Riemannian metrics on positive definite matrices related to means. II

Fumio Hiai[1] and Dénes Petz[2]

[1] Graduate School of Information Sciences, Tohoku University
Aoba-ku, Sendai 980-8579, Japan

[2] Alfréd Rényi Institute of Mathematics,
H-1364 Budapest, POB 127, Hungary

Abstract

On the manifold of positive definite matrices, a Riemannian metric K^ϕ is associated with a positive kernel function ϕ on $(0,\infty)\times(0,\infty)$ by defining $K_D^\phi(H,K)=\sum_{i,j}\phi(\lambda_i,\lambda_j)^{-1}\mathrm{Tr}\,P_iHP_jK$, where D is a foot point with the spectral decomposition $D=\sum_i\lambda_iP_i$ and H,K are Hermitian matrices (tangent vectors). We are concerned with the case $\phi(x,y)=M(x,y)^\theta$ where $M(x,y)$ is a mean of scalars $x,y>0$. We clarify the isometric structure among such kernel metrics and discuss the convergence properties of geodesic distances and geodesic shortest curves along each isometric line of metrics. The metric corresponding to the square of the logarithmic mean shows up as the attractor of the whole metrics concerned.

AMS classification: 15A45; 15A48; 53B21; 53C22

Keywords: Positive definite matrix, Riemannian metric, Geodesic shortest curve, Symmetric homogeneous mean, Geodesic distance, Geometric mean, Logarithmic mean

[1]E-mail: hiai@math.is.tohoku.ac.jp
[2]E-mail: petz@math.bme.hu

Introduction

The $n \times n$ positive definite complex matrices $\mathbb{P}_n$ form an open subset of the space $\mathbb{H}_n$ of $n \times n$ Hermitian matrices regarded as the Euclidean space of dimension n^2. In fact, $\mathbb{H}_n$ is a real subspace of the Hilbert space $\mathbb{M}_n$ consisting of $n \times n$ complex matrices with the Hilbert-Schmidt inner product $\langle X, Y\rangle_{\mathrm{HS}} := \mathrm{Tr}\, X^*Y$ and the Hilbert-Schmidt norm $\|X\|_{\mathrm{HS}}$ for $X, Y \in \mathbb{M}_n$. Consequently, $\mathbb{P}_n$ naturally has a smooth Riemannian manifold structure so that the tangent space at any foot point is identified with $\mathbb{H}_n$. A Riemannian metric $K_D(H, K)$ is a family of inner products on $\mathbb{H}_n$ depending smoothly on the foot point D. When $\phi(x, y)$ is a smooth positive kernel function on $(0, \infty) \times (0, \infty)$, one can define a Riemannian metric K^ϕ on $\mathbb{P}_n$ by

$$K_D^\phi(H, K) := \sum_{i,j=1}^{k} \phi(\lambda_i, \lambda_j)^{-1} \mathrm{Tr}\, P_i H P_j K, \qquad D \in \mathbb{P}_n,\ H, K \in \mathbb{H}_n, \tag{0.1}$$

where D has the spectral decomposition $\sum_{i=1}^{k} \lambda_i P_i$.

The above type of Riemannian metrics on the positive definite matrices have been studied by many authors. The most studied Riemannian metric on $\mathbb{P}_n$ is the case where the kernel function is $\phi(x, y) := xy$, so the metric is given as

$$K_D(H, K) := \mathrm{Tr}\, D^{-1} H D^{-1} K.$$

This appeared as a geometry of multivariate Gaussian distributions [22, 18] and has played a significant role in the recent development of the geometric matrix mean [14, 16, 5, 4]. In fact, the unique geodesic shortest curve joining $A, B \in \mathbb{P}_n$ with respect to this metric is

$$\gamma(t) = A \,\#_t\, B := A^{1/2}(A^{-1/2} B A^{-1/2})^t A^{1/2}, \qquad 0 \le t \le 1, \tag{0.2}$$

and the geodesic midpoint $\gamma(1/2)$ is exactly the *geometric mean* [21, 1]

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

A powerful idea in matrix means is to introduce a matrix mean as the geodesic midpoint of $A, B \in \mathbb{P}_n$ whenever the geodesic shortest curve jointing A, B exists and is unique. The idea further extends to a matrix mean of several variables $A_1, \dots, A_k \in \mathbb{P}_n$ by taking the minimizer (the geodesic barycenter) of $\sum_{i=1}^{k} \delta^2(A, A_i)$, the square sum of the geodesic distances from A_i. In particular, the monotonicity property of the several variable geometric mean has recently been established in [15, 6] via probabilistic methods.

A distinguished class of Riemannian metrics on $\mathbb{P}_n$ (or rather restricted on the set of $n \times n$ density matrices $\mathcal{D}_n := \{D \in \mathbb{P}_n : \mathrm{Tr}\, D = 1\}$) is the *monotone metrics* due to [20]. This is the case where the kernel functions ϕ is given as $\phi(x, y) := f(x/y)y$ for a symmetric operator monotone function $f : (0, \infty) \to (0, \infty)$, where symmetry of f means that $xf(x^{-1}) = f(x)$, $x > 0$, so the metric is given as

$$K_D^f(H, K) := \langle H, f(\mathbf{L}_D \mathbf{R}_D^{-1}) \mathbf{R}_D K\rangle_{\mathrm{HS}},$$

where $\mathbf{L}_D$ and $\mathbf{R}_D$ are the left multiplication and the right multiplication operators. In fact, it was proved in [20] that the Riemannian metrics on $\mathcal{D}_n$, $n \in \mathbb{N}$, determined by operator monotone functions as above are characterized by the monotonicity under completely positive and trace preserving maps $\beta : \mathbb{M}_n \to \mathbb{M}_m$ in such a way that

$$K_{\beta(D)}(\beta(H), \beta(H)) \le K_D(H, H), \qquad D \in \mathcal{D}_n,\ H \in \mathbb{H}_n.$$

A monotone metric K^f on $\mathcal{D}_n$ is often called a *quantum Fisher information* and it plays an essential role in quantum information geometry.

Including the above two important cases, most Riemannian metrics on $\mathbb{P}_n$ or $\mathcal{D}_n$ appearing in matrix analysis and quantum information are among kernel metrics K^ϕ given in (0.1) whose kernel functions ϕ are given as powers of certain mean functions of two positive scalars. From this point of view, in the previous paper [10] we attempted to develop a rather general theory of Riemannian metrics on $\mathbb{P}_n$ of the form (0.1) with $\phi(x, y) = M(x, y)^\theta$ where $M(x, y)$ is a symmetric homogeneous mean of $x, y > 0$ and θ is a real constant. In the paper we found two special isometric families of Riemannian metrics, from which the following two parametrized families of curves showed up as geodesic shortest curves for corresponding metrics:

$$\mu_\alpha(t) = \big((1-t)A^\alpha + tB^\alpha\big)^{1/\alpha}, \quad 0 \le t \le 1, \quad \alpha \ne 0, \tag{0.3}$$

$$\gamma_\kappa(t) = (A^\kappa \#_t B^\kappa)^{1/\kappa} = \big(A^{\kappa/2}(A^{-\kappa/2}B^\kappa A^{-\kappa/2})^t A^{\kappa/2}\big)^{1/\kappa}, \quad 0 \le t \le 1, \quad \kappa > 0. \tag{0.4}$$

Moreover, we discussed in [10] the comparison of geodesic distances for two metrics in terms of the corresponding means and the degrees of power.

The present paper is a continuation of the previous [10]. The stress here is on the existence of geodesic shortest curves and their convergence properties for such Riemannian metrics on $\mathbb{P}_n$ as introduced above. In Section 1 we prove the existence of a geodesic shortest curve joining $A, B \in \mathbb{P}_n$ with respect to the metric K^ϕ with $\phi(x, y) = M(x, y)^\theta$ under a certain assumption on A, B and θ. In particular, a geodesic shortest curve joining A, B exists if A, B are commuting or if θ is sufficiently near 2. In Section 2 we completely characterize when the functional calculus $D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{P}_n$ by a C^∞ homeomorphism F from $(0, \infty)$ onto itself gives an isometric transformation between two Riemannian manifolds $\mathbb{P}_n$ with a scalar multiple of K^ϕ and $\mathbb{P}_n$ with K^ψ, where $\phi = M^\theta$ and $\psi = N^\tau$ as above. From the characterization it turns out that, for any metric K^ψ with $\psi = N^\tau$, we have a one-parameter isometric family of metrics starting from K^ψ and converging to a distinguished metric K^* corresponding to the square of the logarithmic mean. There are two different types for the isometric families depending on $\tau \ne 2$ or $\tau = 2$, and the two isometric families obtained in [10] are special examples of each type. In Section 3 we prove that the geodesic distances and geodesic curves of metrics in each one-parameter isometric family converge along the isometric line to those of the limiting metric K^*. The convergences

$$\lim_{\alpha \to 0} \mu_\alpha(t) = \lim_{\kappa \searrow 0} \gamma_\kappa(t) = \gamma_*(t) := \exp((1-t)\log A + t \log B) \tag{0.5}$$

for the families of quasi-arithmetic mean curves (0.3) and of quasi-geometric mean curves (0.4) are two particular cases of the general convergence result. The limiting curve γ_* is the geodesic shortest curve for K^*. In this way, we obtain the Riemannian geometric interpretation of the fact that the curve γ_* shows up as the common limit of various curves in $\mathbb{P}_n$ related to means, which may be considered as a kind of universal Lie-Trotter formula. This is somewhat the same flavor of finding matrix means approaching to the Riemannian barycenter (the geometric mean) discussed in [15, 6]. Moreover, the midpoint $\gamma_*(1/2)$ in the limit (0.5) is sometimes called the *Log-Euclidean mean* of A, B, and its various properties are found in [3].

1 Riemannian metrics induced by mean functions

For each $D \in \mathbb{P}_n$ the *left multiplication* and the *right multiplication* operators $\mathbf{L}_D$ and $\mathbf{R}_D$ are defined as $\mathbf{L}_D X := DX$ and $\mathbf{R}_D X := XD$ for $X \in \mathbb{M}_n$, which are commuting positive operators on the Hilbert space $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ with the Hilbert-Schmidt inner product $\langle X, Y\rangle_{\mathrm{HS}} := \mathrm{Tr}\, X^*Y$ and the Hilbert-Schmidt norm $\|X\|_{\mathrm{HS}} := (\mathrm{Tr}\, X^*X)^{1/2}$. For a kernel function $\phi : (0,\infty) \times (0,\infty) \to (0,\infty)$, a positive operator $\phi(\mathbf{L}_D, \mathbf{R}_D)$ on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ is associated to each $D \in \mathbb{P}_n$ with the spectral decomposition $D = \sum_{i=1}^k \lambda_i P_i$ by

$$\phi(\mathbf{L}_D, \mathbf{R}_D)X := \sum_{i,j=1}^k \phi(\lambda_i, \lambda_j) P_i X P_j, \qquad X \in \mathbb{M}_n.$$

When ϕ is symmetric, i.e., $\phi(x, y) = \phi(y, x)$ and $\phi(x, y)$ is smooth in x and y, a Riemannian metric K^ϕ on the manifold $\mathbb{P}_n$ defined by (0.1) is

$$K_D^\phi(H, K) := \left\langle H, \phi(\mathbf{L}_D, \mathbf{R}_D)^{-1} K \right\rangle_{\mathrm{HS}} = \sum_{i,j=1}^k \phi(\lambda_i, \lambda_j)^{-1} \mathrm{Tr}\, P_i H P_j K$$

for $D \in \mathbb{P}_n$ and $H, K \in \mathbb{H}_n$. With the *Schur* (or *Hadamard*) *product* $\circ$ one can also write

$$\phi(\mathbf{L}_D, \mathbf{R}_D)^{-1/2} H = U \left(\left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}} \right]_{ij} \circ (U^* H U) \right) U^*,$$

$$K_D^\phi(H, H) = \left\| \phi(\mathbf{L}_D, \mathbf{R}_D)^{-1/2} H \right\|_{\mathrm{HS}}^2 = \left\| \left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}} \right]_{ij} \circ (U^* H U) \right\|_{\mathrm{HS}}^2, \tag{1.1}$$

where $D = U \mathrm{Diag}(\lambda_1, \dots, \lambda_n) U^*$ is the diagonalization with a unitary U.

When $\gamma : [0, 1] \to \mathbb{P}_n$ is a C^1 curve (or more generally, a continuous and piecewise C^1 curve), the *length* of γ with respect to the above metric K^ϕ is given by

$$L_\phi(\gamma) := \int_0^1 \sqrt{K_{\gamma(t)}^\phi(\gamma'(t), \gamma'(t))}\, dt = \int_0^1 \left\| \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2} \gamma'(t) \right\|_{\mathrm{HS}} dt, \tag{1.2}$$

which is independent of the choice of the parametrization of γ. The *geodesic distance* between $A, B \in \mathbb{P}_n$ with respect to K^ϕ is given by

$$\delta_\phi(A, B) := \inf \left\{ L_\phi(\gamma) : \gamma \text{ is a } C^1 \text{ curve from } A \text{ to } B \right\}. \tag{1.3}$$

The definition is equivalent when the infimum is taken over all smooth curves γ from A to B. A C^1 curve γ from A to B is called a *geodesic shortest curve* (or a *segment*) joining A, B with respect to K^ϕ if it satisfies $L_\phi(\gamma) = \delta_\phi(A, B)$. For a geodesic shortest curve γ, one can choose a smooth parametrization; in fact, the parametrization of constant speed is smooth (see [12, Sect. IV.3], [19, Sect. 5.6]).

In this paper it is essential to deal with absolutely continuous curves. Recall that a curve $\gamma : [0, 1] \to \mathbb{P}_n$ is *absolutely continuous* if $\gamma(t)$ is differentiable for a.e. $t \in [0, 1]$ (with respect to the Lebesgue measure) and $\gamma'(t)$ is integrable in the sense that $\int_0^1 \|\gamma'(t)\|_{\mathrm{HS}}\, dt < +\infty$. The length of such a curve γ can be also defined by (1.2) and is finite (see [19, p. 121] for example). Moreover, note that the definition of $\delta_\phi(A, B)$ in (1.3) is equivalent even when the infimum is taken over all absolutely continuous curves γ from A to B.

A function $M : (0, \infty) \times (0, \infty) \to (0, \infty)$ is said to be a *symmetric homogeneous mean* if, for every $x, y > 0$,

- $M(x, y) = M(y, x)$,
- $M(\alpha x, \alpha y) = \alpha M(x, y)$ for all $\alpha > 0$,
- $M(x, y)$ is non-decreasing in x, y,
- $\min\{x, y\} \le M(x, y) \le \max\{x, y\}$.

We denote by $\mathfrak{M}_0$ the set of all symmetric homogeneous means which are smooth, i.e., $M(x, 1)$ is smooth in $x > 0$. In this paper, as in the previous paper [10], we are concerned with the Riemannian metrics K^ϕ, where ϕ is a power of an $M \in \mathfrak{M}_0$ with degree $\theta \in \mathbb{R}$, i.e., $\phi(x, y) := M(x, y)^\theta$.

In the rest of this section we discuss the geodesic distance $\delta_\phi(A, B)$ and geodesic shortest curves for metrics K^ϕ given above. First, recall [10, Theorem 3.1] that, for $\phi(x, y) := M(x, y)^\theta$ with $M \in \mathfrak{M}_0$ and $\theta \in \mathbb{R}$, the Riemannian manifold $(\mathbb{P}_n, K^\phi)$ is complete (i.e., the metric $\delta_\phi(\cdot, \cdot)$ is complete) if and only if $\theta = 2$. Hence, in the case $\theta = 2$, for any $A, B \in \mathbb{P}_n$ there is a geodesic shortest curve joining A, B with respect to K^ϕ. Although we have no strong evidence, it may be conjectured that a geodesic shortest curve joining A, B exists and is even unique (up to parametrization) for every Riemannian metric K^ϕ as above and for every $A, B \in \mathbb{P}_n$. However, our knowledge in this direction is quite limited at the moment. The unique existence of a geodesic shortest curve is known in [10, Theorem 4.10] only when A, B are commuting, $\theta = 1$, and moreover M is an operator monotone mean (i.e., $M(x, 1)$ is an operator monotone function) different from the arithmetic mean. For the existence of a geodesic shortest curve we have the following two results; both situations are still rather restricted. The first result is an extension of [10, Theorem 4.10] except the uniqueness assertion mentioned just above.

From now on we will use the re-parametrization $\alpha = (2 - \theta)/2$ of the degree parameter $\theta \in \mathbb{R}$, which makes the expression of geodesic distances and geodesic shortest curves simpler.

Theorem 1.1. *Let $\phi(x,y) := M(x,y)^{2(1-\alpha)}$ with $M \in \mathfrak{M}_0$ and $\alpha \in \mathbb{R}$. If $A, B \in \mathbb{P}_n$ are commuting, then*

$$\delta_\phi(A,B) = \begin{cases} \frac{1}{|\alpha|}\|A^\alpha - B^\alpha\|_{\mathrm{HS}} & \text{if } \alpha \neq 0, \\ \|\log A - \log\|_{\mathrm{HS}} & \text{if } \alpha = 0, \end{cases}$$

and a geodesic shortest curve from A to B is given by

$$\gamma_{A,B}(t) := \begin{cases} ((1-t)A^\alpha + tB^\alpha)^{1/\alpha}, & 0 \le t \le 1 \quad \text{if } \alpha \neq 0, \\ \exp((1-t)\log A + t\log B), & 0 \le t \le 1 \quad \text{if } \alpha = 0, \end{cases}$$

independently of the choice of $M \in \mathfrak{M}_0$.

Proof. We prove only the case $\alpha \neq 0$ since the proof in the case $\alpha = 0$ is essentially same. We may assume that A, B are diagonal matrices. Let $\gamma : [0,1] \to \mathbb{P}_n$ be a C^1 curve from A to B. Then one can fix diagonalization $\gamma(t) = U(t)\mathrm{Diag}(\lambda_1(t),\dots,\lambda_n(t))U(t)^*$, $0 \le t \le 1$, in such a way that the eigenvalues $\lambda_1(t) \le \cdots \le \lambda_n(t)$ and unitary matrices $U(t)$ are C^1 except for branching points (at most countable) of the eigenvalues (see [11]). For each t except such branching points, as in the proof of [10, Lemma 3.2], we have

$$\begin{aligned} &\left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2}\gamma'(t)\right\|_{\mathrm{HS}} \\ &\quad = \left\| \left[\frac{1}{M(\lambda_i(t),\lambda_j(t))^{1-\alpha}}\right]_{ij} \circ (U(t)^*\gamma'(t)U(t)) \right\|_{\mathrm{HS}} \ge \sqrt{\sum_{i=1}^n \left(\frac{\lambda_i'(t)}{\lambda_i(t)^{1-\alpha}}\right)^2}. \end{aligned}$$

Define

$$\xi(t) := \mathrm{Diag}\left(\frac{1}{\alpha}\lambda_1(t)^\alpha, \dots, \frac{1}{\alpha}\lambda_n(t)^\alpha\right), \qquad 0 \le t \le 1,$$

which is a continuous curve from $\alpha^{-1}A^\alpha$ to $\alpha^{-1}B^\alpha$. Since

$$\xi'(t) = \mathrm{Diag}\left(\lambda_1(t)^{\alpha-1}\lambda_1'(t), \dots, \lambda_n(t)^{\alpha-1}\lambda_n'(t)\right)$$

except for a countable set, we have

$$L_\phi(\gamma) \ge \int_0^1 \|\xi'(t)\|_{\mathrm{HS}}\,dt \ge \left\|\int_0^1 \xi'(t)\,dt\right\|_{\mathrm{HS}} = \frac{1}{|\alpha|}\|A^\alpha - B^\alpha\|_{\mathrm{HS}}.$$

Furthermore, let

$$\begin{aligned} \gamma_{A,B}(t) &:= ((1-t)A^\alpha + tB^\alpha)^{1/\alpha} \\ &= \mathrm{Diag}\left(((1-t)\lambda_1^\alpha + t\mu_1^\alpha)^{1/\alpha}, \dots, ((1-t)\lambda_n^\alpha + t\mu_n^\alpha)^{1/\alpha}\right), \end{aligned}$$

where $A = \mathrm{Diag}(\lambda_1,\dots,\lambda_n)$ and $B = \mathrm{Diag}(\mu_1,\dots,\mu_n)$. Then one can directly compute

$$L_\phi(\gamma_{A,B}) = \frac{1}{|\alpha|}\sqrt{\sum_{i=1}^n (\lambda_i^\alpha - \mu_i^\alpha)^2} = \frac{1}{|\alpha|}\|A^\alpha - B^\alpha\|_{\mathrm{HS}}.$$

Hence the desired conclusion in the case $\alpha \neq 0$ follows. □

Proposition 1.2. *For $\rho > 1$ let $\mathbb{P}_n(\rho) := \{A \in \mathbb{P}_n : \rho^{-1} I \le A \le \rho I\}$. Assume that $\alpha \in \mathbb{R}$ and $\rho > 1$ satisfy*

$$\rho^{|\alpha-1|}(\rho - 1)|\alpha| < \frac{1}{3\sqrt{n}}. \tag{1.4}$$

Then for any $M \in \mathfrak{M}_0$ and for any $A, B \in \mathbb{P}_n(\rho)$, there exists a geodesic shortest curve joining A, B with respect to K^ϕ where $\phi(x, y) := M(x, y)^{2(1-\alpha)}$.

Proof. By [10, Theorem 3.1] we may consider the case $\alpha \neq 0$. Let $\phi := M^{2(1-\alpha)}$ with any $M \in \mathfrak{M}_0$. For every $A, B \in \mathbb{P}_n(\rho)$ consider the line segment $\gamma_0(t) := (1 - t)A + tB$, $0 \le t \le 1$. Then for every $t \in [0, 1]$, with the diagonalization $\gamma_0(t) = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$ we have

$$\begin{aligned}
\|\phi(\mathbf{L}_{\gamma_0(t)}, \mathbf{R}_{\gamma_0(t)})^{-1/2}\gamma_0'(t)\|_{\mathrm{HS}} &= \left\| \left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}} \right]_{ij} \circ (B - A) \right\|_{\mathrm{HS}} \\
&\le \max_{x,y \in [\alpha^{-1}, \alpha]} M(x, y)^{\alpha-1} (\|A - I\|_{\mathrm{HS}} + \|B - I\|_{\mathrm{HS}}) \\
&\le \rho^{|\alpha-1|} \cdot 2\sqrt{n}(\rho - 1),
\end{aligned}$$

since $|A - I| \le (\rho - 1)I$ and $|B - I| \le (\rho - 1)I$. Hence, letting $\rho_0 := \sqrt{n}\rho^{|\alpha-1|}(\rho - 1)$, we have

$$\delta_\phi(A, B) \le L_\phi(\gamma_0) < 2\rho_0$$

and similarly $\delta_\phi(A, I) < \rho_0$ for all $M \in \mathfrak{M}_0$ and all $A, B \in \mathbb{P}_n(\rho)$. For each $A, B \in \mathbb{P}_n(\rho)$ we choose a sequence $\{\gamma_k\}$ of C^1 curves from A to B (depending on $\phi = M^{2(1-\alpha)}$) such that $L_\phi(\gamma_k) \to \delta_\phi(A, B)$ as $k \to \infty$. Here, each γ_k has the parametrization of constant speed. By Theorem 1.1, for every $t \in [0, 1]$ we have

$$\begin{aligned}
\frac{1}{|\alpha|}\|\gamma_k(t)^\alpha - I\|_{\mathrm{HS}} &= \delta_\phi(\gamma_k(t), I) \le \delta_\phi(A, \gamma_k(t)) + \delta_\phi(A, I) \\
&\le L_\phi(\gamma_k) + \delta_\phi(A, I) \\
&\longrightarrow \delta_\phi(A, B) + \delta_\phi(A, I) \quad \text{as } k \to \infty.
\end{aligned}$$

Hence it may be assumed that $|\alpha|^{-1}\|\gamma_k(t)^\alpha - I\|_{\mathrm{HS}} < 3\rho_0$ for all $k \ge 1$ and all $t \in [0, 1]$. Then we have

$$\|\gamma_k(t)^\alpha - I\| \le \|\gamma_k(t)^\alpha - I\|_{\mathrm{HS}} < 3\rho_0|\alpha|,$$

where $\|\cdot\|$ denotes the operator norm. Assumption (1.4) means that $3\rho_0|\alpha| < 1$. Therefore,

$$(1 - 3\rho_0|\alpha|)I \le \gamma_k(t)^\alpha \le (1 + 3\rho_0|\alpha|)I$$

so that

$$R^{-1}I \le \gamma_k(t) \le RI, \qquad k \ge 1,\ 0 \le t \le 1, \tag{1.5}$$

where

$$R := (1 - 3\rho_0|\alpha|)^{-1/|\alpha|}.$$

Since γ_k is of constant speed, we have, for every $k \ge 1$ and every $t \in [0,1]$,

$$L_\phi(\gamma_k) = \sqrt{K^\phi_{\gamma_k(t)}(\gamma_k'(t), \gamma_k'(t))} = \left\| \left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}} \right]_{ij} \circ (U^* \gamma_k'(t) U) \right\|_{\mathrm{HS}} \tag{1.6}$$

with the diagonalization $\gamma_k(t) = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$. From (1.5) and (1.6) together with $L_\phi(\gamma_k) \to \delta_\phi(A, B) < 2\rho_0$ we see that there exists a $C > 0$ such that

$$\|\gamma_k'(t)\|_{\mathrm{HS}} \le C, \qquad k \ge 1,\ 0 \le t \le 1. \tag{1.7}$$

Let $L^1([0,1]; \mathbb{H}_n)$ denote the L^1-space of $\mathbb{H}_n$ ($\cong \mathbb{R}^{n^2}$)-valued integrable functions with respect to the Lebesgue measure on $[0,1]$ with the norm $\|f\|_{L^1} := \int_0^1 \|f(t)\|_{\mathrm{HS}}\, dt$. It follows from (1.7) that $\{\gamma_k'\}$ is relatively compact in $L^1([0,1]; \mathbb{H}_n)$ in the weak topology. So, by taking a subsequence, we may assume that $\{\gamma_k'\}$ itself converges to an $\eta \in L^1([0,1]; \mathbb{H}_n)$ in the weak topology. Then it is easily verified from (1.7) that $\|\eta(t)\|_{\mathrm{HS}} \le C$ a.e. $t \in [0,1]$. Define

$$\gamma(t) := A + \int_0^t \eta(s)\, ds, \qquad 0 \le t \le 1,$$

which is differentiable with $\gamma'(t) = \eta(t)$ for a.e. $t \in [0,1]$. We have

$$\gamma(t) = \lim_{k\to\infty} \left(A + \int_0^t \gamma_k'(s)\, ds \right) = \lim_{k\to\infty} \gamma_k(t), \qquad 0 \le t \le 1, \tag{1.8}$$

which implies from (1.5) that

$$R^{-1} I \le \gamma(t) \le RI, \qquad 0 \le t \le 1. \tag{1.9}$$

Hence γ is an absolutely continuous curve inside $\mathbb{P}_n$ joining A, B.

Thanks to (1.7), for every $t \in [0,1]$ we estimate

$$\begin{aligned}
&\big\langle \gamma_k'(t), \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} \gamma_k'(t) \big\rangle_{\mathrm{HS}} \\
&\quad = \big\langle \gamma_k'(t), \big(\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k'(t)})^{-1} \big) \gamma_k'(t) \big\rangle_{\mathrm{HS}} \\
&\qquad + \big\langle \gamma_k'(t), \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} \gamma_k'(t) \big\rangle_{\mathrm{HS}} \\
&\quad \le C^2 \big\| \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} \big\| + K^\phi_{\gamma_k(t)}(\gamma_k'(t), \gamma_k'(t))
\end{aligned}$$

and hence

$$\begin{aligned}
&\big\| \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2} \gamma_k'(t) \big\| \\
&\quad \le C \big\| \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} \big\|^{1/2} + \sqrt{K^\phi_{\gamma_k(t)}(\gamma_k'(t), \gamma_k'(t))},
\end{aligned} \tag{1.10}$$

where the norm $\|\cdot\|$ denotes the operator norm as an operator on $(\mathbb{H}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$. From (1.5) and (1.9) we notice that

$$\big\| \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2} \big\| \le R^{|\alpha-1|}, \tag{1.11}$$

$$\left\|\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1}-\phi(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}\right\|\le 2R^{2|\alpha-1|}.$$

Approximating $\phi(x,y)^{-1}$ by polynomials in x,y uniformly for $(x,y)\in[R^{-1},R]^2$ and applying (1.8) one can show that

$$\lim_{k\to\infty}\left\|\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1}-\phi(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}\right\|=0,\qquad 0\le t\le 1.$$

Hence the dominated convergence theorem implies that

$$\lim_{k\to\infty}\int_0^1\left\|\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1}-\phi(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}\right\|^{1/2}dt=0. \tag{1.12}$$

Furthermore, thanks to (1.11) one can define a bounded operator $\mathbf{A}$ on $L^1([0,1];\mathbb{H}_n)$ by

$$(\mathbf{A}f)(t):=\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1/2}f(t),\qquad f\in L^1([0,1];\mathbb{H}_n),\ 0\le t\le 1.$$

Since $\mathbf{A}\gamma_k'\to\mathbf{A}\gamma'$ in the weak topology, it follows that $\|\mathbf{A}\gamma'\|_{L^1}\le\liminf_{k\to\infty}\|\mathbf{A}\gamma_k'\|_{L^1}$, that is,

$$\int_0^1\left\|\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1/2}\gamma'(t)\right\|_{\mathrm{HS}}dt\le\liminf_{k\to\infty}\int_0^1\left\|\phi(\mathbf{L}_{\gamma(t)},\mathbf{R}_{\gamma(t)})^{-1/2}\gamma_k'(t)\right\|_{\mathrm{HS}}dt. \tag{1.13}$$

Since

$$\int_0^1\sqrt{K_{\gamma_k(t)}^{\phi}(\gamma_k'(t),\gamma_k'(t))}\,dt=L_\phi(\gamma_k)\longrightarrow\delta_\phi(A,B)\quad\text{as } k\to\infty,$$

it follows from (1.10), (1.12) and (1.13) together that $L_\phi(\gamma)\le\delta_\phi(A,B)$ and hence $L_\phi(\gamma)=\delta_\phi(A,B)$.

Finally, thanks to $\mathbf{A}\gamma_k'\to\mathbf{A}\gamma'$ in the weak topology, (1.10) and (1.13) once again, we have for every $t\in[0,1]$

$$\begin{aligned}\int_0^t\left\|\phi(\mathbf{L}_{\gamma(s)},\mathbf{R}_{\gamma(s)})^{-1/2}\gamma'(s)\right\|_{\mathrm{HS}}ds&\le\liminf_{k\to\infty}\int_0^t\left\|\phi(\mathbf{L}_{\gamma(s)},\mathbf{R}_{\gamma(s)})^{-1/2}\gamma_k'(s)\right\|_{\mathrm{HS}}ds\\&\le\liminf_{k\to\infty}\int_0^t\sqrt{K_{\gamma_k(s)}^{\phi}(\gamma_k'(s),\gamma_k'(s))}\,ds\\&=t\delta_\phi(A,B)=tL_\phi(\gamma).\end{aligned}$$

Hence $\sqrt{K_{\gamma(t)}^{\phi}(\gamma'(t),\gamma'(t))}=L_\phi(\gamma)$ for a.e. $t\in[0,1]$, that is, γ has constant speed almost everywhere. Recall a general fact ([12, Theorem IV.3.6], [19, Theorem 14]) that a geodesic shortest curve in a smooth Riemannian manifold is locally unique and it is smooth if it has constant speed. From this we see that γ is a smooth geodesic shortest curve joining A,B. □

Remark that for any $\rho>1$, (1.4) is satisfied if $|\alpha|$ is sufficiently small, and for any $\alpha\in\mathbb{R}$, (1.4) is satisfied if ρ is sufficiently near 1. The main point of the above proof is to show that approximating curves γ_k stay in a compact range as (1.5), where assumption (1.4) is essential. So it seems difficult to apply the above method without assumption like (1.4).

2 Characterization of isometric transformations

Let $M, N \in \mathfrak{M}_0$ and $\alpha, \beta \in \mathbb{R}$. Set two kernel functions

$$\phi(x,y) := M(x,y)^{2(1-\alpha)}, \quad \psi(x,y) := N(x,y)^{2(1-\beta)}, \qquad x, y > 0,$$

and the corresponding Riemannian metrics on $\mathbb{P}_n$

$$K_D^\phi(H,K) := \left\langle H, \phi(\mathbf{L}_D, \mathbf{R}_D)^{-1} K \right\rangle_{\mathrm{HS}}, \quad K_D^\psi(H,K) := \left\langle H, \psi(\mathbf{L}_D, \mathbf{R}_D)^{-1} K \right\rangle_{\mathrm{HS}}$$

for $D \in \mathbb{P}_n$ and $H, K \in \mathbb{H}_n$. Moreover, let F be a smooth function from $(0,\infty)$ onto itself. Assume that $F'(x) \neq 0$ for all $x > 0$, so F is a C^∞ homeomorphism from $(0,\infty)$ onto itself. The aim of this section is to characterize when the functional calculus

$$D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{P}_n$$

gives an isometric transformation between two Riemannian manifolds $\mathbb{P}_n$ with a scalar multiple of metric K^ϕ and $\mathbb{P}_n$ with K^ψ. The most natural transformations on $\mathbb{P}_n$ from the operator/matrix theory viewpoint are those defined via functional calculus as treated here.

Theorem 2.1. *Under the above assumptions and with $\kappa > 0$, the transformation $D \in \mathbb{P}_n \mapsto F(D) \in \mathbb{P}_n$ is isometric from $(\mathbb{P}_n, \kappa^2 K^\phi)$ onto $(\mathbb{P}_n, K^\psi)$ if and only if one of the following holds:*

(i) $\alpha = \beta = 1$ *and* $F(x) = \kappa x$, $x > 0$. *(In this case, K^ϕ, K^ψ are the Euclidean metric irrelevantly to M, N.)*

(ii) $\alpha \neq 0, 1$, $\beta = 1$, *and*

$$F(x) = \frac{\kappa}{|\alpha|} x^\alpha, \qquad x > 0,$$

$$M(x,y) = \left(\frac{\alpha(x-y)}{x^\alpha - y^\alpha} \right)^{\frac{1}{1-\alpha}}, \qquad x, y > 0.$$

(In this case, K^ψ is the Euclidean metric irrelevantly to N, and K^ϕ is a pull-back of the Euclidean metric.)

(iii) $\alpha = 1$, $\beta \neq 0, 1$, *and*

$$F(x) = (\kappa|\beta|x)^{1/\beta}, \qquad x > 0,$$

$$N(x,y) = \left(\frac{\beta(x-y)}{x^\beta - y^\beta} \right)^{\frac{1}{1-\beta}}, \qquad x, y > 0.$$

(This is the case where the roles of ϕ and ψ in (ii) *are exchanged.)*

(iv) $\alpha, \beta \neq 0, 1$ *and*

$$F(x) = \left(\kappa \left| \frac{\beta}{\alpha} \right| \right)^{1/\beta} x^{\alpha/\beta}, \qquad x > 0,$$

$$M(x,y) = \left(\frac{\alpha}{\beta} \cdot \frac{x-y}{x^{\alpha/\beta} - y^{\alpha/\beta}} \right)^{\frac{1}{1-\alpha}} N\left(x^{\alpha/\beta}, y^{\alpha/\beta}\right)^{\frac{1-\beta}{1-\alpha}}, \qquad x, y > 0.$$

(v) $\alpha = \beta = 0$ *and either*

$$F(x) = cx^{\kappa}, \quad x > 0 \quad \textit{with a constant } c > 0,$$

$$M(x,y) = \frac{\kappa(x-y)}{x^{\kappa} - y^{\kappa}} N(x^{\kappa}, y^{\kappa}), \qquad x, y > 0,$$

or

$$F(x) = cx^{-\kappa}, \quad x > 0 \quad \textit{with a constant } c > 0,$$

$$M(x,y) = \frac{\kappa(y-x)}{x^{-\kappa} - y^{-\kappa}} N(x^{-\kappa}, y^{-\kappa}), \qquad x, y > 0.$$

Proof. For any C^1 curve $\gamma : [0,1] \to \mathbb{P}_n$ let $\xi(t) := F(\gamma(t))$. With the diagonalization $\gamma(t) = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$ for each fixed $t \in [0,1]$, we have (see (1.1) and [10, (2.8)])

$$K^{\phi}_{\gamma(t)}(\gamma'(t), \gamma'(t)) = \left\| \left[\frac{1}{\sqrt{\phi(\lambda_i, \lambda_j)}} \right]_{ij} \circ (U^*\gamma'(t)U) \right\|_{\mathrm{HS}}^2,$$

$$K^{\psi}_{\xi(t)}(\xi'(t), \xi'(t)) = \left\| \left[\frac{F^{[1]}(\lambda_i, \lambda_j)}{\sqrt{\psi(F(\lambda_i), F(\lambda_j))}} \right]_{ij} \circ (U^*\gamma'(t)U) \right\|_{\mathrm{HS}}^2,$$

where $F^{[1]}(x,y)$ is the *divided difference* of F given by

$$F^{[1]}(x,y) := \begin{cases} \frac{F(x)-F(y)}{x-y} & \text{if } x \neq y, \\ F'(x) & \text{if } x = y. \end{cases}$$

Here, the eigenvalues of $\gamma(t)$ can be arbitrary positive numbers and $U^*\gamma'(t)U$ can be arbitrary element of $\mathbb{H}_n$ as γ is arbitrary. Hence one can see that $D \mapsto F(D)$ is isometric from $(\mathbb{P}_n, \kappa^2 K^{\phi})$ onto $(\mathbb{P}_n, K^{\psi})$ if and only if

$$\frac{F^{[1]}(x,y)}{\sqrt{\psi(F(x), F(y))}} = \pm \frac{\kappa}{\sqrt{\phi(x,y)}} \tag{2.1}$$

for all $x, y > 0$. Letting $x = y$ in (2.1) gives

$$\frac{F'(x)}{F(x)^{1-\beta}} = \pm \frac{\kappa}{x^{1-\alpha}}, \qquad x > 0. \tag{2.2}$$

Note [10, Theorem 3.1] that $(\mathbb{P}_n, K^{\phi})$ is complete if and only if $\alpha = 0$. Hence the cases $\alpha = 0 \neq \beta$ and $\alpha \neq 0 = \beta$ can not occur as far as K^{ψ} is isometric to $\kappa^2 K^{\phi}$. Hence we may consider the cases $\alpha, \beta \neq 0$ and $\alpha = \beta = 0$ below. To make discussions clearer, we further divide the case $\alpha, \beta \neq 0$ into the four cases $\alpha = \beta = 1$, $\alpha \neq 1 = \beta$, $\alpha = 1 \neq \beta$, and $\alpha, \beta \neq 1$.

(1) Case $\alpha = \beta = 1$. Equation (2.2) is simply $F'(x) = \pm\kappa$, which is solved as $F(x) = \pm\kappa x + c$ with $c \in \mathbb{R}$. Since F maps $(0, \infty)$ onto itself and $\kappa > 0$ by assumption, only $F(x) = \kappa x$ is possible. Conversely, (2.1) is satisfied for this F with $\alpha = \beta = 1$.

(2) Case $\alpha \neq 0, 1$ and $\beta = 1$. Equation (2.2) is $F'(x) = \pm\kappa x^{\alpha-1}$, which is solved as

$$F(x) = \pm\frac{\kappa}{\alpha}\,(x^\alpha + c) \quad \text{with } c \in \mathbb{R}.$$

By the assumptions on F and α, F is determined as in (ii). Inserting the form of F into (2.1) determines M as in (ii). Conversely, (2.1) is satisfied for these F and M with $\beta = 1$.

(3) Case $\alpha = 1$ and $\beta \neq 0, 1$. Equation (2.2) is $F'(x)/F(x)^{1-\beta} = \pm\kappa$, which is solved as

$$\frac{1}{\beta}\,F(x)^\beta = \pm\kappa x + c \quad \text{with } c \in \mathbb{R}.$$

Hence F and N are determined as in (iii) similarly to the above case. Conversely, (2.1) is satisfied for these F and N with $\alpha = 1$.

(4) Case $\alpha, \beta \neq 0, 1$. The solution of (2.2) is

$$\frac{1}{\beta}\,F(x)^\beta = \pm\frac{\kappa}{\alpha}(x^\alpha + c) \quad \text{with } c \in \mathbb{R}.$$

Hence F is determined as in (iv). This and (2.1) determine M from N as stated. Conversely, (2.1) is satisfied if F, N, and M are given as in (iv).

(5) Case $\alpha = \beta = 0$. Equation (2.2) is solved as $F(x) = cx^\kappa$ or $F(x) = cx^{-\kappa}$ with $c > 0$. These and (2.1) determine M from N as respectively stated in (v). Conversely, (2.1) is satisfied if F, N, and M are given as in (v). □

Remark 2.2. Theorem 2.1 contains both Theorems 2.1 and 3.3 of [10]. Indeed, [10, Theorem 2.1] is just case (ii) and [10, Theorem 3.3] is a particular case of (v) with $N(x, y) = \sqrt{xy}$.

Remark 2.3. The relation between M and N in (iv) is rephrased as

$$N(x, y) = \left(\frac{\beta}{\alpha} \cdot \frac{x - y}{x^{\beta/\alpha} - y^{\beta/\alpha}}\right)^{\frac{1}{1-\beta}} M\left(x^{\beta/\alpha}, y^{\beta/\alpha}\right)^{\frac{1-\alpha}{1-\beta}}$$

so that the relation is completely symmetric for M and N. The relations between $M^{2(1-\alpha)}$ and $N^{2(1-\beta)}$ given in (i)–(iv) define an equivalence relation on the set of kernel functions $\{M^\theta : M \in \mathfrak{M}_0,\ \theta \in \mathbb{R} \setminus \{2\}\}$. On the other hand, the two relations between M and N in (v) are unified as the first one with $\kappa \in \mathbb{R} \setminus \{0\}$, which defines an equivalence relation on $\mathfrak{M}_0$.

3 Two kinds of isometric families of Riemannian metrics and their convergence property

By taking account of cases (iv) and (v) of Theorem 2.1, for each $N \in \mathfrak{M}_0$, $\alpha \in \mathbb{R} \setminus \{0,1\}$, $\beta \in \mathbb{R} \setminus \{0\}$, and $\kappa \in \mathbb{R} \setminus \{0\}$, we introduce the following two kinds of kernel functions:

$$N_{\alpha,\beta}(x,y) := \left(\frac{\alpha}{\beta} \cdot \frac{x-y}{x^{\alpha/\beta} - y^{\alpha/\beta}} \right)^{\frac{1}{1-\alpha}} N\left(x^{\alpha/\beta}, y^{\alpha/\beta}\right)^{\frac{1-\beta}{1-\alpha}}, \tag{3.1}$$

$$N_\kappa(x,y) := \frac{\kappa(x-y)}{x^\kappa - y^\kappa} N(x^\kappa, y^\kappa), \qquad x,y > 0. \tag{3.2}$$

In particular, when $\beta = 1$, $N_{\alpha,1}$'s are *Stolarsky means* (see [23, 8, 7, 10])

$$S_\alpha(x,y) := \left(\frac{\alpha(x-y)}{x^\alpha - y^\alpha} \right)^{\frac{1}{1-\alpha}},$$

which interpolate the following familiar means:

$$S_2(x,y) = M_{\mathrm{A}}(x,y) := \frac{x+y}{2} \quad \text{(arithmetic mean)},$$

$$S_{1/2}(x,y) = M_{\surd}(x,y) := \left(\frac{\sqrt{x} + \sqrt{y}}{2} \right)^2 \quad \text{(root mean)},$$

$$S_{-1}(x,y) = M_{\mathrm{G}}(x,y) := \sqrt{xy} \quad \text{(geometric mean)}.$$

Furthermore, S_α for $\alpha = 0, 1$ are understood as

$$S_0(x,y) := \lim_{\alpha \to 0} S_\alpha(x,y) = M_{\mathrm{L}}(x,y) := \frac{x-y}{\log x - \log y} \quad \text{(logarithmic mean)},$$

$$S_1(x,y) := \lim_{\alpha \to 1} S_\alpha(x,y) = \frac{1}{e} \left(\frac{x^x}{y^y} \right)^{\frac{1}{x-y}} \quad \text{(identric mean)}.$$

So we may define $N_{\alpha,1}$ for $\alpha = 0, 1$ as S_0 and S_1, respectively.

Proposition 3.1. *Let $N_{\alpha,\beta}$ and N_κ be defined for each N, α, β, and κ as specified above. Then the following hold:*

(a) *For any $\alpha \in \mathbb{R} \setminus \{0,1\}$ and $\beta \in \mathbb{R} \setminus \{0\}$,*

$$N_{\alpha,\beta}(x,y) = S_{\alpha/\beta}(x,y)^{\frac{\beta-\alpha}{(1-\alpha)\beta}} N\left(x^{\alpha/\beta}, y^{\alpha/\beta}\right)^{\frac{1-\beta}{1-\alpha}},$$

$$\lim_{\alpha \to 0} N_{\alpha,\beta}(x,y) = M_{\mathrm{L}}(x,y).$$

Furthermore, if either $0 < \alpha \le \beta \le 1$ or $0 > \alpha \ge \beta$, then $N_{\alpha,\beta} \in \mathfrak{M}_0$.

(b) *For any $\kappa \in \mathbb{R} \setminus \{0\}$,*

$$N_\kappa(x,y) = S_{2-2\kappa}(x,y)^{1-\kappa} N(x^\kappa, y^\kappa),$$

$$\lim_{\kappa \to 0} N_\kappa(x,y) = M_{\mathrm{L}}(x,y).$$

Furthermore, if $0 < \kappa \le 1$, then $N_\kappa \in \mathfrak{M}_0$.

Proof. The first expression in (a) is seen by a simple computation, from which the assertion for convergence follows immediately. If $\beta = 1$, then $N_{\alpha,1} = S_\alpha \in \mathfrak{M}_0$ for all $\alpha \in \mathbb{R}$ as mentioned before the proposition. If either $0 < \alpha \le \beta < 1$ or $0 > \alpha \ge \beta$, then α/β, $(1-\beta)/(1-\alpha)$, and $(\beta-\alpha)/(1-\alpha)\beta$ are all nonnegative, and so $N_{\alpha,\beta} \in \mathfrak{M}_0$ is easily verified. Hence (a) follows. The proof of (b) is similar. $\square$

By Theorem 2.1 and Proposition 3.1 we see that if $N \in \mathfrak{M}_0$ and $\beta \le 1$ with $\beta \ne 0$, then $K^{N_{\alpha,\beta}^{2(1-\alpha)}}$ for $\beta \ge \alpha > 0$ or $\beta \le \alpha < 0$ is a one-parameter family of isometric Riemannian metrics on $\mathbb{P}_n$ starting from $K^{N^{2(1-\beta)}}$ (at $\alpha = \beta$) and converging to $K^{M_{\mathrm{L}}^2}$ as $\alpha \to 0$. Also, for each $N \in \mathfrak{M}_0$, $K^{N_\kappa^2}$ for $1 \ge \alpha > 0$ is a one-parameter family of Riemannian metrics on $\mathbb{P}_n$ starting from K^{N^2} (at $\kappa = 1$) and converging to $K^{M_{\mathrm{L}}^2}$ as $\kappa \searrow 0$. It is remarkable that the Riemannian metric $K^{M_{\mathrm{L}}^2}$ corresponding to the square of the logarithmic mean is the common limiting point of both one-parameter families. But $K^{M_{\mathrm{L}}^2}$ itself is not isometric to any other Riemannian metric K^{N^θ} on $\mathbb{P}_n$ where $N \in \mathfrak{M}_0$ and $\theta \in \mathbb{R}$. So we may consider $K^{M_{\mathrm{L}}^2}$ as a kind of attractor for all Riemannian metrics on $\mathbb{P}_n$ corresponding to powers of mean functions.

We note also that $K^{M_{\mathrm{L}}^2}$ is a unique Riemannian metric with degree 2 that is a pull-back of the Euclidean metric (see [10, Theorem 2.1]); however it did not appear in case (ii) of Theorem 2.1 because the pull-back is given by the transformation $D \in \mathbb{P}_n \mapsto \log D \in \mathbb{H}_n$ that is not inside $\mathbb{P}_n$. From this transformation we immediately see that, for every $A, B \in \mathbb{P}_n$, a unique geodesic shortest curve joining A, B with respect to $K^{M_{\mathrm{L}}^2}$ is given by

$$\gamma_{A,B}(t) := \exp((1-t)\log A + t\log B), \qquad 0 \le t \le 1,$$

and the geodesic distance between A, B is

$$\delta_{M_{\mathrm{L}}^2}(A,B) = \|\log A - \log B\|_{\mathrm{HS}}.$$

In the rest of this section we will prove two theorems, which say that $K^{M_{\mathrm{L}}^2}$ is indeed a limit point (attractor) not only in the sense of Riemannian metrics but also in that of geodesic distances and geodesic shortest curves as well. To show this, we first give a lemma.

Lemma 3.2. *Let $N \in \mathfrak{M}_0$ and $R > 1$ be arbitrary.*

(a) *For each $\beta \le 1$ with $\beta \ne 0$ let $N_{\alpha,\beta}$ be in (3.1) for $0 < \alpha \le \beta$ or $0 > \alpha \ge \beta$. Then*

$$\frac{N_{\alpha,\beta}(x,y)^{2(1-\alpha)}}{M_{\mathrm{L}}(x,y)^2} \longrightarrow 1 \quad \textit{uniformly for } x, y \in [R^{-1}, R]$$

as $\alpha \searrow 0$ for $0 < \alpha \le \beta$ or $\alpha \nearrow 0$ for $0 > \alpha \ge \beta$.

(b) *Let N_κ be in (3.2) for $0 < \kappa \le 1$. Then*

$$\frac{N_\kappa(x,y)^2}{M_{\mathrm{L}}(x,y)^2} \longrightarrow 1 \quad \textit{uniformly for } x, y \in [R^{-1}, R]$$

as $\kappa \searrow 0$.

Proof. (a) By homogeneity and symmetry, letting $\nu := \alpha/\beta > 0$, we have

$$\begin{aligned}
\max_{x,y\in[R^{-1},R]} \frac{N_{\alpha,\beta}(x,y)^{2(1-\alpha)}}{M_{\mathrm{L}}(x,y)^2} &= \max_{x,y\in[R^{-1},R],\,x\ge y} y^{-2\alpha}\,\frac{N_{\alpha,\beta}(x/y,1)^{2(1-\alpha)}}{M_{\mathrm{L}}(x/y,1)^2} \\
&\le R^{2|\alpha|} \max_{x\in[1,R^2]} \frac{\nu^2\left(\frac{x-1}{x^\nu-1}\right)^2 N(x^\nu,1)^{2(1-\beta)}}{\left(\frac{x-1}{\log x}\right)^2} \\
&\le R^{2|\alpha|} N(R^{2\nu},1)^{2(1-\beta)} \left(\max_{x\in[1,R^2]} \frac{\nu\log x}{x^\nu-1}\right)^2 \\
&= R^{2|\alpha|} N(R^{2\nu},1)^{2(1-\beta)} \left(\max_{t\in[0,2\nu\log R]} \frac{t}{e^t-1}\right)^2 \\
&\le R^{2|\alpha|} N(R^{2\nu},1)^{2(1-\beta)} \longrightarrow 1
\end{aligned}$$

and

$$\begin{aligned}
\min_{x,y\in[R^{-1},R]} \frac{N_{\alpha,\beta}(x,y)^{2(1-\alpha)}}{M_{\mathrm{L}}(x,y)^2} &\ge R^{-2|\alpha|} N(R^{-2\nu},1)^{2(1-\beta)} \left(\min_{t\in[0,2\nu\log R]} \frac{t}{e^t-1}\right)^2 \\
&\ge R^{-2|\alpha|} N(R^{-2\nu},1)^{2(1-\beta)} \left(\frac{2\nu\log R}{e^{2\nu\log R}-1}\right)^2 \longrightarrow 1
\end{aligned}$$

as $\alpha \to 0$ and so $\nu \to 0$.

(b) Similarly, we have

$$\begin{aligned}
\max_{x,y\in[R^{-1},R]} \frac{N_\kappa(x,y)^2}{M_{\mathrm{L}}(x,y)^2} &= \max_{x,y\in[R^{-1},R],\,x\ge y} \frac{N_\kappa(x/y,1)^2}{M_{\mathrm{L}}(x/y,1)^2} \\
&= \max_{x\in[1,R^2]} \frac{\kappa^2\left(\frac{x-1}{x^\kappa-1}\right)^2 N(x^\kappa,1)^2}{\left(\frac{x-1}{\log x}\right)^2} \\
&\le N(R^{2\kappa},1)^2 \left(\max_{x\in[1,R^2]} \frac{\kappa\log x}{x^\kappa-1}\right)^2 \\
&\le N(R^{2\kappa},1)^2 \longrightarrow 1
\end{aligned}$$

and

$$\min_{x,y\in[R^{-1},R]} \frac{N_\kappa(x,y)^2}{M_{\mathrm{L}}(x,y)^2} \ge N(R^{-2\kappa},1)^2 \left(\frac{2\kappa\log R}{e^{2\kappa\log R}-1}\right)^2 \longrightarrow 1$$

as $\kappa \searrow 0$. □

Theorem 3.3. *Let $N \in \mathfrak{M}_0$ and $A, B \in \mathbb{P}_n$ be arbitrary.*

(a) *For each $\beta \le 1$ with $\beta \ne 0$ let $N_{\alpha,\beta}$ be in (3.1) for $0 < \alpha \le \beta$ or $0 > \alpha \ge \beta$. Then*

$$\lim_{\alpha\to 0} \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,B) = \lim_{\alpha\to 0} \delta_{N^{2(1-\beta)}}(A_{\alpha,\beta}, B_{\alpha,\beta}) = \|\log A - \log B\|_{\mathrm{HS}},$$

where $\alpha \to 0$ means $\alpha \searrow 0$ for $0 < \alpha \le \beta$ or $\alpha \nearrow 0$ for $0 > \alpha \ge \beta$, and

$$A_{\alpha,\beta} := \left(\frac{\beta}{\alpha}\right)^{1/\beta} A^{\alpha/\beta}, \quad B_{\alpha,\beta} := \left(\frac{\beta}{\alpha}\right)^{1/\beta} B^{\alpha/\beta}. \tag{3.3}$$

(b) *Let N_κ be in (3.2) for $0 < \kappa \le 1$. Then*

$$\lim_{\kappa\searrow 0} \delta_{N_\kappa^2}(A,B) = \lim_{\kappa\searrow 0} \frac{1}{\kappa}\delta_{N^2}(A^\kappa, B^\kappa) = \|\log A - \log B\|_{\mathrm{HS}}.$$

Proof. (a) Assume that $0 < \alpha \le \beta \le 1$ or $0 > \alpha \ge \beta$. Since Theorem 2.1 implies that $\delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,B) = \delta_{N^{2(1-\beta)}}(A_{\alpha,\beta}, B_{\alpha,\beta})$, it suffices to prove that

$$\lim_{\alpha\to 0} \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,B) = \delta_{M_{\mathrm{L}}^2}(A,B). \tag{3.4}$$

By Theorem 1.1 and [10, Lemma 3.2],

$$\delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,B) \le \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,I) + \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(B,I) = \frac{1}{|\alpha|}\left(\|A^\alpha - I\|_{\mathrm{HS}} + \|B^\alpha - I\|_{\mathrm{HS}}\right),$$
$$\delta_{M_{\mathrm{L}}^2}(A,B) \le \delta_{M_{\mathrm{L}}^2}(A,I) + \delta_{M_{\mathrm{L}}^2}(B,I) = \|\log A\|_{\mathrm{HS}} + \|\log B\|_{\mathrm{HS}},$$

and moreover,

$$\lim_{\alpha\to 0} \frac{1}{|\alpha|}\left(\|A^\alpha - I\|_{\mathrm{HS}} + \|B^\alpha - I\|_{\mathrm{HS}}\right) = \|\log A\|_{\mathrm{HS}} + \|\log B\|_{\mathrm{HS}}.$$

Hence one can choose a $\rho > 0$ (depending on A, B) such that

$$\|\log A\|_{\mathrm{HS}} + \|\log B\|_{\mathrm{HS}} < \rho, \quad \frac{1}{|\alpha|}\left(\|A^\alpha - I\|_{\mathrm{HS}} + \|B^\alpha - I\|_{\mathrm{HS}}\right) < \rho$$

for all α in $(0,\beta]$ or $[\beta,0)$. Let $\gamma : [0,1] \to \mathbb{P}_n$ be any C^1 curve joining A, B such that $L_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma) < \rho$. By Theorem 1.1 again, for every $t \in [0,1]$ we have

$$\frac{1}{|\alpha|}\|\gamma(t)^\alpha - I\| \le \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma(t), I) \le \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A, \gamma(t)) + \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A, I) < 2\rho,$$

where $\|\cdot\|$ denotes the operator norm. Therefore, when $2\rho|\alpha| < 1$,

$$(1 - 2\rho|\alpha|)^{1/|\alpha|} I \le \gamma(t) \le (1 - 2\rho|\alpha|)^{-1/|\alpha|} I.$$

Since $\lim_{\alpha\to 0}(1 - 2\rho|\alpha|)^{-1/|\alpha|} = e^{2\rho}$, one can choose a $\delta > 0$ with $\delta < \min\{|\beta|, 1/2\rho\}$ and an $R > e^{2\rho}$ such that if $0 < \alpha < \delta$ ($< \beta$) or $0 > \alpha > -\delta$ ($> \beta$) and if γ is a C^1 curve joining A, B with $L_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma) < \rho$, then

$$R^{-1} I \le \gamma(t) \le RI, \qquad 0 \le t \le 1. \tag{3.5}$$

Also, assume that γ is a C^1 curve joining A, B such that $L_{M_{\rm L}^2}(\gamma) < \rho$. By [10, Lemma 3.2], for every $t \in [0,1]$ we have

$$\|\log\gamma(t)\| \le \delta_{M_{\rm L}^2}(\gamma(t), I) \le \delta_{M_{\rm L}^2}(A, \gamma(t)) + \delta_{M_{\rm L}^2}(A, I) < 2\rho$$

so that (3.5) holds again.

By Lemma 3.2 (a), for any $\varepsilon > 0$ there exists a $\delta_0 \in (0, \delta)$ such that if $0 < \alpha < \delta_0$ or $0 > \alpha > -\delta_0$ then

$$\frac{1-\varepsilon}{M_{\rm L}(x,y)} \le \frac{1}{N_{\alpha,\beta}(x,y)^{1-\alpha}} \le \frac{1+\varepsilon}{M_{\rm L}(x,y)}, \qquad x, y \in [R^{-1}, R].$$

Assume that $0 < \alpha < \delta_0$ or $0 > \alpha > -\delta_0$, and let γ be a C^1 curve joining A, B such that either $L_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma) < \rho$ or $L_{M_{\rm L}^2}(\gamma) < \rho$. Since γ satisfies (3.5), we have

$$\begin{aligned} K_{\gamma(t)}^{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma'(t), \gamma'(t)) &= \left\| \left[\frac{1}{N_{\alpha,\beta}(\lambda_i, \lambda_j)^{1-\alpha}}\right]_{ij} \circ (U^*\gamma'(t)U) \right\|_{\rm HS}^2 \\ &\le (1+\varepsilon)^2 \left\| \left[\frac{1}{M_{\rm L}(\lambda_i, \lambda_j)}\right]_{ij} \circ (U^*\gamma'(t)U) \right\|_{\rm HS}^2 \\ &= (1+\varepsilon)^2 K_{\gamma(t)}^{M_{\rm L}^2}(\gamma'(t), \gamma'(t)) \end{aligned}$$

and similarly

$$K_{\gamma(t)}^{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma'(t), \gamma'(t)) \ge (1-\varepsilon)^2 K_{\gamma(t)}^{M_{\rm L}^2}(\gamma'(t), \gamma'(t)),$$

where $\gamma(t) = U{\rm Diag}(\lambda_1, \dots, \lambda_n)U^*$ is the diagonalization for each fixed $t \in [0,1]$. The above inequalities imply that

$$(1-\varepsilon)L_{M_{\rm L}^2}(\gamma) \le L_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma) \le (1+\varepsilon)L_{M_{\rm L}^2}(\gamma)$$

whenever α and γ satisfy the conditions stated above. From both sides inequalities above it follows that

$$(1-\varepsilon)\delta_{M_{\rm L}^2}(A,B) \le \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A,B) \le (1+\varepsilon)\delta_{M_{\rm L}^2}(A,B)$$

whenever $0 < \alpha < \delta_0$ or $0 > \alpha > -\delta_0$. Hence (3.4) has been proved.

(b) The proof is similar to that of (a) by using of Lemma 3.2 (b), so the details are left to the reader. □

We have the monotone convergence in Theorem 3.3 (b) under a certain assumption.

Proposition 3.4. *If $N \in \mathfrak{M}_0$ and $(\log x/(x-1))N(x,1)$ is decreasing (resp. increasing) in $x > 1$, then $\kappa^{-1}\delta_{N^2}(A^\kappa, B^\kappa)$ decreases (resp. increases) to $\|\log A - \log B\|_{\rm HS}$ as $1 \ge \kappa \searrow 0$ for every $A, B \in \mathbb{P}_n$.*

Proof. Assume that $(\log x/(x-1))N(x,1)$ is decreasing in $x>0$. If $x>1$ and $0<\kappa<\kappa'\le 1$, then

$$\frac{\log x^\kappa}{x^\kappa-1}N(x^\kappa,1)\ge\frac{\log x^{\kappa'}}{x^{\kappa'}-1}N(x^{\kappa'},1)$$

that is,

$$\frac{\kappa N(x^\kappa,1)}{x^\kappa-1}\ge\frac{\kappa' N(x^{\kappa'},1)}{x^{\kappa'}-1}.$$

For every $x>y>0$, replacing x by x/y gives

$$\frac{\kappa N(x^\kappa,y^\kappa)}{x^\kappa-y^\kappa}\ge\frac{\kappa' N(x^{\kappa'},y^{\kappa'})}{x^{\kappa'}-y^{\kappa'}}.$$

This implies that $N_\kappa(x,y)\ge N_{\kappa'}(x,y)$ for all $x,y>0$ if $0<\kappa<\kappa'\le 1$. Hence $\delta_{N_\kappa^2}(A,B)$ decreases as $1\ge\kappa\searrow 0$ (see [10, Theorem 4.1]). The proof of the increasing case is same with the opposite inequalities. □

Note that $N(x,y)\le M_{\mathrm{L}}(x,y)$ (resp., $N(x,y)\ge M_{\mathrm{L}}(x,y)$) for $x,y>0$ is necessary for the decreasingness (resp., increasingness) of $(\log x/(x-1))N(x,1)$. For example, $(\log x/(x-1))M_{\mathrm{A}}(x,1)$ and $(\log x/(x-1))M_{\sqrt{}}(x,1)$ are increasing in $x>1$ while $(\log x/(x-1))M_{\mathrm{G}}(x,1)$ is decreasing in $x>1$, where M_{A}, $M_{\sqrt{}}$, and M_{G} are the arithmetic, the root, and the geometric means, respectively, as introduced just before Proposition 3.1.

Theorem 3.5. *Let $N\in\mathfrak{M}_0$ and $A,B\in\mathbb{P}_n$ be arbitrary. In the following, assume that geodesic shortest curves are always parametrized under constant speed.*

(a) *For each $\beta\le 1$ with $\beta\ne 0$, if $|\alpha|$ is sufficiently small with $0<\alpha\le\beta$ or $0>\alpha\ge\beta$ depending on A,B (and N), then there exists a geodesic shortest curve $\gamma_{A_{\alpha,\beta},B_{\alpha,\beta}}$ joining $A_{\alpha,\beta},B_{\alpha,\beta}$ given in (3.3) with respect to $K^{N^{2(1-\beta)}}$, and*

$$\left(\frac{\alpha}{\beta}\right)^{1/\alpha}\left(\gamma_{A_{\alpha,\beta},B_{\alpha,\beta}}(t)\right)^{\beta/\alpha},\qquad 0\le t\le 1,\tag{3.6}$$

is a geodesic shortest curve joining A,B with respect to $K^{N_{\alpha,\beta}^{2(1-\alpha)}}$. Furthermore,

$$\lim_{\alpha\to 0}\left(\frac{\alpha}{\beta}\right)^{1/\alpha}\left(\gamma_{A_{\alpha,\beta},B_{\alpha,\beta}}(t)\right)^{\beta/\alpha}=\exp((1-t)\log A+t\log B),\quad 0\le t\le 1,\tag{3.7}$$

where $\alpha\to 0$ is meant as in Theorem 3.3.

(b) *For each $\kappa\in(0,1]$, if $\gamma_{A^\kappa,B^\kappa}$ is a geodesic shortest curve joining A^κ,B^κ with respect K^{N^2}, then*

$$\left(\gamma_{A^\kappa,B^\kappa}(t)\right)^{1/\kappa},\qquad 0\le t\le 1,$$

is a geodesic shortest curve joining A,B with respect to $K^{N_\kappa^2}$. Furthermore,

$$\lim_{\kappa\searrow 0}\left(\gamma_{A^\kappa,B^\kappa}(t)\right)^{1/\kappa}=\exp((1-t)\log A+t\log B),\quad 0\le t\le 1.\tag{3.8}$$

Proof. (a) When $A, B \in \mathbb{P}_n$ are fixed and $|\alpha|$ is small with $0 < \alpha \le \beta$ or $0 > \alpha \ge \beta$, we see by Lemma 3.2 (a) that $\delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A, B)$ is bounded from above. Hence in the same way as in the proof of Proposition 1.2, one can prove that there exists a geodesic shortest curve γ_α (of constant speed) joining A, B with respect to $K^{N_{\alpha,\beta}^{2(1-\alpha)}}$ whenever $|\alpha|$ is sufficiently small with $0 < \alpha \le \beta$ or $0 > \alpha \ge \beta$. Then, via the isometric transformation in Theorem 2.1 (iv), one can define a geodesic shortest curve joining $A_{\alpha,\beta}, B_{\alpha,\beta}$ given in (3.3) with respect to $K^{N^{2(1-\beta)}}$ by

$$\gamma_{A_{\alpha,\beta},B_{\alpha,\beta}}(t) := \left(\frac{\beta}{\alpha}\right)^{1/\beta} (\gamma_\alpha(t))^{\alpha/\beta}, \qquad 0 \le t \le 1,$$

so that γ_α is conversely written as (3.6). What remains to prove is

$$\lim_{\alpha\to 0} \gamma_\alpha(t) = \gamma_*(t) := \exp((1-t)\log A + t\log B), \qquad 0 \le t \le 1. \tag{3.9}$$

By Theorems 1.1 and 3.3 (a) we have

$$\begin{aligned}\frac{1}{|\alpha|}\|\gamma_\alpha(t)^\alpha - I\| = \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(\gamma_\alpha(t), I) &\le \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A, B) + \delta_{N_{\alpha,\beta}^{2(1-\alpha)}}(A, I) \\ &\longrightarrow \|\log A - \log B\|_{\mathrm{HS}} + \|\log A\|_{\mathrm{HS}} \quad \text{as } \alpha \to 0.\end{aligned}$$

Hence as in the proofs of Proposition 1.2 and Theorem 3.3 (a), one can choose a $\delta > 0$ with $\delta < |\beta|$, an $R > 1$, and $C > 0$ such that if $0 < \alpha < \delta$ $(< \beta)$ or $0 > \alpha > -\delta$ $(> \beta)$ then

$$R^{-1}I \le \gamma_\alpha(t) \le RI, \quad \|\gamma_\alpha'(t)\|_{\mathrm{HS}} \le C, \qquad 0 \le t \le 1. \tag{3.10}$$

To prove (3.9), it suffices to show that $\gamma_\alpha' \to \gamma_*'$ in $L^1([0,1];\mathbb{H}_n)$ in the weak topology. Since $\{\gamma_\alpha' : \alpha \in (0,\delta) \text{ or } (-\delta,0)\}$ is relatively compact in the weak topology, we may show that if $\eta \in L^1([0,1];\mathbb{H}_n)$ is a limit point of $\{\gamma_\alpha'\}$ as $\alpha \to 0$ in the weak topology, then $\eta = \gamma_*'$. So, let $\alpha(k)$ be a sequence in $(0,\beta]$ or $[\beta,0)$ with $\alpha(k) \searrow 0$ or $\alpha(k) \nearrow 0$ such that $\gamma_{\alpha(k)}' \to \eta$ in the weak topology. Define $\gamma(t) := A + \int_0^t \eta(s)\,ds$ for $t \in [0,1]$. Then we have

$$\gamma(t) = \lim_{k\to\infty}\left(A + \int_0^t \gamma_{\alpha(k)}'(s)\,ds\right) = \lim_{k\to\infty}\gamma_{\alpha(k)}(t), \qquad 0 \le t \le 1. \tag{3.11}$$

From this and (3.10) it follows that

$$R^{-1} \le \gamma(t) \le RI, \qquad t \in [0,1], \tag{3.12}$$

and hence $\gamma(t) \in \mathbb{P}_n$ for all $t \in [0,1]$. Now, to prove that $\gamma = \gamma_*$, we proceed as in the proof of Proposition 1.2 with slight complication.

Write $\gamma_k := \gamma_{\alpha(k)}$, $\phi_k(x,y) := N_{\alpha(k),\beta}(x,y)^{2(1-\alpha(k))}$, and $\phi(x,y) := M_{\mathrm{L}}(x,y)^2$ for short. Thanks to (3.10) we estimate

$$\begin{aligned}
&\left\langle \gamma_k'(t), \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1}\gamma_k'(t)\right\rangle_{\mathrm{HS}} \\
&\quad = \left\langle \gamma_k'(t), \left(\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right)\gamma_k'(t)\right\rangle_{\mathrm{HS}} \\
&\qquad + \left\langle \gamma_k'(t), \left(\phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} - \phi_k(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right)\gamma_k'(t)\right\rangle_{\mathrm{HS}} \\
&\qquad + \left\langle \gamma_k'(t), \phi_k(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\gamma_k'(t)\right\rangle_{\mathrm{HS}} \\
&\quad \le C^2 \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\| \\
&\qquad + C^2 \left\|\phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} - \phi_k(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\| \\
&\qquad + K^{\phi_k}_{\gamma_k(t)}(\gamma_k'(t), \gamma_k'(t))
\end{aligned}$$

and hence

$$\begin{aligned}
\left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2}\gamma_k'(t)\right\|_{\mathrm{HS}} &\le C \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\|^{1/2} \\
&\quad + C \left\|\phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1} - \phi_k(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\|^{1/2} \\
&\quad + \sqrt{K^{\phi_k}_{\gamma_k(t)}(\gamma_k'(t), \gamma_k'(t))}, \qquad 0 \le t \le 1. \qquad (3.13)
\end{aligned}$$

It follows from (3.12) that $\left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2}\right\| \le R^2$ so that a bounded operator $\mathbf{A}$ on $L^1([0,1]; \mathbb{H}_n)$ is defined by

$$(\mathbf{A}f)(t) := \phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2} f(t), \qquad f \in L^1([0,1]; \mathbb{H}_n),\ 0 \le t \le 1.$$

Since $\mathbf{A}\gamma_k' \to \mathbf{A}\gamma'$ in the weak topology, we have

$$\begin{aligned}
L_{K^\phi}(\gamma) &= \int_0^1 \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2}\gamma'(t)\right\|_{\mathrm{HS}} dt \\
&\le \liminf_{k\to\infty} \int_0^1 \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1/2}\gamma_k'(t)\right\|_{\mathrm{HS}} dt. \qquad (3.14)
\end{aligned}$$

Since we have by (3.10) and (3.12)

$$\left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\| \le 2R^2$$

and by (3.11)

$$\lim_{k\to\infty} \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\| = 0, \qquad 0 \le t \le 1,$$

the dominated convergence theorem implies that

$$\lim_{k\to\infty} \int_0^1 \left\|\phi(\mathbf{L}_{\gamma(t)}, \mathbf{R}_{\gamma(t)})^{-1} - \phi(\mathbf{L}_{\gamma_k(t)}, \mathbf{R}_{\gamma_k(t)})^{-1}\right\|^{1/2} dt = 0. \qquad (3.15)$$

It follows from (3.10) and Lemma 3.2 (a) that

$$\begin{aligned}
&\big\|\phi(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}-\phi_k(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}\big\| \\
&\quad\le \max_{x,y\in[R^{-1},R]}\frac{1}{\phi_k(x,y)}\left|\frac{\phi_k(x,y)}{\phi(x,y)}-1\right| \\
&\quad\le R^{2(1+|\beta|)}\max_{x,y\in[R^{-1},R]}\left|\frac{N_{\alpha(k),\beta}(x,y)^{2(1-\alpha(k))}}{M_{\mathrm{L}}(x,y)^2}-1\right|\longrightarrow 0\quad\text{as } k\to\infty
\end{aligned}$$

so that

$$\int_0^1\big\|\phi(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}-\phi_k(\mathbf{L}_{\gamma_k(t)},\mathbf{R}_{\gamma_k(t)})^{-1}\big\|^{1/2}\,dt=0. \tag{3.16}$$

Furthermore, Theorem 3.3 (a) implies that

$$\int_0^1\sqrt{K^{\phi_k}_{\gamma_k(t)}(\gamma_k'(t),\gamma_k'(t))}\,dt=\delta_{\phi_k}(A,B)\longrightarrow\delta_\phi(A,B)\quad\text{as } k\to\infty. \tag{3.17}$$

Combining (3.13)–(3.17) altogether implies that $L_\phi(\gamma)\le\delta_\phi(A,B)$. Note that $\gamma(t)$ has constant speed almost everywhere as shown in the last part of the proof of Proposition 1.2. Since $\gamma_*(t)$ is a unique (up to parametrization) geodesic shortest curve joining A,B with respect to K^ϕ by [10, Theorem 2.1], we obtain $\gamma=\gamma_*$ and hence $\eta=\gamma_*'$ as we desired.

(b) First, by [10, Theorem 3.1] note that we always have a geodesic shortest curve joining A,B with respect to K^{N^2} for any $N\in\mathfrak{M}_0$. The proof is similar to that of (a) based on Lemma 3.2 (b) and Theorem 3.3 (b). The details are left to the reader. □

Remark 3.6. The convergence formulas for geodesic shortest curves in (3.7) and (3.8) may be considered as variants of the so-called *Lie-Trotter formula*. As another variant, it is also known [9, Theorem 4.11] that if σ is an operator mean with $s:=f'(1)$ for the corresponding operator monotone function f (see [13]), then

$$\lim_{\kappa\searrow 0}(A^\kappa\,\sigma\,B^\kappa)^{1/\kappa}=\exp((1-s)\log A+s\log B)$$

for all $A,B\in\mathbb{P}_n$. (In fact, this was proved in [9] in the infinite-dimensional operator setting.)

Example 3.7. (a) The Stolarsky mean S_α is $N_{\alpha,\beta}$ for $\beta=1$. In [10, Theorem 2.1] we showed that, for every $\alpha\ne 0$, the geodesic distance between A,B with respect to $K^{S_\alpha^{2(1-\alpha)}}$ is

$$\delta_{S_\alpha^{2(1-\alpha)}}(A,B)=\frac{1}{|\alpha|}\|A^\alpha-B^\alpha\|_{\mathrm{HS}}$$

and a unique geodesic shortest curve joining A,B is

$$\gamma_{A,B}(t)=((1-t)A^\alpha+tB^\alpha)^{1/\alpha}.$$

We have

$$\lim_{\alpha\to 0}\frac{1}{|\alpha|}\|A^\alpha-B^\alpha\|_{\mathrm{HS}}=\|\log A-\log B\|_{\mathrm{HS}},$$

$$\lim_{\alpha\to 0}((1-t)A^\alpha + tB^\alpha)^{1/\alpha} = \exp((1-t)\log A + t\log B),$$

which are special cases of Theorems 3.3 (a) and 3.5 (a).

(b) When $N = M_{\mathrm{G}}$ is the geometric mean, $K^{M_{\mathrm{G}}^2}$ is the Riemannian metric discussed in [17, 22, 18, 14, 16, 5] and for $\kappa \neq 0$

$$N_\kappa(x,y) = \frac{\kappa(x-y)}{x^\kappa - y^\kappa}\,(xy)^{\kappa/2}, \qquad x,y>0.$$

In [10, Theorem 3.3] we showed that, for every $\kappa > 0$,

$$\delta_{N_\kappa^2}(A,B) = \frac{1}{\kappa}\delta_{M_{\mathrm{G}}}(A^\kappa, B^\kappa) = \left\|\log(A^{-\kappa/2}B^\kappa A^{-\kappa/2})^{1/\kappa}\right\|_{\mathrm{HS}}$$

and a unique shortest curve joining A, B is

$$\gamma_{A,B}(t) = (A^\kappa \,\#_t\, B^\kappa)^{1/\kappa},$$

where $\#_t$ is the tth power mean or the weighted geometric mean given in (0.2). We have

$$\lim_{\kappa\searrow 0}\left\|\log(A^{-\kappa/2}B^\alpha A^{-\kappa/2})^{1/\kappa}\right\|_{\mathrm{HS}} = \|\log A - \log B\|_{\mathrm{HS}} \quad \text{(decreasingly)}$$

and

$$\lim_{\kappa\searrow 0}(A^\kappa \,\#_t\, B^\kappa)^{1/\kappa} = \exp((1-t)\log A + t\log B), \tag{3.18}$$

which are special cases of Theorems 3.3 (b) (also Proposition 3.4) and 3.5 (b). This is also an example of the convergence formula in Remark 3.6.

Problem 3.8. Let δ denote the geodesic distance on $\mathbb{P}_n$ with respect to the Riemannian metric $K^{M_{\mathrm{G}}^2}$. Note that $\mathbb{P}_n$ with this δ is an example of so-called NPC spaces (non-positively curved metric spaces), whose theory has recently been developed extensively, as seen in [24]. Let $\mathbf{w} = (w_1, \dots, w_m)$ be a weight vector, i.e., $w_i \ge 0$ and $\sum_{i=1}^m w_i = 1$ with $m \ge 2$. Given m matrices $A_1, \dots, A_m \in \mathbb{P}_n$, the *weighted geometric mean* $G_{\mathbf{w}}(A_1, \dots, A_m)$ was introduced in [15] as a unique minimizer of the weighted sum of the squares of distance $\sum_{i=1}^m w_i\delta^2(X, A_i)$ for $X \in \mathbb{P}_n$. (Unique existence of the minimizer is a general result in NPC spaces.) In particular, when $m = 2$, $G_{(1-t,t)}(A,B) = A \,\#_t\, B$ for all $t \in [0,1]$ and $A, B \in \mathbb{P}_n$, and the non-weighted m-variable geometric mean is the case for $\mathbf{w} = (1/m, \dots, 1/m)$. It may naturally be conjectured that convergence (3.18) extends to the m-variable geometric mean, i.e.,

$$\lim_{\kappa\searrow 0}\left(G_{\mathbf{w}}(A_1^\kappa, \dots, A_m^\kappa)\right)^{1/\kappa} = \exp\left(\sum_{i=1}^m w_i \log A_i\right). \tag{3.19}$$

The above right-hand side is the Log-Euclidean mean discussed in [3] in comparison with the m-variable geometric mean (called the affine-invariant mean there). An interesting problem in connection with discussions in [3] is to extend the log-majorizations for the 2-variable weighted geometric mean in [2] to the m-variable case:

$$G_{\mathbf{w}}(A_1, \dots, A_m) \prec_{(\log)} \left(G_{\mathbf{w}}(A_1^\kappa, \dots, A_m^\kappa)\right)^{1/\kappa}, \qquad 0 < \kappa < 1. \tag{3.20}$$

(See [2] for some details on log-majorization.) The two conjectures (3.19) and (3.20) would imply that $G_{\mathbf{w}}(A_1, \dots, A_m) \prec_{(\log)} \exp\bigl(\sum_{i=1}^m w_i \log A_i\bigr)$.

Acknowledgments

This work was partially supported by the Grant-in-Aid for Scientific Research (C)21540208 (FH) and the Hungarian Research Grant OTKA T068258 (DP). The authors thank the referee for his helpful comments and for calling their attention to the papers [3, 6, 15].

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* **26** (1979), 203–241.

[2] T. Ando and F. Hiai, Log majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.* **197/198** (1994), 113–131.

[3] V. Arsigny, P. Fillard, X. Pennec and N. Ayache, Geometric means in a novel vector space structure on symmetric positive-definite matrices, *SIAM J. Matrix Anal. Appl.* **29** (2007), 328–347.

[4] R. Bhatia, *Positive Definite Matrices*, Princeton Univ. Press, Princeton, 2007.

[5] R. Bhatia and J. A. R. Holbrook, Riemannian geometry and matrix geometric means, *Linear Algebra. Appl.* **423** (2006), 594–618.

[6] R. Bhatia and R. L. Karandikar, Monotonicity of the matrix geometric mean, *Math. Ann.*, to appear.

[7] R. Bhatia and H. Kosaki, Mean matrices and infinite divisibility, *Linear Algebra. Appl.* **424** (2007), 36–54.

[8] J. I. Fujii, M. Fujii, T. Miura, H. Takagi and S.-E. Takahasi, Continuously differentiable means, *J. Inequal. Appl.* **2006**, Art. ID 75941, 15 pp.

[9] F. Hiai, Log-majorizations and norm inequalities for exponential operators, in *Linear Operators*, J. Janas, F. H. Szafraniec and J. Zemánek (eds.), Banach Center Publications, Vol. 38, 1997, pp. 119–181.

[10] F. Hiai and D. Petz, Riemannian metrics on positive definite matrices related to means, *Linear Algebra Appl.* **430** (2009), 3105–3130.

[11] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[12] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. 1, Wiley Interscience, New York-London, 1963.

[13] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.* **246** (1980), 205–224.

[14] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly* **108** (2001) 797–812.

[15] J. Lawson and Y. Lim, Monotonic properties of the least squares mean, *Math. Ann.* **351** (2011), 267–279.

[16] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive definite matrices, *SIAM J. Matrix Anal. Appl.* **26** (2005), 735–747.

[17] G. D. Mostow, *Strong Rigidity of Locally Symmetric Spaces*, Annals Math. Studies 78, Princeton University Press, Princeton, 1973.

[18] A. Ohara, N. Suda and S. Amari, Dualistic differential geometry of positive definite matrices and its applications to related problems, *Linear Algebra Appl.* **247** (1996), 31–53.

[19] P. Petersen, *Riemannian Geometry*, Second edition, Springer-Verlag, 2006.

[20] D. Petz, Monotone metrics on matrix spaces, *Linear Algebra Appl.* **244** (1996), 81–96.

[21] W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.* **8** (1975) 159–170.

[22] L. T. Skovgaard, A Riemannian geometry of the multivariate normal model, *Scand. J. Statistics*, **11** (1984), 211–223.

[23] K. B. Stolarsky, Generalizations of the logarithmic mean, *Math. Mag.* **48** (1975), 87–92.

[24] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, in: *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces*, P. Auscher et. al. (eds.), Contemp. Math., Vol. 338, Amer, Math. Soc., Providence, 2003.