some flavor of the preceding, we prove parts of (1), (3) and (4). First for (1) we observe for $A, B \in \mathbb{P}$ and $M$ invertible, $(MBM^*)(MAM^*)^{-1} = M(BA^{-1})M^{-1}$, which is similar to $BA^{-1}$, and hence to $A^{-1/2}BA^{-1/2}$, and thus may replace it in last part of equation (4.2).

For (3), let $A \in \mathbb{H}$. Then

$$\delta(e^{sA}, e^{tA}) = \delta(e^{(-s/2)A}e^{sA}e^{(-s/2)A}, e^{(-s/2)A}e^{tA}e^{(-s/2)A}) = \delta(I, e^{(t-s)A}) = \|\log(e^{(t-s)A})\|_2 = \|tA - sA\|_2.$$ 

Now let $d(\cdot, \cdot)$ be a metric satisfying the two properties of (3). For $A, B \in \mathbb{P}$,

$$d(A, B) = d(A^{-1/2}AA^{-1/2}, A^{-1/2}BA^{-1/2}) = d(I, A^{-1/2}BA^{-1/2}) = \|\log(A^{-1/2}BA^{-1/2})\|_2 = \delta(A, B).$$

For the midpoint property in (4), since congruence mappings preserve $\delta$ and $\#$, it suffices to consider the case $A \# B = I$ (otherwise first apply $\Gamma_{(A\#B)^{-1/2}}$). Then $B = IA^{-1}I = A^{-1}$. Applying log we obtain $\log B = -\log A$, so restricting exp to the one-dimensional subspace $\mathbb{R} \cdot A$ yields the result by (3).

A proof of (2) is sketched in [15], and a shorter and more elegant proof appears in [5, Chapter 6].
For the triangular inequality, consider first the case $A, B, C$ with $A = I$. Then from (3) $\delta(I, B) = \|\log B\|_2$, $\delta(I, C) = \|\log C\|_2$ and from (2) $\|\log(C) - \log(B)\|_2 \leq \delta(B, C)$, so

$$\delta(I, C) = \|\log C\|_2 \leq \|\log B\|_2 + \|\log(C) - \log(B)\|_2 \leq \delta(I, B) + \delta(B, C).$$

The general case follows using (1) and the congruence transformation $X \mapsto A^{-1/2}XA^{-1/2}$ and its inverse.

By (2) and continuity of the exponential map, the metric $\delta$ is complete. □

For a metric space $(X, d)$ the metric is said to satisfy the \textbf{semiparallelogram law} if for all $x_1, x_2 \in X$, there exists $m \in X$ such that for any $x \in X$,

$$d^2(x_1, x_2) + 4d^2(x, m) \leq 2d^2(x, x_1) + 2d^2(x, x_2) \quad (NPC)$$

One can show that $m = m(x_1, x_2)$ is unique and is the unique metric midpoint between $x_1$ and $x_2$. If the inequality is replaced by an equality, one obtains a version of the parallelogram law holding in Hilbert spaces. The semiparallelogram law is a metric version of nonpositive curvature (NPC). We define a \textit{Hadamard space} to be a complete metric space satisfying the semiparallelogram law. These spaces have been and continue to be widely studied and appear under the alternative names of global NPC-spaces or CAT(0)-spaces.

Using Proposition 4.1 it is straightforward to show that $(\mathbb{P}, \delta)$ is a Hadamard space. One first considers the case that $A#B = I$ and another point $C$. Then the parallelogram law holds in the Hilbert space $\mathbb{H}$ for $\log A$, $\log B = -\log A$, and $\log C$, and one uses Proposition 4.1(2) to obtain the semiparallelogram law for $A, B, C$. Via 4.1(1) the general case reduces to this one. See, for example, [15] for further details.

Corollary 4.2. \textit{The space $(\mathbb{P}, \delta)$ is a Hadamard space.}

Remark 4.3. A more sophisticated approach to the results of this section is the path of Riemannian geometry. The open cone $\mathbb{P}_m$ of $m \times m$ positive definite matrices becomes a well-known Riemannian manifold when equipped with the trace Riemannian metric: $\langle X, Y \rangle_A = \text{tr} A^{-1}XA^{-1}Y$, where $A \in \mathbb{P}_m$ and $X, Y$ are $m \times m$ Hermitian matrices. The corresponding distance metric on $\mathbb{P}_m$ is precisely our metric $\delta$, and
this is the source of the name “Riemannian metric.” The distance metric of a simply connected Riemannian manifold satisfies (NPC) iff the manifold has nonpositive curvature in the usual sense.

5. MEANS OF SEVERAL VARIABLES

Formally a mean of order \( n \), or \( n \)-mean for short, on a set \( X \) is a function \( \gamma : X^n \to X \) satisfying the idempotency condition: \( \forall x \in X, \gamma(x, x, \ldots, x) = x \). It is frequently assumed in the definition of a mean that it is symmetric, that is, invariant under any permutation of variables. When we speak of an omnivariable mean \( \gamma = \{\gamma_n\} \), we are referring to one defined for all \( X^n, n \geq 1 \). (For \( n = 1 \), \( \gamma_1 \) is the identity, and is thus frequently ignored.) The mean \( \gamma : X^n \to X \) is continuous or a topological mean if \( X \) is a topological space and \( \gamma \) is continuous. Frequently a mean represents some type of averaging operator.

In 2004 T. Ando, C. K. Li and R. Mathias [3] gave the first extension of the binary geometric mean to an omnivariable mean on \( \mathbb{P} \), which came to be called the ALM mean. For three variables \( A, B, C \), they first took the new three point set \( \{A_1 := B\#C, B_1 := A\#C, C_1 := A\#B\} \) consisting of the geometric mean (midpoint) of each pair, then repeated this construction on the new three point set and continued repeating the operation inductively. They showed the triples approached a common point, their mean for the case \( n = 3 \). They extended it inductively to an \( n \)-variable mean for all \( n > 2 \). We note that Lawson and Lim [16] later extended the ALM construction to a rather wide class of metric spaces.

Ando, Li, and Mathias made two important contributions in their paper. First of all, they identified axiomatic properties that an omnivariable geometric mean \( \gamma \) should satisfy. They then established that the ALM mean they had defined satisfied all these properties. The proofs typically involved extending from the known case of \( n = 2 \) by induction.

**Theorem 5.1.** Let \( \mathbb{A} = (A_1, \ldots, A_n), \mathbb{B} = (B_1, \ldots, B_n) \in \mathbb{P}^n \) and let \( \gamma \) be the ALM mean. The following properties hold.

(P1) (Consistency with scalars) \( \gamma(\mathbb{A}) = (A_1 \cdots A_n)^{1/n} \) if the \( A_i \)'s commute;

(P2) (Joint homogeneity) \( \gamma(a_1 A_1, \ldots, a_n A_n) = (a_1 \cdots a_n)^{1/n} \gamma(\mathbb{A}) \);

(P3) (Permutation invariance) \( \gamma(\mathbb{A}_\sigma) = \gamma(\mathbb{A}) \), where \( \mathbb{A}_\sigma = (A_{\sigma(1)}, \ldots, A_{\sigma(n)}) \).
(P4) \textbf{(Monotonicity)} If \( A_i \leq B_i \) for all \( 1 \leq i \leq n \), then \( \gamma(A) \leq \gamma(B) \);

(P5) \textbf{(Continuity)} \( \gamma \) is continuous;

(P6) \textbf{(Congruence invariance)} \( \gamma(MA'M^*) = M\gamma(A)M^* \) for \( M \) invertible, where \( M(A_1, \ldots, A_n)M^* = (MA_1M^*, \ldots, MA_nM^*) \);

(P7) \textbf{(Joint concavity)} \( \gamma(\lambda A + (1 - \lambda) B) \geq \lambda \gamma(A) + (1 - \lambda) \gamma(B) \) for \( 0 \leq \lambda \leq 1 \);

(P8) \textbf{(Self-duality)} \( \gamma(A_1^{-1}, \ldots, A_n^{-1})^{-1} = \gamma(A_1, \ldots, A_n) \);

(P9) \textbf{(Determinantal identity)} \( \text{Det}\gamma(A) = \prod_{i=1}^{n}(\text{Det}A_i)^{1/n} \); and

(P10) \textbf{(AGH mean inequalities)} \( n(\sum_{i=1}^{n}A_i^{-1})^{-1} \leq \gamma(A) \leq \frac{1}{n} \sum_{i=1}^{n}A_i \).

(D. Bini, B. Meini, and F. Poloni \cite{9} later gave a variant of the ALM mean that retained its properties, but was much more computationally efficient.

A \textit{weight} \( \mathbf{w} \) of length \( n \) is an \( n \)-tuple \( (w_1, \ldots, w_n) \) where \( 0 < w_i \leq 1 \) for each \( i \) and \( \sum_{k=1}^{n}w_k = 1 \), and a \textit{weighted} \( n \)-tuple of a set \( X \) is a pair \( (\mathbf{w}, \mathbf{x}) \), where \( \mathbf{w} \) is a weight of length \( n \) and \( \mathbf{x} = (x_1, \ldots, x_n) \in X^n \). We think of this as convenient notation for an ordered \( n \)-tuple of weighted points \( (x_i, w_i) \). An \( n \)-variable \textit{weighted mean} \( \gamma \) on a set \( X \) assigns to each weighted \( n \)-tuple \( (\mathbf{w}, \mathbf{x}) \) some \( \gamma(\mathbf{w}, \mathbf{x}) \in X \) with the extra condition that \( \gamma(\mathbf{w}, (x, x, \ldots, x)) = x \). We may think of \( \gamma(\mathbf{w}, \mathbf{x}) \) as the assigning of a “center of mass.”

\textbf{Remark 5.2.} The \textit{weighted geometric mean} \( A\#_t B \) (with \( \mathbf{w} = (1-t, t) \)) is given by

\[ A\#_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^tA^{1/2} \]

The map \( \alpha : [0, \delta(A, B)] \rightarrow \mathbb{P} \) defined by \( \alpha(t) = A\#_t B \) is an isometry onto the geodesic arc between \( A \) and \( B \).

Rather obvious variants of the axiomatic properties (P1)-(P10) exist in the weighted mean setting and were introduced and studied in \cite{14}. There a weighted version of the mean given by Bini, Meini, and Poloni was introduced, which could also be extended to more general metric spaces.

Closely related to the notion of a omnivariable weighted mean is that of a barycentric map. Denote by \( \mathcal{P}^{<\infty}(X) \) the set of all finitely supported probability measures on \( X \), measures of the form \( \sum_{k=1}^{n}w_k\delta_{x_k} \), where \( (w_1, \ldots, w_n) \) is a weight and \( \delta_x \) is the unit point mass at \( x \). A \textit{barycentric map} in its simplest manifestation is a map.
\( \beta : \mathcal{P}^{<\infty}(X) \to X \) such that \( \beta(\delta_x) = x \) for each \( x \in X \). A barycentric \( \beta \) gives rise to a corresponding multivariable weighted mean defined by

\[
\gamma_\beta(w; x_1, \ldots, x_n) = \beta\left( \sum_{k=1}^{n} w_k \delta_{x_k} \right),
\]

where \( w = (w_1, \ldots, w_n) \) is a weight. We note, however, that it is not the case that all weighted means arise from barycentric maps. If we restrict to uniform weights \( (w_i = 1/n \text{ for each } i) \), we are essentially in the setting of non-weighted means.

Soon after the introduction of the ALM mean, a better candidate for the multivariable matrix geometric mean was put forth to which we now turn. Let \((M, d)\) be a metric space. The least squares mean \( \Lambda(a_1, \ldots, a_n) \) is defined as the solution to the optimization problem of minimizing the sum of distances squared

\[
\Lambda(a_1, \ldots, a_n) = \arg \min_{x \in M} \sum_{i=1}^{n} d^2(x, a_i),
\]

and the weighted least squares mean \( \Lambda(w; a_1, \ldots, a_n) \) for \( w = (w_1, \ldots, w_n) \) by

\[
\Lambda(w; a_1, \ldots, a_n) = \arg \min_{x \in M} \sum_{i=1}^{n} w_id^2(x, a_i),
\]

provided the solution uniquely exists in each of the respective cases. Note that the least squares mean is equal to the weighted least squares mean for the uniform weight with all entries \( 1/n \). The unique solution of the minimizing problem exists in Hadamard spaces \( [29, \text{Proposition 1.7}] \), since the non-negative function defined by \( x \mapsto \sum_{i=1}^{n} w_id^2(x, a_i) \) is uniformly convex in this case. (We recall \( F : M \to \mathbb{R} \) is uniformly convex if there exists a strictly increasing \( \varphi : [0, \infty) \to [0, \infty) \) such that for all \( x, y \in M \),

\[
F(m(x, y)) \leq \frac{1}{2}F(x) + \frac{1}{2}F(y) - \varphi(d(x, y)),
\]

where \( m(x, y) \) is the unique midpoint between \( x \) and \( y \).)

E. Cartan considered such “barycenters” in the case of Riemannian manifolds, where they uniquely exist for the ones of nonpositive curvature and exist locally much more generally, and M. Fréchet considered them in more general metric spaces. Thus the least squares mean is also called the Cartan mean or Fréchet mean. Such means also frequently go by the name “Karcher means,” but Karcher’s approach from
differential geometry involved finding the solution of the “Karcher equation” that was satisfied at this extremum [11]. We return to this in a later section.

M. Moakher (2005) [24] first introduced and studied the least squares mean for the set of positive definite matrices $\mathbb{P}$ equipped with the Riemannian metric as a omnivariable generalization of the two-variable geometric mean. Independently R. Bhatia and J. Holbrook (2006) [6, 7] introduced and studied the least squares mean in the weighted setting. These authors established its (unique) existence and verified most of the axiomatic properties (P1)-(P10) satisfied by the Ando-Li-Mathias geometric mean: consistency with scalars, joint homogeneity, permutation invariance, congruence invariance, and self-duality (the last two being true since congruence transformations and inversion are isometries). Further, based on computational experimentation, Bhatia and Holbrook conjectured monotonicity for the least squares mean (property (P4) in the earlier list), but this was left as an open problem.

6. The Inductive Mean, Random Variables, and Monotonicity

One other mean played an important role in what followed, one that we shall call the inductive mean, following the terminology of K.-T. Sturm (2003) [29]. It appeared elsewhere in the work of M. Sagae and K. Tanabe (1994) [28] and Ahn, Kim, and Lim (2007) [1]. It is defined inductively for Hadamard spaces (or more generally for metric spaces with weighted binary means $x\#_t y$) for each $k \geq 2$ by $S_2(x, y) = x\#y$ and for $k \geq 3$, $S_k(x_1, \ldots, x_k) = S_{k-1}(x_1, \ldots, x_{k-1})\#_t x_k$. (Here $x\#_t y$ is the unique point $z$ such that $d(x, y) = (1-t)d(x, z) + td(y, z)$ for $0 \leq t \leq 1$.) Note that this mean at each stage is defined from the previous stage by taking the appropriate two-variable weighted mean, which is monotone for the Hadamard space $\mathbb{P}$. Thus the inductive mean is monotone (property (P4)).

Let $(M, d)$ be Hadamard space and $\mathbf{w} = (w_1, \ldots, w_n)$ be a weight. Set $\mathbb{N}_n = \{1, 2, \ldots, n\}$ and assign to $k \in \mathbb{N}_n$ the probability $w_k$. Let $\mathbf{x} = (x_1, \ldots, x_n) \in M^n$. For each $\omega \in \prod_{j=1}^{\infty} \mathbb{N}_n$, a countable product, define a sequence $\sigma_{\omega, \mathbf{x}}$ in $M$ by $\sigma_{\omega, \mathbf{x}}(1) = x_{\omega(1)}$, $\sigma_{\omega, \mathbf{x}}(k) = S_k(x_{\omega(1)}, \ldots, x_{\omega(k)})$, where $S_k$ is the inductive mean. The sequence $\sigma_{\omega, \mathbf{x}}$ may be viewed as a “walk” starting at $\sigma_{\omega, \mathbf{x}}(1) = x_{\omega(1)}$ and moving toward $x_{\omega(k)}$ from $\sigma_{\omega, \mathbf{x}}(k - 1)$ a distance of $(1/k)d(\sigma_{\omega, \mathbf{x}}(k - 1), x_{\omega(k)})$ to obtain $\sigma_{\omega, \mathbf{x}}(k)$. Alternatively we may give $\prod_{j=1}^{\infty} \mathbb{N}_n$ the product probability, making it a probability
space, and define a family of i.i.d. random variables \( \{X_k\} \) by \( X_k(\omega) = x_{\omega(k)} \). We replace the traditional sum of the first \( k \) random variables by \( \sigma_{\omega, X}(k) \) and take for the expected value \( \Lambda(w; x_1, \ldots, x_m) \). From this viewpoint we have the following special case of Sturm’s Law of Large Numbers for Hadamard spaces [29, Theorem 4.7]:

**Sturm’s Theorem.** Giving \( \prod_{n=1}^{\infty} \mathbb{N}_m \) the product probability, the set

\[
\left\{ \omega \in \prod_{k=1}^{\infty} \mathbb{N}_m : \lim_{k \to \infty} \sigma_{\omega, X}(k) = \Lambda(w; x_1, \ldots, x_m) \right\}
\]

has measure 1, i.e., \( \sigma_{\omega, X}(k) \to \Lambda(w; x_1, \ldots, x_m) \) as \( k \to \infty \) for almost all \( \omega \).

We shall briefly return to the theory of random variables on a probability space taking values in \( \mathbb{P} \), or more generally in a Hadamard space, at a later point.

Using the preceding version of Sturm’s Theorem, Lawson and Lim (2011) [17] provided a positive solution to the earlier mentioned conjecture of Bhatia and Holbrook about the monotonicity of the least squares mean.

**Theorem 6.1.** Let \( \mathbb{P} \) be the open cone of positive definite matrices of some fixed dimension, and let \( n \geq 3 \)

1. The [weighted] least squares mean \( \Lambda \) on \( \mathbb{P} \) is monotone: \( A_i \preceq B_i \) for \( 1 \leq i \leq n \)
   implies \( \Lambda(A_1, \ldots, A_n) \preceq \Lambda(B_1, \ldots, B_n) \) \( [\Lambda(w; A_1, \ldots, A_n) \preceq \Lambda(w; B_1, \ldots, B_n)] \).

2. The other nine (weighted) ALM axioms hold for \( \Lambda \).

**Proof.** Assume for \( \mathbb{A} = (A_1, \ldots, A_m) \) and \( \mathbb{B} = (B_1, \ldots, B_m) \) that \( A_i \preceq B_i \) for \( 1 \leq i \leq m \). Let \( w \) be a weight. By Sturm’s theorem applied to the “walks” for \( \mathbb{A} \) and for \( \mathbb{B} \), we have

\[
\sigma_{\omega, \mathbb{A}}(k) \to \Lambda(w; A_1, \ldots, A_n) \quad \text{and} \quad \sigma_{\omega, \mathbb{B}}(k) \to \Lambda(w; B_1, \ldots, B_n)
\]

as \( k \to \infty \) for almost all \( \omega \in \prod_{n=1}^{\infty} \mathbb{N}_m \) (since the intersection of two sets of measure 1 has measure 1). Fixing any such \( \omega \), we obtain part (1) since the partial order relation is closed and by the monotonicity of the inductive mean for each \( k \)

\[
\sigma_{\omega, \mathbb{A}}(k) = S_k(A_{\omega(1)}, \ldots, A_{\omega(k)}) \preceq S_k(B_{\omega(1)}, \ldots, B_{\omega(k)}) = \sigma_{\omega, \mathbb{B}}(k).
\]

Later in [8] R. Bhatia and R. Karandikar proved the monotonicity property using in place of Sturm’s theorem some elementary counting arguments and basic inequalities
for the metric $\delta$. Meanwhile J. Holbrook [10] identified a specific $\omega \in \prod_{n=1}^{\infty} N_m$ that always yielded a “walk” converging to $\Lambda(A_1, \ldots, A_n)$ for any choice of $A_1, \ldots, A_n \in \mathbb{P}$. Later Y. Lim and M. Palfia [22] obtained more general results about deterministic walks yielding the least squares mean in general Hadamard spaces.

**Note:** The ALM mean is typically distinct from the least squares mean for $n \geq 3$. Thus the ALM axioms do not characterize a mean. The latter fact had already been noted by Bini, Meini and Poloni (2010), who observed their variant of the ALM mean was different from it [9]. In [21], Lim and Palfia showed that the Karcher mean is uniquely determined by congruence invariancy (P6), self-duality (P8), and the following Yamazaki inequality [30]:

$$\sum_{k=1}^{n} w_k \log A_k \leq 0 \implies \Lambda(w; A_1, \ldots, A_n) \leq I.$$ 

7. The Hilbert Space Setting

Let $\mathcal{B}(E)$ be the $C^*$-algebra of bounded linear operators equipped with the operator norm on an infinite-dimensional Hilbert space $E$. Let $\mathbb{H}(E)$ denote the closed subspace of hermitian operators, and let $\mathbb{P}_E$ be the cone of invertible positive hermitian operators, an open cone in $\mathbb{H}(E)$. Again exp and log are analytic inverses between $\mathbb{H}(E)$ and $\mathbb{P}_E$. We equip $\mathbb{P}_E$ with the Thompson metric defined by $d_T(A, B) = \| \log(A^{-1/2}BA^{-1/2}) \|$, where $\| \cdot \|$ is the operator norm. The metric $d_T$ retains the four properties of $\delta$ in Proposition 4.1 in the finite-dimensional setting, except that $A\#B$ is no longer the unique midpoint between $A$ and $B$ [26]. Additionally $\mathbb{P}_E$ is no longer a Hadamard space, basically because $\mathbb{H}(E)$ is no longer a Hilbert space. The definition and basic properties of the geometric mean $A\#B$, in particular those of Proposition 3.3, remain valid, and it still gives a distinguished metric midpoint (no longer unique [26]) between $A$ and $B$. The smallest closed set containing $A$ and $B$ and closed under taking $\#$-midpoints yields a distinguished metric geodesic connecting $A$ and $B$ consisting of all $A\#_t B$, $0 \leq t \leq 1$.

8. The Karcher Equation

The uniform convexity of the Riemannian metric $\delta$ on $\mathbb{P}$ yields that the least squares mean is the unique critical point for the function $X \mapsto \sum_{k=1}^{n} w_k \delta^2(X, A_k)$. The
least squares mean is thus characterized by the vanishing of the gradient, which is equivalent to its being a solution of the following Karcher equation:

\begin{equation}
\sum_{k=1}^{n} w_k \log(X^{-1/2} A_k X^{-1/2}) = 0.
\end{equation}

The Karcher equation (8.3) can also be used to define a mean on the cone \( \mathbb{P} \) of positive invertible bounded operators on an infinite-dimensional Hilbert space (where one no longer has a Hadamard space), called the Karcher mean. As we just previously noted, restricted to the matrix setting it yields the least squares mean. Thus it is reasonable to continue to denote it by \( \Lambda(w, A_1, \ldots, A_n) \).

Power means for positive definite matrices were introduced by Lim and Palfia (2012) \[21\].

**Theorem.** Let \( A_1, \ldots, A_n \in \mathbb{P} \) and let \( w = (w_1, \ldots, w_n) \) be a weight. Then for each \( t \in (0, 1] \), the following equation has a unique positive definite solution \( X = P_t(w; A_1, \ldots, A_n) \), called the \( t \)-weighted power mean:

\[ X = \sum_{k=1}^{n} w_k (X \#_t A_i). \]

When restricted to the positive reals, the power mean reduces to the usual power mean

\[ P_t(w; a_1, \ldots, a_n) = (w_1 a_1^t + \cdots + w_n a_n^t)^{1/t}. \]

In 2014 Lawson and Lim showed that the preceding notion of power mean extended to the setting of bounded operators on a Hilbert space \[18\] and established that the power means are decreasing, \( s < t \) implies \( P_s(\cdot; \cdot) \leq P_t(\cdot; \cdot) \). Using power means Lawson and Lim were able to establish the existence and uniqueness of the Karcher mean in the \( C^* \)-algebra of bounded operators on a Hilbert space.

**Theorem 8.1.** In the strong operator topology

\[ \Lambda(\cdot; \cdot) = \lim_{t \to 0^+} P_t(\cdot; \cdot) = \inf_{t>0} P_t(\cdot; \cdot), \]

where \( \Lambda \) is the Karcher mean, the unique solution of the Karcher equation

\[ X = \Lambda(w; A_1, \ldots, A_n) \Leftrightarrow \sum_{k=1}^{n} w_k \log(X^{-1/2} A_k X^{-1/2}) = 0. \]
Via this machinery many of the axiomatic properties of the least squares mean in the finite-dimensional setting were extended to the corresponding Karcher mean in the infinite-dimensional setting.

Recent work by Lim-Palfia [23] and independently by Lawson [13] shows that the preceding constructions and results remain valid for the open cone of positive invertible elements in any unital $C^*$-algebra.

9. Barycenters

A Borel probability measure on a metric space $(X, d)$ is a countably additive non-negative measure $\mu$ on the Borel algebra $\mathcal{B}(X)$, the smallest $\Sigma$-algebra containing the open sets, such that $\mu(X) = 1$. We denote the set of all probability measures on $(X, \mathcal{B}(X))$ by $\mathcal{P}(X)$. Let $\mathcal{P}_0(X)$ be the set of all uniform finitely supported probability measures, i.e., all $\mu \in \mathcal{P}(X)$ of the form $\mu = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k}$ for some $n \in \mathbb{N}$, where $\delta_x$ is the point measure of mass 1 at $x$.

A measure $\mu \in \mathcal{P}(X)$ is said to be integrable if for some (and hence all) $x$

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

The set of integrable measures is denoted by $\mathcal{P}^1(X)$.

The Wasserstein distance (alternatively Kantorovich-Rubinstein distance) $d^W$ on $\mathcal{P}^1(X)$ is a standard metric for probability measures. It is known that $d^W$ is a complete metric on $\mathcal{P}^1(X)$ whenever $X$ is a complete metric space and that $\mathcal{P}_0(X)$ is $d^W$-dense in $\mathcal{P}^1(X)$.

**Definition 9.1.** ([23]) A map $\beta : (\mathcal{P}^1(X), d^W) \to (X, d)$ is a contractive barycentric map if (i) $\beta(\delta_x) = x$ for all $x \in X$, and (ii) $d(\beta(\mu), \beta(\nu)) \leq d^W(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}^1(X)$.

Suppose $G = \{G_n : X^n \to X : n \geq 2\}$ is a omnivariable mean on $(X, d)$, a metric space. The mean $G$ is said to be iterative if for all $n, k \geq 2$ and $x = (x_1, \ldots, x_n) \in X^n$,

$$G_n(x) = G_{nk}(x^k),$$

where $x^k = (x_1, \ldots, x_n, x_1, \ldots, x_n, \ldots, x_1, \ldots, x_n) \in X^{nk}$. The next proposition is a key result from [19].
Proposition 9.2. Suppose $G$ is an omnivariable mean on a complete metric space $(X, d)$. If $G$ is symmetric, iterative, and satisfies for all $n$, all $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$,

\begin{equation}
    d(G(\mathbf{x}), G(\mathbf{y})) \leq \frac{1}{n} \sum_{k=1}^{n} d(x_k, y_k),
\end{equation}

then there exists a unique contractive barycentric map $\beta : \mathcal{P}^1(X) \to X$ satisfying $G(x_1, \ldots, x_n) = \beta(\sum_{k=1}^{n}(1/n)\delta_{x_k})$.

The Wasserstein distance between $\sum_{k=1}^{n}(1/n)\delta_{x_k}$ and $\sum_{k=1}^{n}(1/n)\delta_{y_k}$ is the minimum over all permutations of $\{y_k\}_{k=1}^{n}$ of the sum on the right in equation (9.4), so the inequality implies $\beta$ is contractive on $\mathcal{P}_0(X)$, a dense subset, and hence extends uniquely to a contractive map on $\mathcal{P}^1(X)$.

The hypotheses of the preceding proposition are satisfied by the Karcher mean $\Lambda$, and thus the mean “extends” to a contractive barycentric map $\beta_\Lambda : \mathcal{P}^1(\mathcal{P}) \to \mathcal{P}$, called the Karcher barycentric map. It is characterized by

\[ X = \beta_\Lambda(\mu) \Leftrightarrow \int_{\mathcal{P}} \log(X^{-1/2}AX^{-1/2})d\mu(A) = 0. \]

The existence and basic theory and properties of the Karcher barycentric map can be found in [19]. Many of the basic properties of the least squares mean in Theorem 5.1 have analogous properties for the Karcher barycentric map.

10. $\mathcal{P}$-valued Random Variables

The Karcher barycentric map makes possible a theory of $\mathcal{P}$-valued random variables. Let $(Z, \mathcal{A}, q)$ be a probability space, where $Z$ is a set, $\mathcal{A}$ is a $\sigma$-algebra, and $q$ is a probability measure on $\mathcal{A}$. Let $X : Z \to \mathcal{P}$ be measurable. With respect to $X$ there is push-forward measure $X_*(q)$ on $\mathcal{P}$ defined by $X_*(q)(Q) = q(X^{-1}(Q))$ for any Borel subset $Q$ of $\mathcal{P}$. This measure is called the distribution of $X$. We are particularly interested in the $L^1$ random variables, those for which $X_*(q) \in \mathcal{P}^1(\mathcal{P})$. In this case we can define the expectation $E(X) = \beta(X_*(q))$, where $\beta$ is the Karcher barycentric map. These ideas can be worked out in Hadamard spaces to a rather full-blown theory as has been done by Sturm in [29], and we refer the reader to that source for
details. However, the infinite dimensional \( \mathbb{P} \) is not a Hadamard space, so requires new approaches; see [23].

11. Summary

In the preceding we have attempted to outline the high points of the striking development of the theory of the matrix/operator geometric mean on the cone of positive matrices/operators in the past twenty years. Over this period of time it has evolved from a two-variable matrix mean to a omnivariable matrix mean (the least squares, Cartan, or Frechét mean) to an operator mean in the setting of unital \( C^* \)-algebras (the Karcher mean) and finally to a barycentric map on the space of integrable Borel probability measures. At each stage of the evolution significant new insights and developments were necessary. The theory has drawn heavily from matrix and operator theory and at the same time from geometric notions. And along the way we have seen a variety of characterizations of this mean: the least squares mean, the probabilistic or deterministic characterization as a limit of a “walk” with the inductive mean, the solution of the Karcher equation, and the limit/infimum of the power means \( \{P_t\} \) for \( t \searrow 0 \). Whatever future developments may hold, it is clear that a substantial theory has already emerged.

References

[1] E. Ahn, S. Kim, and Y. Lim, An extended Lie-Trotter formula and its applications, Linear Algebra Appl. 427 (2007), 190-196.
[2] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl. 26 (1979), 203-241.
[3] T. Ando, C. K. Li and R. Mathias, Geometric means, Linear Algebra Appl. 385 (2004), 305-334.
[4] F. Barbaresco, Innovative tools for radar signal processing based on Cartan’s geometry of symmetric positive definite matrices and information geometry, In: Proceedings of the IEEE International Radar 407 Conference (Rome, Italy 2008).
[5] R. Bhatia, Positive Definite Matrices, Princeton University Press, Princeton, 2007.
[6] R. Bhatia and J. Holbrook, Noncommutative geometric means, Math. Intelligencer 28 (2006), 32-39.
[7] R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006), 594-618.
[8] R. Bhatia and R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. 353 (2012), 1453-1467.
[9] D. Bini, B. Meini and F. Poloni, An effective matrix geometric mean satisfying the Ando-Li-Mathias properties, Math. Comp. 79 (2010), 437-452.
[10] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc. 27 (2012), 509-521.
[11] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. 30 (1977), 509-541.
[12] F. Kubo and T. Ando, Means of positive linear operators, Math. Ann. 246 (1980), 205-224.
[13] J. D. Lawson, Existence and uniqueness of the Karcher mean on unital $C^*$-algebras, to appear J. Math. Anal. and Appl.
[14] J. D. Lawson, H. Lee, and Y. Lim, Weighted geometric means, Forum Math. 24 (2012), 1067-1090.
[15] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, Amer. Math. Monthly 108 (2001), 797-812.
[16] —, A general framework for extending means to higher orders, Colloq. Math. 113 (2008), 191-221.
[17] —, Monotonic properties of the least squares mean, Math. Ann. 351 (2011), 267-279.
[18] —, Karcher means and Karcher equations of positive operators, Trans. Amer. Math. Soc. Series B 1 (2014), 1-22.
[19] —, Contractive barycentric maps, J. Operator Theory 77 (2017), 87-107.
[20] —, Following the trail of the operator geometric mean, Geometric Methods in Physics XXXVIII, Trends Math., Birkhäuser/Springer, Cham, 2020, 143-153.
[21] Y. Lim and M. Pálfia, The matrix power means and the Karcher mean, J. Funct. Anal. 262 (2012), 1498-1514.
[22] —, Weighted deterministic walks for the squares mean on Hadamard spaces, Bull. Lond. Math. Soc. 46 (2014), 561-570.
[23] —, Existence and uniqueness of the $L^1$-Karcher mean, Adv. in Math. 376 (2021), 107435.
[24] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, SIAM J. Matrix Anal. Appl. 26 (2005), 735-747.
[25] M. Moakher and M. Zéraï, The Riemannian geometry of the space of positive-definite matrices and its application to the regularization of positive-definite matrix-valued data, J. Math. Imaging Vision 40 (2011), 171-187.
[26] R. D. Nussbaum, Finsler structures for the part metric and Hilbert’s projective metric and applications to ordinary differential equations, Differential and Integral Equations 7 (1994), 1649-1707.
[27] W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, Reports Math. Phys. 8 (1975), 159-170.

[28] M. Sagae and K. Tanabe, Upper and lower bounds for the arithmetic-geometric-harmonic means of positive definite matrices, Linear and Multilinear Algebras 37 (1994), 279-282.

[29] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, Eds. P. Auscher et. al., (Paris, 2002), 357-390, Contemp. Math. 338, Amer. Math. Soc. (AMS), Providence, 2003.

[30] T. Yamazaki, The Riemannian mean and matrix inequalities related to the Ando-Hiai inequality and chaotic order, Operators and Matrices 6(2012), 577–588.

Department of Mathematics, Louisiana State University, Baton Rouge, LA70803, USA

Email address: lawson@math.lsu.edu

Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Korea

Email address: ylim@skku.edu