SIAM J. MATRIX ANAL. APPL.
Vol. 26, No. 3, pp. 735–747

A DIFFERENTIAL GEOMETRIC APPROACH TO THE GEOMETRIC MEAN OF SYMMETRIC POSITIVE-DEFINITE MATRICES*

MAHER MOAKHER†

Abstract. In this paper we introduce metric-based means for the space of positive-definite matrices. The mean associated with the Euclidean metric of the ambient space is the usual arithmetic mean. The mean associated with the Riemannian metric corresponds to the geometric mean. We discuss some invariance properties of the Riemannian mean and we use differential geometric tools to give a characterization of this mean.

Key words. geometric mean, positive-definite symmetric matrices, Riemannian distance, geodesics

AMS subject classifications. 47A64, 26E60, 15A48, 15A57

DOI. 10.1137/S0895479803436937

1. Introduction.

Almost 2500 years ago, the ancient Greeks defined a list of 10 (actually 11) distinct "means" [14, 23]. All these means are constructed using geometric proportions. Among these are the well-known arithmetic, geometric, and harmonic (originally called "subcontrary") means. These three principal means, which are used particularly in the works of Nicomachus of Gerasa and Pappus, are the only ones out of the original 11 that are still commonly used.

The arithmetic, geometric, and harmonic means, originally defined for two positive numbers, generalize naturally to a finite set of positive numbers. In fact, for a set of m positive numbers, $\{x_k\}_{1\le k\le m}$, the arithmetic mean is the positive number $\bar{x} = \frac{1}{m}\sum_{k=1}^{m} x_k$. The arithmetic mean has a variational property; it minimizes the sum of the squared distances to the given points x_k,

$$\bar{x} = \operatorname*{arg\,min}_{x>0} \sum_{k=1}^{m} \mathrm{d}_e(x, x_k)^2, \tag{1.1}$$

where $\mathrm{d}_e(x, y) = |x - y|$ represents the usual Euclidean distance in $\mathbb{R}$. The geometric mean of $x_1, \dots, x_m$, which is given by $\tilde{x} = \sqrt[m]{x_1 x_2 \cdots x_k}$, also has a variational property; it minimizes the sum of the squared *hyperbolic distances* to the given points x_k,

$$\tilde{x} = \operatorname*{arg\,min}_{x>0} \sum_{k=1}^{m} \mathrm{d}_h(x_k, x)^2, \tag{1.2}$$

where $\mathrm{d}_h(x, y) = |\log x - \log y|$ is the hyperbolic distance[1] between x and y. The harmonic mean of the set of m positive numbers $\{x_k\}_{1\le k\le m}$ is simply given by the

*Received by the editors October 27, 2003; accepted for publication (in revised form) by U. Helmke June 12, 2004; published electronically April 8, 2005. This work was partially supported by the Swiss National Science Foundation.

http://www.siam.org/journals/simax/26-3/43693.html

†Laboratory for Mathematical and Numerical Modeling in Engineering Science, National Engineering School at Tunis, Tunis El-Manar University, ENIT-LAMSIN, B.P. 37, 1002 Tunis-Belvédère, Tunisia (Maher.Moakher@enit.rnu.tn).

[1]We borrow this terminology from the hyperbolic geometry of the Poincaré upper half-plane. In fact, the hyperbolic length of the geodesic segment joining the points $P(a, y_1)$ and $Q(a, y_2)$, $y_1, y_2 > 0$, is $|\log \frac{y_1}{y_2}|$ (see [24, 27]).

inverse of the arithmetic mean of their inverses, i.e., $\hat{x} = [\frac{1}{m}\sum_{k=1}^{m}(x_k)^{-1}]^{-1}$, and thus the harmonic mean has a variational characterization as well.

The arithmetic mean has been widely used to average elements of linear Euclidean spaces. Depending on the application, it is usually referred to as the average, the barycenter, or the center of mass. The use of the geometric mean, on the other hand, has been limited to positive numbers and positive integrable functions [13, 7]. In 1975, Anderson and Trapp [2] and Pusz and Woronowicz [22] introduced the harmonic and geometric means for a pair of positive operators on a Hilbert space. Thereafter, an extensive theory on operator means originated. It has been shown that the geometric mean of two positive-definite operators shares many of the properties of the geometric mean of two positive numbers. A recent paper by Lawson and Lim [16] surveys eight shared properties. The geometric mean of positive operators has been used mainly as a binary operation.

In [26], there was a discussion about how to define the geometric mean of more than two Hermitian semidefinite matrices. There have been attempts to use iterative procedures, but none seemed to work when the matrices do not commute. In [1] there is a definition for the geometric mean of a finite set of operators; however, the given definition is not invariant under reordering of the matrices. The present author, while working with means of a finite number of 3-dimensional rotation matrices [18], discovered that there is a close connection between the Riemannian mean of two rotations and the geometric mean of two Hermitian definite matrices. This observation motivated the present work on the generalization of the geometric mean for more than two matrices using metric-based means. In an abstract setting, if $\mathscr{M}$ is a Riemannian manifold with metric $\mathrm{d}(\cdot,\cdot)$, then by analogy to (1.1) and (1.2), a plausible definition of a mean associated with $\mathrm{d}(\cdot,\cdot)$ of m points in $\mathscr{M}$ is given by

$$\mathfrak{M}(x_1,\dots,x_m) := \operatorname*{arg\,min}_{x\in\mathscr{M}} \sum_{k=1}^{m} \mathrm{d}(x_k,x)^2. \tag{1.3}$$

Note that this definition does not guarantee that the mean is unique.

As we have seen, for the set of positive real numbers, which is at the same time a Lie group and an open convex cone,[2] different notions of mean can be associated with different metrics. In what follows, we will extend these metric-based means to the cone of positive-definite transformations. The methods and ideas used in this paper carry over to the complex counterpart of the space considered here, i.e., the convex cone of Hermitian definite transformations. We here concentrate on the real space just for simplicity of exposition but not for any fundamental reason.

The remainder of this paper is organized as follows. In section 2 we gather all the necessary background from differential geometry and optimization on manifolds that will be used throughout the text. Further information on this condensed material can be found in [9, 5, 11, 25, 27]. In section 3 we give a Riemannian metric-based notion of mean for positive-definite matrices. We discuss some invariance properties of this mean and show that in the case where two matrices are to be averaged, this mean coincides with the geometric mean.

2. Preliminaries. Let $\mathcal{M}(n)$ be the set of $n\times n$ real matrices and $GL(n)$ be its subset containing only nonsingular matrices. $GL(n)$ is a Lie group, i.e., a group which is also a differentiable manifold and for which the operations of group multiplication

[2]Here and throughout we use the term *open convex cone*, or simply *cone*, when we really mean the interior of a convex cone.

and inverse are smooth. The tangent space at the identity is called the corresponding Lie algebra and denoted by $\mathfrak{gl}(n)$. It is the space of all linear transformations in $\mathbb{R}^n$, i.e., $\mathcal{M}(n)$.

In $\mathcal{M}(n)$ we shall use the Euclidean inner product, known as the Frobenius inner product and defined by $\langle \boldsymbol{A}, \boldsymbol{B} \rangle_F = \operatorname{tr}(\boldsymbol{A}^T \boldsymbol{B})$, where $\operatorname{tr}(\cdot)$ stands for the trace and the superscript ^{T} denotes the transpose. The associated norm $\|\boldsymbol{A}\|_F = \langle \boldsymbol{A}, \boldsymbol{A} \rangle_F^{1/2}$ is used to define the Euclidean distance on $\mathcal{M}(n)$,

$$\mathrm{d}_F(\boldsymbol{A}, \boldsymbol{B}) = \|\boldsymbol{A} - \boldsymbol{B}\|_F. \tag{2.1}$$

2.1. Exponential and logarithms. The exponential of a matrix in $\mathfrak{gl}(n)$ is given, as usual, by the convergent series

$$\exp \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{A}^k. \tag{2.2}$$

We remark that the product of the exponentials of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is equal to $\exp(\boldsymbol{A} + \boldsymbol{B})$ only when $\boldsymbol{A}$ and $\boldsymbol{B}$ commute.

Logarithms of $\boldsymbol{A}$ in $GL(n)$ are solutions of the matrix equation $\exp \boldsymbol{X} = \boldsymbol{A}$. When $\boldsymbol{A}$ does not have eigenvalues in the (closed) negative real line, there exists a unique real logarithm, called the principal logarithm and denoted by $\operatorname{Log} \boldsymbol{A}$, whose spectrum lies in the infinite strip $\{z \in \mathbb{C} \ : \ -\pi < \operatorname{Im}(z) < \pi\}$ of the complex plane [9]. Furthermore, if for any given matrix norm $\|\cdot\|$ we have $\|\boldsymbol{I} - \boldsymbol{A}\| < 1$, where $\boldsymbol{I}$ denotes the identity transformation in $\mathbb{R}^n$, then the series $-\sum_{k=1}^{\infty} \frac{(\boldsymbol{I}-\boldsymbol{A})^k}{k}$ converges to $\operatorname{Log} \boldsymbol{A}$, and therefore one can write

$$\operatorname{Log} \boldsymbol{A} = -\sum_{k=1}^{\infty} \frac{(\boldsymbol{I} - \boldsymbol{A})^k}{k}. \tag{2.3}$$

We note that, in general, $\operatorname{Log}(\boldsymbol{A}\boldsymbol{B}) \neq \operatorname{Log} \boldsymbol{A} + \operatorname{Log} \boldsymbol{B}$. We here recall the important fact [9]

$$\operatorname{Log}(\boldsymbol{A}^{-1} \boldsymbol{B} \boldsymbol{A}) = \boldsymbol{A}^{-1} (\operatorname{Log} \boldsymbol{B}) \boldsymbol{A}. \tag{2.4}$$

This fact is also true when Log in the above is replaced with an analytic matrix function.

The following result is essential in the development of our analysis.

PROPOSITION 2.1. *Let $\boldsymbol{X}(t)$ be a real matrix-valued function of the real variable t. We assume that, for all t in its domain, $\boldsymbol{X}(t)$ is an invertible matrix which does not have eigenvalues on the closed negative real line. Then*

$$\frac{d}{dt} \operatorname{tr} \left[\operatorname{Log}^2 \boldsymbol{X}(t) \right] = 2 \operatorname{tr} \left[\operatorname{Log} \boldsymbol{X}(t) \boldsymbol{X}^{-1}(t) \frac{d}{dt} \boldsymbol{X}(t) \right].$$

Proof. We recall the following facts:

(i) $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$.

(ii) $\operatorname{tr}(\int_a^b \boldsymbol{M}(s) ds) = \int_a^b \operatorname{tr}(\boldsymbol{M}(s)) ds$.

(iii) $\operatorname{Log} \boldsymbol{A}$ commutes with $[(\boldsymbol{A} - \boldsymbol{I})s + \boldsymbol{I}]^{-1}$.

(iv) $\int_0^1 [(\boldsymbol{A} - \boldsymbol{I})s + \boldsymbol{I}]^{-2} ds = (\boldsymbol{I} - \boldsymbol{A})^{-1} [(\boldsymbol{A} - \boldsymbol{I})s + \boldsymbol{I}]^{-1} |_0^1 = \boldsymbol{A}^{-1}$.

(v) $\frac{d}{dt} \operatorname{Log} \boldsymbol{X}(t) = \int_0^1 [(\boldsymbol{X}(t) - \boldsymbol{I})s + \boldsymbol{I}]^{-1} \frac{d}{dt} \boldsymbol{X}(t) [(\boldsymbol{X}(t) - \boldsymbol{I})s + \boldsymbol{I}]^{-1} ds$.

Facts (i), (ii), (iii), and (iv) are easily checked. See [10] for a proof of (v).

Using the above, we have

$$\begin{aligned}
\frac{d}{dt}\operatorname{tr}\left([\operatorname{Log}\boldsymbol{X}(t)]^2\right) &\overset{\text{(i)}}{=} 2\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\frac{d}{dt}\operatorname{Log}\boldsymbol{X}(t)\right)\\
&\overset{\text{(v)}}{=} 2\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\int_0^1[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\frac{d}{dt}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\,ds\right)\\
&= 2\operatorname{tr}\left(\int_0^1\operatorname{Log}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\frac{d}{dt}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+I]^{-1}\,ds\right)\\
&\overset{\text{(ii)}}{=} 2\int_0^1\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\frac{d}{dt}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\right)ds\\
&\overset{\text{(i)}}{=} 2\int_0^1\operatorname{tr}\left([(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\operatorname{Log}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-1}\frac{d}{dt}\boldsymbol{X}(t)\right)ds\\
&\overset{\text{(iii)}}{=} 2\int_0^1\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\,[(\boldsymbol{X}(t)-\boldsymbol{I})s+\boldsymbol{I}]^{-2}\frac{d}{dt}\boldsymbol{X}(t)\right)ds\\
&= 2\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\int_0^1[(\boldsymbol{X}(t)-\boldsymbol{I})s+I]^{-2}\,ds\frac{d}{dt}\boldsymbol{X}(t)\right)\\
&\overset{\text{(iv)}}{=} 2\operatorname{tr}\left(\operatorname{Log}\boldsymbol{X}(t)\boldsymbol{X}^{-1}(t)\frac{d}{dt}\boldsymbol{X}(t)\right). \qquad\square
\end{aligned}$$

2.2. Gradient and geodesic convexity. For a real-valued function $f(x)$ defined on a Riemannian manifold $\mathscr{M}$, the gradient ∇f is the unique tangent vector u at x such that

$$\langle u, \nabla f\rangle = \left.\frac{d}{dt}f(\gamma(t))\right|_{t=0}, \tag{2.5}$$

where $\gamma(t)$ is a geodesic emanating from x in the direction of u, and $\langle\cdot,\cdot\rangle$ denotes the Riemannian inner product on the tangent space.

A subset $\mathscr{A}$ of a Riemannian manifold $\mathscr{M}$ is said to be convex if the shortest geodesic curve between any two points x and y in $\mathscr{A}$ is unique in $\mathscr{M}$ and lies in $\mathscr{A}$. A real-valued function defined on a convex subset $\mathscr{A}$ of $\mathscr{M}$ is said to be convex if its restriction to any geodesic path is convex, i.e., if $t \mapsto \hat{f}(t) \equiv f(\exp_x(tu))$ is convex over its domain for all $x \in \mathscr{M}$ and $u \in T_x(\mathscr{M})$, where $\exp_x$ is the exponential map at x.

2.3. The cone of the positive-definite symmetric matrices. We denote by

$$\mathcal{S}(n) = \{\boldsymbol{A} \in \mathcal{M}(n),\ \boldsymbol{A}^T = \boldsymbol{A}\}$$

the space of all $n \times n$ symmetric matrices and denote by

$$\mathcal{P}(n) = \{\boldsymbol{A} \in \mathcal{S}(n),\ \boldsymbol{A} > 0\}$$

the set of all $n \times n$ positive-definite symmetric matrices. Here $\boldsymbol{A} > 0$ means that the quadratic form $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ for all $\boldsymbol{x} \in \mathbb{R}^n\backslash\{\boldsymbol{0}\}$. It is well known that $\mathcal{P}(n)$ is an open convex cone; i.e., if $\boldsymbol{P}$ and $\boldsymbol{Q}$ are in $\mathcal{P}(n)$, so is $\boldsymbol{P} + t\boldsymbol{Q}$ for any $t > 0$.

We recall that the exponential map from $\mathcal{S}(n)$ to $\mathcal{P}(n)$ is one-to-one and onto. In other words, the exponential of any symmetric matrix is a positive-definite symmetric matrix, and the inverse of the exponential (i.e., the principal logarithm) of any positive-definite symmetric matrix is a symmetric matrix.

As $\mathcal{P}(n)$ is an open subset of $\mathcal{S}(n)$, for each $\boldsymbol{P} \in \mathcal{P}(n)$ we identify the set $T_{\boldsymbol{P}}$ of tangent vectors to $\mathcal{P}(n)$ at $\boldsymbol{P}$ with $\mathcal{S}(n)$. On the tangent space at $\boldsymbol{P}$ we define the positive-definite inner product and corresponding norm,

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle_{\boldsymbol{P}} = \operatorname{tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{B}), \qquad \|\boldsymbol{A}\|_{\boldsymbol{P}} = \langle \boldsymbol{A}, \boldsymbol{A} \rangle_{\boldsymbol{P}}^{1/2}, \tag{2.6}$$

that depend on the point $\boldsymbol{P}$. The positive definiteness is a consequence of the positive definiteness of the Frobenius inner product for

$$\langle \boldsymbol{A}, \boldsymbol{A} \rangle_{\boldsymbol{P}} = \operatorname{tr}(\boldsymbol{P}^{-1/2}\boldsymbol{A}\boldsymbol{P}^{-1/2}\boldsymbol{P}^{-1/2}\boldsymbol{A}\boldsymbol{P}^{-1/2}) = \left\langle \boldsymbol{P}^{-1/2}\boldsymbol{A}\boldsymbol{P}^{-1/2}, \boldsymbol{P}^{-1/2}\boldsymbol{A}\boldsymbol{P}^{-1/2} \right\rangle.$$

Let $[a, b]$ be a closed interval in $\mathbb{R}$, and let $\boldsymbol{\Gamma} : [a, b] \to \mathcal{P}(n)$ be a sufficiently smooth curve in $\mathcal{P}(n)$. We define the length of $\boldsymbol{\Gamma}$ by

$$\mathcal{L}(\boldsymbol{\Gamma}) := \int_a^b \sqrt{\left\langle \dot{\boldsymbol{\Gamma}}(t), \dot{\boldsymbol{\Gamma}}(t) \right\rangle_{\boldsymbol{\Gamma}(t)}}\, dt = \int_a^b \sqrt{\operatorname{tr}(\boldsymbol{\Gamma}(t)^{-1}\dot{\boldsymbol{\Gamma}}(t))^2}\, dt. \tag{2.7}$$

We note that the length $\mathcal{L}(\boldsymbol{\Gamma})$ is invariant under congruent transformations, i.e., $\boldsymbol{\Gamma} \mapsto \boldsymbol{C}\boldsymbol{\Gamma}\boldsymbol{C}^T$, where $\boldsymbol{C}$ is any fixed element of $\in GL(n)$. As $\frac{d}{dt}\boldsymbol{\Gamma}^{-1} = -\boldsymbol{\Gamma}^{-1}\dot{\boldsymbol{\Gamma}}\boldsymbol{\Gamma}^{-1}$, one can readily see that this length is also invariant under inversion.

The distance between two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in $\mathcal{P}(n)$ considered as a differentiable manifold is the infimum of lengths of curves connecting them:

$$\mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{A}, \boldsymbol{B}) := \inf \left\{ \mathcal{L}(\boldsymbol{\Gamma}) \mid \boldsymbol{\Gamma} : [a, b] \to \mathcal{P}(n) \text{ with } \boldsymbol{\Gamma}(a) = \boldsymbol{A},\ \boldsymbol{\Gamma}(b) = \boldsymbol{B} \right\}. \tag{2.8}$$

This metric makes $\mathcal{P}(n)$ a Riemannian manifold which is of dimension $\frac{1}{2}n(n+1)$. The geodesic emanating from $\boldsymbol{I}$ in the direction of $\boldsymbol{S}$, a (symmetric) matrix in the tangent space, is given explicitly by $e^{t\boldsymbol{S}}$ [17]. Using invariance under congruent transformations, the geodesic $\boldsymbol{P}(t)$ such that $\boldsymbol{P}(0) = \boldsymbol{P}$ and $\dot{\boldsymbol{P}}(0) = \boldsymbol{S}$ is therefore given by

$$\boldsymbol{P}(t) = \boldsymbol{P}^{1/2} e^{t\boldsymbol{P}^{-1/2}\boldsymbol{S}\boldsymbol{P}^{-1/2}} \boldsymbol{P}^{1/2}.$$

It follows that the Riemannian distance of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in $\mathcal{P}(n)$ is

$$\mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \|\operatorname{Log}(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)\|_F = \left[\sum_{i=1}^n \ln^2 \lambda_i\right]^{1/2}, \tag{2.9}$$

where $\lambda_i,\ i = 1, \dots, n$, are the eigenvalues of $\boldsymbol{P}_1^{-1}\boldsymbol{P}_2$. Even though in general $\boldsymbol{P}_1^{-1}\boldsymbol{P}_2$ is not symmetric, its eigenvalues are real and positive. This can be seen by noting that $\boldsymbol{P}_1^{-1}\boldsymbol{P}_2$ is similar to the positive-definite symmetric matrix $\boldsymbol{P}_2^{1/2}\boldsymbol{P}_1^{-1}\boldsymbol{P}_2^{1/2}$. It is important to note here that the real-valued function defined on $\mathcal{P}(n)$ by $\boldsymbol{P} \mapsto \mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}, \boldsymbol{S})$, where $\boldsymbol{S} \in \mathcal{P}(n)$ is fixed, is (geodesically) convex [21].

We note in passing that $\mathcal{P}(n)$ is a homogeneous space of the Lie group $GL(n)$ (by identifying $\mathcal{P}(n)$ with the quotient $GL(n)/O(n)$). It is also a symmetric space of noncompact type [25].

We shall also consider the symmetric space of special positive matrices

$$\mathcal{SP}(n) = \{\boldsymbol{A} \in \mathcal{P}(n),\ \det \boldsymbol{A} = 1\}.$$

This submanifold can also be identified with the quotient $SL(n)/SO(n)$. Here $SL(n)$ denotes the special linear group of all determinant-one matrices in $GL(n)$. We note that $\mathcal{SP}(n)$ is a totally geodesic submanifold of $\mathcal{P}(n)$ [17]. Now since

$$\mathcal{P}(n) = \mathcal{SP}(n) \times \mathbb{R}^+,$$

$\mathcal{P}(n)$ can be seen as a foliated manifold whose codimension-one leaves are isomorphic to the hyperbolic space $\mathbb{H}^p$, where $p = \frac{1}{2}n(n+1) - 1$.

3. Means of positive-definite symmetric matrices. Using definition (1.3) with the two distance functions (2.1) and (2.9), we introduce the two different notions of mean in $\mathcal{P}(n)$.

DEFINITION 3.1. *The mean in the Euclidean sense, i.e., associated with the metric* (2.1)*, of m given positive-definite symmetric matrices $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_m$ is defined as*

$$\mathfrak{A}(\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_m) := \mathop{\arg\min}_{\boldsymbol{P} \in \mathcal{P}(n)} \sum_{k=1}^{m} \|\boldsymbol{P}_k - \boldsymbol{P}\|_F^2. \tag{3.1}$$

DEFINITION 3.2. *The mean in the Riemannian sense, i.e., associated with the metric* (2.9)*, of m given positive-definite symmetric matrices $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_m$ is defined as*

$$\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_m) := \mathop{\arg\min}_{\boldsymbol{P} \in \mathcal{P}(n)} \sum_{k=1}^{m} \|\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P})\|_F^2. \tag{3.2}$$

Before we proceed further, we note that both means satisfy the following desirable properties:

P1. Invariance under reordering: For any permutation σ of the numbers $1, \ldots, m$, we have

$$\mathfrak{M}(\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_m) = \mathfrak{M}(\boldsymbol{P}_{\sigma(1)}, \boldsymbol{P}_{\sigma(2)}, \ldots, \boldsymbol{P}_{\sigma(m)}).$$

P2. Invariance under congruent transformations: If $\boldsymbol{P}$ is the positive-definite symmetric mean of $\{\boldsymbol{P}_k\}_{1\le k\le m}$, then $\boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T$ is the positive-definite symmetric mean of $\{\boldsymbol{C}\boldsymbol{P}_k\boldsymbol{C}^T\}_{1\le k\le m}$ for every $\boldsymbol{C}$ in $GL(n)$. From the special case when $\boldsymbol{C}$ is in the full orthogonal group $O(n)$, we deduce the invariance under orthogonal transformations.

We remark that P2 is the counterpart of the homogeneity property of means of positive numbers (but here left and right multiplication are both needed so that the resultant matrix lies in $\mathcal{P}(n)$). The mean in the Riemannian sense does satisfy the following additional property:

P3. Invariance under inversion: If $\boldsymbol{P}$ is the mean of $\{\boldsymbol{P}_k\}_{1\le k\le m}$, then $\boldsymbol{P}^{-1}$ is the mean of $\{\boldsymbol{P}_k^{-1}\}_{1\le k\le m}$.

The mean in the Euclidean sense does in fact satisfy properties other than P1 and P2; however, they are not relevant for the cone of positive-definite symmetric matrices. Furthermore, the solution of the minimization problem (3.1) is simply given by $\boldsymbol{P} = \frac{1}{m}\sum_{k=1}^{m} \boldsymbol{P}_k$, which is the usual arithmetic mean. Therefore, the mean in the Euclidean sense will not be considered any further.

Remark 3.3. The Riemannian mean of $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ may also be called the *Riemannian barycenter* of $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$, which is a notion introduced by Grove, Karcher, and Ruh [12]. In [15] it was proven that for manifolds with nonpositive sectional curvature, the Riemannian barycenter is unique.

3.1. Characterization of the Riemannian mean. In the following proposition we will give a characterization of the Riemannian mean.

PROPOSITION 3.4. *The Riemannian mean of given m symmetric positive-definite matrices $\boldsymbol{P}_1, \dots, \boldsymbol{P}_m$ is the unique symmetric positive-definite solution to the nonlinear matrix equation*

$$\sum_{k=1}^{m} \operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) = \boldsymbol{0}. \tag{3.3}$$

Proof. First, we compute the derivative of the real-valued function $H(\boldsymbol{S}(t)) = \frac{1}{2}\|\operatorname{Log}(\boldsymbol{W}^{-1}\boldsymbol{S}(t))\|_F^2$ with respect to t, where $\boldsymbol{S}(t) = \boldsymbol{P}^{1/2}\exp(t\boldsymbol{A})\boldsymbol{P}^{1/2}$ is the geodesic emanating from $\boldsymbol{P}$ in the direction of $\boldsymbol{\Delta} = \dot{\boldsymbol{S}}(0) = \boldsymbol{P}^{1/2}\boldsymbol{A}\boldsymbol{P}^{1/2}$, and $\boldsymbol{W}$ is a constant matrix in $\mathcal{P}(n)$.

Using (2.4) and some properties of the trace, it follows that

$$H(\boldsymbol{S}(t)) = \frac{1}{2}\|\operatorname{Log}(\boldsymbol{W}^{-1/2}\boldsymbol{S}(t)\boldsymbol{W}^{-1/2})\|_F^2.$$

Because $\operatorname{Log}(\boldsymbol{W}^{-1/2}\boldsymbol{S}(t)\boldsymbol{W}^{-1/2})$ is symmetric, we have

$$\frac{d}{dt}H(\boldsymbol{S}(t))\bigg|_{t=0} = \frac{1}{2}\,\frac{d}{dt}\operatorname{tr}\left([\operatorname{Log}(\boldsymbol{W}^{-1/2}\boldsymbol{S}(t)\boldsymbol{W}^{-1/2})]^2\right)\bigg|_{t=0}.$$

Therefore, the general result of Proposition 2.1 applied to the above yields

$$\frac{d}{dt}H(\boldsymbol{S}(t))\bigg|_{t=0} = \operatorname{tr}[\operatorname{Log}(\boldsymbol{W}^{-1}\boldsymbol{P})\boldsymbol{P}^{-1}\boldsymbol{\Delta}] = \operatorname{tr}[\boldsymbol{\Delta}\operatorname{Log}(\boldsymbol{W}^{-1}\boldsymbol{P})\boldsymbol{P}^{-1}],$$

and hence the gradient of H is given by

$$\nabla H = \operatorname{Log}(\boldsymbol{W}^{-1}\boldsymbol{P})\boldsymbol{P}^{-1} = \boldsymbol{P}^{-1}\operatorname{Log}(\boldsymbol{P}\boldsymbol{W}^{-1}), \tag{3.4}$$

which is indeed in the tangent space, i.e., in $\mathcal{S}(n)$.

Now, let G denote the objective function of the minimization problem (3.2), i.e.,

$$G(\boldsymbol{P}) = \sum_{k=1}^{m}\|\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P})\|_F^2. \tag{3.5}$$

Using the above, the gradient of G is found to be

$$\nabla G = \boldsymbol{P}\sum_{k=1}^{m}\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}). \tag{3.6}$$

As (3.5) is the sum of convex functions, the necessary condition and sufficient condition for $\boldsymbol{P}$ to be the minimum of (3.5) is the vanishing of the gradient (3.6), or, equivalently,

$$\sum_{k=1}^{m}\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) = \boldsymbol{0}. \qquad \square$$

It is worth noting that the characterization for the Riemannian mean given in (3.3) is similar to the characterization

$$\sum_{k=1}^{m} \ln(x_k^{-1}x) = 0 \tag{3.7}$$

of the geometric mean (1.2) of positive numbers. However, unlike the case of positive numbers, where (3.7) yields to an explicit expression of the geometric mean, in general, due to the noncommutative nature of $\mathcal{P}(n)$, (3.3) cannot be solved in closed form. In the next section we will show that when $m = 2$, (3.3) yields explicit expressions of the Riemannian mean.

3.1.1. Riemannian mean of two positive-definite symmetric matrices. The following proposition shows that for the case $m = 2$, (3.3) can be solved analytically.

PROPOSITION 3.5. *The mean in the Riemannian sense of two positive-definite symmetric matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ is given explicitly by the following six equivalent expressions:*

$$\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2} = \boldsymbol{P}_2(\boldsymbol{P}_2^{-1}\boldsymbol{P}_1)^{1/2} \tag{3.8}$$
$$= (\boldsymbol{P}_2\boldsymbol{P}_1^{-1})^{1/2}\boldsymbol{P}_1 = (\boldsymbol{P}_1\boldsymbol{P}_2^{-1})^{1/2}\boldsymbol{P}_2$$
$$= \boldsymbol{P}_1^{1/2}(\boldsymbol{P}_1^{-1/2}\boldsymbol{P}_2\boldsymbol{P}_1^{-1/2})^{1/2}\boldsymbol{P}_1^{1/2} \tag{3.9}$$
$$= \boldsymbol{P}_2^{1/2}(\boldsymbol{P}_2^{-1/2}\boldsymbol{P}_1\boldsymbol{P}_2^{-1/2})^{1/2}\boldsymbol{P}_2^{1/2}.$$

Proof. First, we rewrite (3.3) as $\operatorname{Log}(\boldsymbol{P}_1^{-1}\boldsymbol{P}) = -\operatorname{Log}(\boldsymbol{P}_2^{-1}\boldsymbol{P})$. Then we take the exponential of both sides to obtain

$$\boldsymbol{P}_1^{-1}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{P}_2. \tag{3.10}$$

After left multiplying both sides with $\boldsymbol{P}_1^{-1}\boldsymbol{P}$ we get $(\boldsymbol{P}_1^{-1}\boldsymbol{P})^2 = \boldsymbol{P}_1^{-1}\boldsymbol{P}_2$. Such a matrix equation has $\boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2}$ as the unique solution in $\mathcal{P}(n)$. Therefore, the mean in the Riemannian sense of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ is given explicitly by

$$\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2}.$$

The second equality in (3.9) can be easily verified by premultiplying $\boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2}$ by $\boldsymbol{P}_2\boldsymbol{P}_2^{-1} = \boldsymbol{I}$. This makes it clear that $\mathfrak{G}$ is symmetric with respect to $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$, i.e., $\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \mathfrak{G}(\boldsymbol{P}_2, \boldsymbol{P}_1)$. The third equality in (3.9) can be obtained by right multiplying both sides of (3.10) by $\boldsymbol{P}^{-1}\boldsymbol{P}_2$ and solving the resultant equation for $\boldsymbol{P}$. The fourth equality in (3.9) can be established from the third by right multiplying $(\boldsymbol{P}_2\boldsymbol{P}_1^{-1})^{1/2}\boldsymbol{P}_1$ by $\boldsymbol{P}_2^{-1}\boldsymbol{P}_2$. Alternatively, these equalities can be verified using (2.4). Furthermore, by use of (2.4) and after some algebra, one can show that the two expressions in (3.10) are but two other alternative forms of the geometric mean of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ that highlight not only its symmetry with respect to $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ but also its symmetry as a matrix. ∎

The explicit equivalent expressions given in (3.9) and (3.10) for the Riemannian mean of a pair of positive-definite matrices that we obtained by solving the minimization problem (3.5) coincide with the different definitions of the geometric mean of a pair of positive Hermitian operators first introduced by Pusz and Woronowicz [22]. This mean arises in electrical network theory as described in the survey paper by Trapp [26]. For this reason and the following proposition, the Riemannian mean will be termed the *geometric mean*.

PROPOSITION 3.6. *If all matrices $\boldsymbol{P}_k$, $k = 1, \dots, m$, belong to a single geodesic curve of $\mathcal{P}(n)$, i.e.,*

$$\boldsymbol{P}_k = \boldsymbol{C} \exp(t_k \boldsymbol{S}) \boldsymbol{C}^T, \quad k = 1, \dots, m,$$

where $\boldsymbol{S} \in \mathcal{S}(n)$, $\boldsymbol{C} \in GL(n)$, and $t_k \in \mathbb{R}$, $k = 1, \dots, m$, then their Riemannian mean is

$$\boldsymbol{P} = \boldsymbol{C} \exp\left(\frac{1}{m} \sum_{k=1}^{m} t_k \boldsymbol{S} \right) \boldsymbol{C}^T.$$

In particular, when $\boldsymbol{C}$ is orthogonal, i.e., such that $\boldsymbol{C}^T\boldsymbol{C} = \boldsymbol{I}$, we have

$$\boldsymbol{P} = (\boldsymbol{P}_1 \cdots \boldsymbol{P}_m)^{1/m} = \boldsymbol{P}_1^{1/m} \cdots \boldsymbol{P}_m^{1/m}.$$

Proof. Straightforward computations show that the given expression for $\boldsymbol{P}$ does satisfy (3.3) characterizing the Riemannian mean. ∎

For more than two matrices, in general, it is not possible to obtain an explicit expression for their geometric mean. In the commutative case of the cone of positive numbers, the problem of finding the geometric mean of three positive numbers can be done by first finding the geometric mean of two numbers and then finding the weighted geometric mean of the latter (with weight 2/3) and the other number (with weight 1/3). This procedure does not depend on the ordering of the numbers. For the space of positive-definite matrices this procedure gives different positive-definite matrices, depending on the way we order the elements. Furthermore, in general, none of these matrices satisfy the characterization (3.3) of the geometric mean. This is due to the fact that the geodesic triangle, whose sides join the three given matrices, is not flat. (See the discussion at the end of section 3.3.)

Thus, it is only in some special cases that one expects to have an explicit formula for the geometric mean. In the following proposition we give an example in which we obtain a closed-form expression of the geometric mean.

PROPOSITION 3.7. *Let $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_3$ be matrices in $\mathcal{P}(n)$ such that $\boldsymbol{P}_1 = r\boldsymbol{P}_2$ for some $r > 0$. Then their geometric mean is given explicitly by*

$$\begin{aligned}\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3) &= r^{1/3} \boldsymbol{P}_3 (\boldsymbol{P}_3^{-1} \boldsymbol{P}_2)^{2/3} = r^{-1/3} \boldsymbol{P}_3 (\boldsymbol{P}_3^{-1} \boldsymbol{P}_1)^{2/3} \\ &= r^{1/3} \boldsymbol{P}_2 (\boldsymbol{P}_2^{-1} \boldsymbol{P}_3)^{1/3} = r^{-1/3} \boldsymbol{P}_1 (\boldsymbol{P}_1^{-1} \boldsymbol{P}_3)^{1/3}.\end{aligned}$$

This proposition can easily be checked by straightforward calculations.

3.2. Reduction to the space of special positive-definite symmetric matrices. In this section we will show that the problem of finding the geometric mean of positive-definite matrices can be reduced to that of finding the geometric mean of special positive-definite matrices.

LEMMA 3.8. *If $\boldsymbol{P}$ is the geometric mean of m positive-definite symmetric matrices $\boldsymbol{P}_1, \dots, \boldsymbol{P}_m$, then for any m-tuple $(a_1, \dots, a_m)$ in the positive orthant of $\mathbb{R}^m$, the positive-definite symmetric matrix $\sqrt[m]{a_1 \cdots a_m}\boldsymbol{P}$ is the geometric mean of $a_1\boldsymbol{P}_1, \dots, a_m\boldsymbol{P}_m$.*

Proof. We have

$$\operatorname{Log}\left[(a_k \boldsymbol{P}_k)^{-1} \sqrt[m]{a_1 \cdots a_m} \boldsymbol{P}\right] = \operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) + \left[\frac{\ln a_1 + \cdots + \ln a_m}{m} - \ln a_k\right] \boldsymbol{I}.$$

Therefore,

$$\sum_{k=1}^{m} \operatorname{Log}\left[(a_k \boldsymbol{P}_k)^{-1} \sqrt[m]{a_1 \cdots a_m} \boldsymbol{P}\right] = \sum_{k=1}^{m} \operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) = \boldsymbol{0},$$

and hence $\sqrt[m]{a_1 \cdots a_m}\boldsymbol{P}$ is the geometric mean of $a_1\boldsymbol{P}_1, \ldots, a_m\boldsymbol{P}_m$. ∎

LEMMA 3.9. *If $\boldsymbol{P}$ and $\boldsymbol{Q}$ are in $\mathcal{SP}(n)$, i.e., are positive-definite symmetric matrices of determinant one, then for any $\alpha > 0$ we have $\mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}, \boldsymbol{Q}) \leq \mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}, \alpha\boldsymbol{Q})$, and equality holds if and only if $\alpha = 1$.*

Proof. This lemma follows immediately from the fact that $\mathcal{SP}(n)$ is a totally geodesic submanifold of $\mathcal{P}(n)$. Here is an alternative proof. Let $0 < \lambda_i,\ i = 1, \ldots, n$, be the eigenvalues of $\boldsymbol{P}^{-1}\boldsymbol{Q}$. Then

$$\mathrm{d}^2_{\mathcal{P}(n)}(\boldsymbol{P}, \alpha\boldsymbol{Q}) = \sum_{i=1}^{n} \ln^2(\alpha\lambda_i) = \sum_{i=1}^{n} \ln^2 \lambda_i + 2\ln\alpha \sum_{i=1}^{n} \ln\lambda_i + n\ln^2\alpha.$$

But as $\boldsymbol{P}$ and $\boldsymbol{Q}$ have determinant one, it follows that $\sum_{i=1}^{n} \ln\lambda_i = 0$, and hence

$$\mathrm{d}^2_{\mathcal{P}(n)}(\boldsymbol{P}, \alpha\boldsymbol{Q}) = \mathrm{d}^2_{\mathcal{P}(n)}(\boldsymbol{P}, \boldsymbol{Q}) + n\ln^2\alpha.$$

Therefore $\mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}, \boldsymbol{Q}) \leq \mathrm{d}_{\mathcal{P}(n)}(\boldsymbol{P}, \alpha\boldsymbol{Q})$, where the equality holds only when $\alpha = 1$. ∎

PROPOSITION 3.10. *Given m positive-definite symmetric matrices $\{\boldsymbol{P}_k\}_{1\leq k\leq m}$, set $\alpha_k = \sqrt[n]{\det \boldsymbol{P}_k}$ and $\widetilde{\boldsymbol{P}}_k = \boldsymbol{P}_k/\alpha_k$. Then the geometric mean of $\{\boldsymbol{P}_k\}_{1\leq k\leq m}$ is the geometric mean of $\{\alpha_k\}_{1\leq k\leq m}$ multiplied by the geometric mean of $\{\widetilde{\boldsymbol{P}}_k\}_{1\leq k\leq m}$, i.e.,*

$$\mathfrak{G}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) = \sqrt[m]{\alpha_1 \cdots \alpha_m}\,\mathfrak{G}(\widetilde{\boldsymbol{P}}_1, \ldots, \widetilde{\boldsymbol{P}}_m).$$

The proof of this proposition is given by the combination of the results of the two previous lemmas.

3.3. Geometric mean of 2 × 2 positive-definite matrices. We start with the following geometric characterization of the geometric mean of two positive-definite matrices of determinant one.

PROPOSITION 3.11. *The geometric mean of two positive-definite symmetric matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in $\mathcal{SP}(2)$ is given by*

$$\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \frac{\boldsymbol{P}_1 + \boldsymbol{P}_2}{\sqrt{\det(\boldsymbol{P}_1 + \boldsymbol{P}_2)}}.$$

Proof. Let $\boldsymbol{X} = (\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2}$. Note that $\det \boldsymbol{X} = 1$ and that the two eigenvalues λ and $1/\lambda$ of $\boldsymbol{X}$ are positive. By the Cayley–Hamilton theorem we have

$$\boldsymbol{X}^2 - \operatorname{tr}(\boldsymbol{X})\boldsymbol{X} + \boldsymbol{I} = \boldsymbol{0},$$

which after premultiplication by $\boldsymbol{P}_1$ and rearrangement is written as

$$\operatorname{tr}(\boldsymbol{X})\boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2} = \boldsymbol{P}_1 + \boldsymbol{P}_2.$$

But $\operatorname{tr}\boldsymbol{X} = \lambda + 1/\lambda$ and $\det(\boldsymbol{P}_1 + \boldsymbol{P}_2) = \det \boldsymbol{P}_1 \det(\boldsymbol{I} + \boldsymbol{X}^2) = (1 + \lambda^2)(1 + 1/\lambda^2) = (\lambda + 1/\lambda)^2$. Therefore,

$$\boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2} = \frac{\boldsymbol{P}_1 + \boldsymbol{P}_2}{\sqrt{\det(\boldsymbol{P}_1 + \boldsymbol{P}_2)}}. \qquad ∎$$

This proposition gives a nice geometric characterization of the geometric mean of two positive-definite matrices in $\mathcal{SP}(2)$. It is given by the intersection of $\mathcal{SP}(2)$ with the ray joining the arithmetic average $\frac{1}{2}(\boldsymbol{P}_1 + \boldsymbol{P}_2)$ and the apex of the cone, i.e., the zero matrix.

By Lemma 3.8 and the above proposition, we have the following result giving an alternative expression for the geometric mean of two matrices in $\mathcal{P}(2)$ that does not require the evaluation of a matrix square root.

Corollary 3.12. *The geometric mean of two positive-definite symmetric matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in $\mathcal{P}(2)$ is given by*

$$\mathfrak{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \sqrt{\alpha_1 \alpha_2} \frac{\sqrt{\alpha_2}\boldsymbol{P}_1 + \sqrt{\alpha_1}\boldsymbol{P}_2}{\sqrt{\det(\sqrt{\alpha_2}\boldsymbol{P}_1 + \sqrt{\alpha_1}\boldsymbol{P}_2)}},$$

where α_1 and α_2 are the determinants of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$, respectively.

Unfortunately, this elegant characterization of the geometric mean for two matrices in $\mathcal{SP}(2)$ cannot be generalized for more than two matrices in $\mathcal{SP}(2)$ nor for two matrices in $\mathcal{SP}(n)$ with $n > 2$. Indeed, this characterization, in general, does not hold for the mentioned cases.

Let us now consider in some detail the geometry of the space of 2×2 positive-definite matrices

$$\mathcal{P}(2) = \left\{ \begin{bmatrix} a & c \\ c & b \end{bmatrix} \ : \ a > 0, \ b > 0, \text{ and } ab > c^2 \right\}.$$

Set $0 < u = \frac{a+b}{2}$, $v = \frac{a-b}{2}$. Then the condition $ab > c^2$ can be rewritten as

$$\sqrt{v^2 + c^2} < u,$$

and therefore $\mathcal{P}(2)$ parameterized by u, v, and c can be seen as an open convex second-order cone (ice cream cone or future-light cone). Furthermore, the determinant-one condition $ab - c^2 = 1$ can be formulated as $u^2 - (v^2 + c^2) = 1$, and hence by using the identity $\cosh^2 \alpha - \sinh^2 \alpha = 1$, a matrix $\boldsymbol{P} = \left[\begin{smallmatrix} a & c \\ c & b \end{smallmatrix}\right]$ in $\mathcal{SP}(2)$ can be parameterized by $(\alpha, \theta) \in \mathbb{R} \times [0, \pi)$ as

$$\boldsymbol{P} = \cosh \alpha \boldsymbol{I} + \cos \theta \sinh \alpha \boldsymbol{J}_1 + \sin \theta \sinh \alpha \boldsymbol{J}_2,$$

where

$$\boldsymbol{J}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \boldsymbol{J}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \cosh \alpha = \frac{a+b}{2}, \qquad \tan \theta = \frac{2c}{a-b}.$$

Note that $(\boldsymbol{I}, \ \boldsymbol{J}_1, \ \boldsymbol{J}_2)$ is a basis for the space $\mathcal{S}(2)$ of 2×2 symmetric matrices. Thus, we see that $\mathcal{SP}(2)$ is isomorphic to the hyperboloid $\mathbb{H}^2$. In particular, geodesics on $\mathcal{SP}(2)$ correspond to geodesics on $\mathbb{H}^2$.

For more than two matrices in $\mathcal{SP}(2)$, in general, it is not possible to obtain their geometric mean in closed form. Nevertheless, using the isomorphism between $\mathcal{SP}(2)$ and $\mathbb{H}^2$, we can identify the geometric mean with the hyperbolic centroid of point masses in the hyperboloid. In particular, the geometric mean of three matrices in $\mathcal{SP}(2)$ corresponds to the (hyperbolic) center of the hyperbolic triangles associated with the given three matrices. This center is the point of intersection of the three medians; see Figure 1. However, unlike Euclidean geometry, the ratio of the geodesic length between a vertex and the center to the length between this vertex and the midpoint on the opposite side is not $2/3$ in general [6].

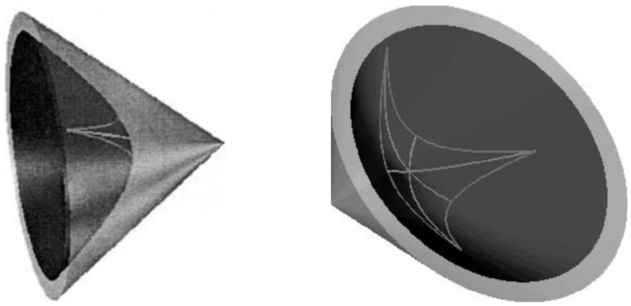

FIG. 1. *Two views for the representation, in the space parameterized by u, v, and c, of the cone $\mathcal{P}(2)$ and the hyperboloid $\mathcal{SP}(2)$. For the geodesic triangle shown, the point of concurrence of the three medians corresponds to the geometric mean of the positive-definite symmetric matrices in $\mathcal{P}(2)$ associated with the vertices of this triangle.*

4. Conclusion. Using the Riemannian metric on the space of positive-definite matrices, we defined the geometric mean. This mean satisfies some invariance properties. Some of these properties are related to the geodesic reversing isometry in the symmetric space considered here. Therefore, the notion of geometric mean, studied here and which is based on the Riemannian metric, can be used to define the geometric mean on other symmetric spaces which enjoy similar invariance properties. Here we used the space of positive-definite matrices as a prototype of a symmetric spaces of noncompact type. The case of the geometric mean of matrices in the group of special orthogonal matrices, which was studied in [18], is a prototype of symmetric spaces of compact type.

Equation (3.3) characterizing this mean is similar to (3.7) characterizing the geometric mean of positive numbers. Unfortunately, due to the noncommutative nature of the matrix multiplication, in general, it is not possible to obtain the geometric mean in closed form for more than two matrices.

Applications of the geometric mean to the problem of averaging data of anisotropic symmetric positive-definite tensors, such as in elasticity theory [8] and in diffusion tensor magnetic resonance imaging [3], are discussed in [19, 20, 4]. In [19], further invariance properties of the geometric mean are discussed and a fixed-point algorithm for solving the nonlinear matrix equation for the geometric mean of more than two matrices is presented.

REFERENCES

[1] M. Alić, B. Mond, J. Pečarić, and V. Volenec, *Bounds for the differences of matrix means*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 119–123.

[2] W. N. Anderson, Jr., and G. E. Trapp, *Shorted operators*. II, SIAM J. Appl. Math., 28 (1975), pp. 60–71.

[3] P. J. Basser, J. Matiello, and D. Le Bihan, *MR diffusion tensor spectroscopy and imaging*, Biophys. J., 66 (1994), pp. 259–267.

[4] P. G. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly, *A rigorous framework for diffusion tensor calculus*, Magn. Reson. Med., 53 (2005), pp. 221–225.

[5] M. Berger and B. Gostiaux, *Differential Geometry: Manifolds, Curves, and Surfaces*, Springer-Verlag, New York, 1988.

[6] O. Bottema, *On the medians of a triangle in hyperbolic geometry*, Canad. J. Math., 10 (1958), pp. 502–506.

[7] P. S. Bullen, D. S. Mitrinović, and P. M. Vasić, *Means and Their Inequalities*, Mathematics and Its Applications (East European Series) 31, D. Reidel Publishing Co., Dordrecht, The Netherlands, 1988.

[8] S. C. Cowin and G. Yang, *Averaging anisotropic elastic constant data*, J. Elasticity, 46 (1997), pp. 151–180.
[9] M. L. Curtis, *Matrix Groups*, Springer-Verlag, New York, Heidelberg, 1979.
[10] L. Dieci, B. Morini, and A. Papini, *Computational techniques for real logarithms of matrices*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 570–593.
[11] P. B. Eblerlein, *Goemetry of Nonpositively Curved Manifolds*, The University of Chicago Press, Chicago, 1996.
[12] K. Grove, H. Karcher, and E. A. Ruh, *Jacobi fields and Finsler metrics on compact Lie groups with an application to differentiable pinching problems*, Math. Ann., 211 (1974), pp. 7–21.
[13] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, UK, 1934.
[14] T. Heath, *A History of Greek Mathematics, Vol.* 1: *From Thales to Euclid*, Dover, New York, 1981.
[15] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math., 30 (1977), pp. 509–541.
[16] J. D. Lawson and Y. Lim, *The geometric mean, matrices, metrics, and more*, Amer. Math. Monthly, 108 (2001), pp. 797–812.
[17] H. Maass, *Siegel's Modular Forms and Dirichlet Series*, Lecture Notes in Math. 216, Springer-Verlag, Heidelberg, 1971.
[18] M. Moakher, *Means and averaging in the group of rotations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 1–16.
[19] M. Moakher, *On averaging symmetric positive-definite tensors*, J. Elasticity, submitted.
[20] M. Moakher and P. G. Batchelor, *The symmetric space of positive-definite tensors: From geometry to applications and visualization*, in Visualization and Image Processing of Tensor Fields, J. Weickert and H. Hagen, eds., Springer, Berlin, 2005, to appear.
[21] G. D. Mostow, *Strong Rigidity of Locally Symmetric Spaces*, Ann. Math. Stud. 78, Princeton University Press, Princeton, NJ, 1973.
[22] W. Pusz and S. L. Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys., 8 (1975), pp. 159–170.
[23] M. Spiesser, *Les médiétés dans la pensée grecque*, in Musique et Mathématiques, Sci. Tech. Perspect. 23, Université de Nantes, Nantes, France, 1993, pp. 1–71.
[24] S. Stahl, *The Poincaré Half-Plane. A Gateway to Modern Geometry*, Jones and Bartlett, Boston, 1993.
[25] A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications* II, Springer-Verlag, New York, 1988.
[26] G. E. Trapp, *Hermitian semidefinite matrix means and related matrix inequalities—an introduction*, Linear and Multilinear Algebra, 16 (1984), pp. 113–123.
[27] C. Udrişte, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Math. Appl. 297, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.