Linear Algebra and its Applications 434 (2011) 542–558

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Path of quasi-means as a geodesic

Jun Ichi Fujii

Department of Arts and Sciences (Information Science), Osaka Kyoiku University, Asahigaoka, Kashiwara, Osaka 582-8582, Japan

ARTICLE INFO

Article history:
Received 26 May 2010
Accepted 7 September 2010
Available online 13 October 2010

Submitted by R.A. Brualdi

AMS classification:
15A48
15A60
47A64
53B40

Keywords:
Finsler geometry
Quasi-arithmetic mean
Quasi-geometric mean
Geodesic
Positive definite matrices
Operator mean
Unitarily invariant norm

ABSTRACT

As a generalization of the Hiai–Petz geometries, we discuss two types of them where the geodesics are the quasi-arithmetic means and the quasi-geometric means, respectively. Each derivative of such a geodesic might determine a new relative operator entropy. Also in these cases, the Finsler metric can be induced by each unitarily invariant norm. If the norm is strictly convex, then the geodesic is the shortest. We also give examples of the shortest paths which are not the geodesics when the Finsler metrics are induced by the Ky Fan k-norms.

1. Introduction

In [10,11], we introduced the Uhlmann transformation from operator means in the sense of Kubo and Ando [20] onto the derivative solidarities: let $A\, m_t B$ be an interpolational path, that is, a path from $A = A\, m_0 B$ to $B = A\, m_1 B$ of the Kubo–Ando operator means satisfying

$$A\, m_{(1-r)p+rq} B = (A\, m_p B) m_r (A\, m_q B)$$

for all $p, q, r \in [0, 1]$. Then we showed in [11] that this path is differentiable for t and the limit

E-mail address: fujii@cc.osaka-kyoiku.ac.jp

0024-3795/$ - see front matter
doi:10.1016/j.laa.2010.09.011

$$A\, s_m B \equiv \left.\frac{d(A\, m_t B)}{dt}\right|_{t=0} = \lim_{t\to 0}\frac{A\, m_t B - A}{t}$$

is called the derivative solidarity for m. In particular, if m_t is the path of geometric operator means

$$A\,\#_t\,B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}},$$

then the derivative solidarity is the relative operator entropy [9,10]

$$S(A|B) = A^{\frac{1}{2}}\log\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)A^{\frac{1}{2}}.$$

More generally, the derivative for a path $m_{r,t}$ defined by

$$A\,\#_{r,t}\,B = A^{\frac{1}{2}}\left(1-t+t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right)^{1/r}A^{\frac{1}{2}}$$

gives the Tsallis relative operator entropy in the sense of [26]

$$\lim_{t\to 0}\frac{A\,\#_{r,t}\,B - A}{t} = T_r(A|B) \equiv \frac{A\,\#_r\,B - A}{r},$$

see also [9,10,11]. This construction of the relative entropy is due to Uhlmann [24] and hence we call the map $m_t \mapsto s_m$ the *Uhlmann transformation* in [9]. These discussions are reduced to the representing functions $f_t(x) = 1\, m_t x$ and $F(x) = 1\, s_m x$ by virtue of the 'transformer equality' of the Kubo–Ando mean; If X is invertible, then

$$X^*(A\, m\, B)X = (X^*AX)m(X^*BX).$$

A remarkable property of the representing function F is that F is a strictly increasing smooth function with $F(1) = 0$ and $F'(1) = 1$ which follow from $f_t(1) = 1$ and $f_0(x) = 1$. (Here the strict increasingness means $F'(x) > 0$ for all $x > 0$, which is denoted by $F' > 0$).

But it would be hard to find the interpolational paths in the Kubo–Ando means except $m_{r,t}$ in spite of the importance of them. Recently Hiai and Petz [17] introduced two types of parametrized geometries whose geodesics are

$$((1-t)A^r + tB^r)^{\frac{1}{r}} \quad\text{and}\quad (A^\alpha\,\#_t\,B^\alpha)^{\frac{1}{\alpha}}$$

for $r \in \mathbb{R}$ and $\alpha > 0$. They are the interpolational paths but not the ones of the Kubo–Ando means.

Based on this, we extend the Hiai–Petz geometries and discuss the induced entropies as their derivatives.

2. Path of quasi-arithmetic means

Throughout this section, a *numerical mean* is a binary operation $\mathfrak{m}$ on positive numbers which satisfies only the following conditions:

$$\textit{normalization}:\quad a\,\mathfrak{m}\,a = a,$$

$$\textit{monotonicity}:\quad a\,\mathfrak{m}\,b \leqq a'\,\mathfrak{m}\,b' \text{ whenever } a \leqq a' \text{ and } b \leqq b'.$$

Moreover we assume here that $a\,\mathfrak{m}\,b$ is a smooth function for each term. But the homogeneity; $ta\,\mathfrak{m}\,tb = t(a\,\mathfrak{m}\,b)$ for all $t > 0$, is not assumed here.

Let $a\,\mathfrak{m}_t b$ be a (smooth) path of numerical means from $a = a\,\mathfrak{m}_0 b$ to $b = a\,\mathfrak{m}_1 b$. Though the homogeneity does not always hold, we still consider the *path function* $f_t(x) = 1\,\mathfrak{m}_t x$. A path is also called *interpolational* if

$$a\,\mathfrak{m}_{(1-r)p+rq}\,b = (a\,\mathfrak{m}_p\,b)\,\mathfrak{m}_r\,(a\,\mathfrak{m}_q\,b)$$

holds for all $p, q, r \in [0, 1]$. In this case, we can define $a\,s_{\mathfrak{m}} b = \left.\frac{d(a\,\mathfrak{m}_t b)}{dt}\right|_{t=0}$, which is no longer positive in general. First we confirm the properties of the derivative function $F_{\mathfrak{m}}(x) = 1\,s_{\mathfrak{m}} x$ (cf. [11]):

Lemma 2.1. *If $a\,\mathfrak{m}_t b$ is a smooth interpolational path of numerical means, then the derivative function $F_\mathfrak{m}$ satisfies $F_\mathfrak{m}(1) = 0$ and*

$$(a\,\mathfrak{m}_t b) s_\mathfrak{m} (a\,\mathfrak{m}_{t+h} b) = h \frac{d(a\,\mathfrak{m}_t b)}{dt}$$

for $h \geqq q$ with $t + h \leqq 1$. In particular,

$$F_\mathfrak{m}(f_t(x)) = tF_\mathfrak{m}(x).$$

In addition, if $F_\mathfrak{m}$ is a nonzero function, then $F'_\mathfrak{m}(1) = 1$.

Proof. It follows from $1\,\mathfrak{m}_t 1 = 1$ that $F_\mathfrak{m}(1) = 0$. The monotonicity shows that $F_\mathfrak{m}$ is increasing. From the interpolational property,

$$\begin{aligned}(a\,\mathfrak{m}_t b)\, s_\mathfrak{m} (a\,\mathfrak{m}_{t+h} b) &= \lim_{r\to 0} \frac{(a\,\mathfrak{m}_t b)\,\mathfrak{m}_r (a\,\mathfrak{m}_{t+h} b) - (a\,\mathfrak{m}_t b)}{r} \\ &= \lim_{r\to 0} \frac{a\,\mathfrak{m}_{(1-r)t+r(t+h)} b - (a\,\mathfrak{m}_t b)}{r} \\ &= \lim_{r\to 0} h \frac{a\,\mathfrak{m}_{t+rh} b - (a\,\mathfrak{m}_t b)}{rh} = h \frac{d(a\,\mathfrak{m}_t b)}{dt}.\end{aligned}$$

Putting $t = 0$ and then using t instead of h, we have

$$a\, s_\mathfrak{m} (a\,\mathfrak{m}_t\, b) = t(a\, s_\mathfrak{m} b)$$

and hence $F_\mathfrak{m}(f_t(x)) = tF_\mathfrak{m}(x)$ for $f_t(x) = 1\ \mathfrak{m}_t x$. It follows that

$$F_\mathfrak{m}(x) = \frac{F_\mathfrak{m}(f_t(x)) - F_\mathfrak{m}(1)}{t} = \frac{F_\mathfrak{m}(f_t(x)) - F_\mathfrak{m}(1)}{f_t(x) - 1} \frac{f_t(x) - 1}{t} \to F'_\mathfrak{m}(1) F_\mathfrak{m}(x)$$

as $t \to 0$. Therefore we have $F'_\mathfrak{m}(1) = 1$ when $F_\mathfrak{m}(x) \neq 0$ for some x by the additional assumption. □

We also call a map $f_t \mapsto F_\mathfrak{m}$ the *Uhlmann transformation* here. If $F_\mathfrak{m}$ is strictly increasing, then the 'reproducing equality' holds (see also the formula (0′) in the below):

$$f_t(x) = F_\mathfrak{m}^{-1}(tF_\mathfrak{m}(x)) = F_\mathfrak{m}^{-1}((1-t)F_\mathfrak{m}(1) + tF_\mathfrak{m}(x)).$$

In this paper, only for a strictly increasing smooth function F on $(0, \infty)$ with $F' > 0$, we use the term '*quasi-arithmetic mean*'

$$a\,\mathfrak{m}_t b = F^{-1}((1-t)F(a) + tF(b)). \tag{0}$$

Note that the assumption $F' > 0$ assures the continuous differentiability of the mean, see [15]. Here we call a strictly increasing smooth function F a *fundamental function* if $F(1) = 0$, $F' > 0$ and $F'(1) = 1$. In fact, each quasi-arithmetic mean has the *standard form* (0) for a fundamental function F since the invariant property

$$a\,\mathfrak{m}_t b = F^{-1}((1-t)F(a) + tF(b)) = F_{\alpha,\beta}^{-1}((1-t)F_{\alpha,\beta}(a) + tF_{\alpha,\beta}(b))$$

holds for an affine transform $F_{\alpha,\beta}(x) = (F(x) - \alpha)/\beta$ of F for $\alpha \in \mathbb{R}$ and $\beta > 0$.

Then we have the inverse of the Uhlmann transformation in these means: each quasi-arithmetic mean has the standard form

$$a\,\mathfrak{m}_t b = F_\mathfrak{m}^{-1}((1-t)F_\mathfrak{m}(a) + tF_\mathfrak{m}(b)). \tag{0′}$$

Theorem 2.2. *A path of quasi-arithmetic means defined by (0) is interpolational and the Uhlmann transformation $f_t \mapsto F_\mathfrak{m}$ is a bijection from path functions of quasi-arithmetic means onto the fundamental functions, precisely $F_\mathfrak{m} = F$.*

Proof. This path is interpolational since

$$\begin{aligned}(a\,\mathfrak{m}_p b)\,\mathfrak{m}_r(a\,\mathfrak{m}_q b) &= F^{-1}((1-p)F(a)+pF(b))\,\mathfrak{m}_r F^{-1}((1-q)F(a)+qF(b))\\ &= F^{-1}((1-r)((1-p)F(a)+pF(b))+r((1-q)F(a)+qF(b)))\\ &= F^{-1}([1-(1-r)p-rq)]F(a)+[(1-r)p+rq]F(b))\\ &= a\,\mathfrak{m}_{(1-r)p+rq}b.\end{aligned}$$

By $F(1)=0$ and $F'(1)=1$, we have

$$\begin{aligned}F_{\mathfrak{m}}(x) \equiv 1\, s_{\mathfrak{m}} x &= \lim_{t\to 0}\frac{1\,\mathfrak{m}_t x - 1}{t} = \lim_{t\to 0}\frac{F^{-1}(tF(x))-1}{t}\\ &= \left.\frac{\partial F^{-1}(tF(x))}{\partial t}\right|_{t=0} = \frac{F(x)}{F'(F^{-1}(0))} = \frac{F(x)}{F'(1)} = F(x).\end{aligned}$$

Thus we have that the Uhlmann transformation is bijective. □

For a fundamental function F, we also define a path of *quasi-arithmetic operator means* for positive invertible operators A and B on a Hilbert space:

$$A\,\mathfrak{m}_t B \equiv A\,\mathfrak{m}_{F,t} B = F^{-1}((1-t)F(A)+tF(B)). \tag{1}$$

Immediately we have $A\,\mathfrak{m}_t A = A$, but this mean is not a Kubo–Ando operator mean. Moreover it is not always chaotic operator mean in [12].

Here, for a monotone increasing function G, we define a *G-order* $A \underset{G}{\le} B$ by $G(A) \le G(B)$ for all positive operators A and B in the usual order of operators. We also call H is *G-monotone* if the composition $G\circ H$ is operator monotone. Thus a chaotic operator mean (e.g. the case $F(x)=x^r$ for $r\in[-1,1]$) is a chaotic (i.e., log-monotone) operator mean. Then we obtain properties of the path of quasi-arithmetic means like numerical case:

Theorem 2.3. *A path $A\,\mathfrak{m}_t B$ of quasi-arithmetic means defined by (1) is interpolational and the derivative $A\,s_{\mathfrak{m}} B$ has the representing function F itself. In addition, if F is operator monotone and F^{-1} is G-monotone, then $\mathfrak{m}_t$ has a monotone property: If $A \le A'$ and $B \le B'$, then $A\,\mathfrak{m}_t B \underset{G}{\le} A'\,\mathfrak{m}_t B'$.*

In the next section, we give the derivative $\frac{d(A\,\mathfrak{m}_t B)}{dt}$, which is a kind of entropy, in the discussion of a geometry.

3. QAM geometry

Now we observe a geometry of the $n\times n$ positive definite matrices $\mathcal{M}^+$ with a geodesic of a path of quasi-arithmetic means $A\,\mathfrak{m}_t B$, which is called the *QAM geometry* here. From now on, for $A\in\mathcal{M}^+$ and a smooth path $\gamma(t)$ on $\mathcal{M}^+$, U (resp. U_t) is assumed to be any unitary such that U^*AU (resp. $U_t^*\gamma(t)U_t$) is a diagonal matrix D (resp. D_t) with entries d_j (resp. $d_j(t)$). Let F be a strictly increasing smooth function on $(0,\infty)$ and $F^{[1]}$ the divided difference:

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

As in [18,14], we observe the following differential formula:

$$\frac{dF(\gamma(t))}{dt} = U_t((F^{[1]}(d_i(t),d_j(t)))\circ U_t^*\dot{\gamma}(t)U_t)U_t^*. \tag{2}$$

Immediately we have the inverse formula:

$$\dot{\gamma}(t) = U_t \left(\left(\frac{1}{F^{[1]}(d_i(t), d_j(t))} \right) \circ U_t^* \frac{dF(\gamma(t))}{dt} U_t \right) U_t^*. \tag{2'}$$

For each $A \in \mathcal{M}^+$, define the linear transform on $\mathcal{M}^h$ by

$$\Phi_A(X) = U((F^{[1]}(d_i, d_j)) \circ U^* X U) U^*.$$

Here we introduce a geometry of fiber bundles to obtain the geodesic as an auto-parallel curve. For this, we should define a 'parallelism' between tangent vectors at distinct points in $\mathcal{M}^+$. Each fiber of a principal fiber bundle is considered as the transformation group of each tangent space. A 'connection' determines a relation between transformations in different tangent spaces and defines a parallelism, that is, an affine connection (or a covariant derivative) on tangent vector bundles as an associated bundle. Following such procedure, we can reach the equation to get the geodesic between two points in $\mathcal{M}^+$. It is a simple way to obtain the geodesic without complex calculations introduced by Cartan as in [19].

Now consider the trivial principal bundle $\mathcal{P} = \mathcal{M}^+ \times \mathcal{U}$ for $\mathcal{M}^+$ with the trivial projection $\pi((A, V)) = A$. Define the left action $\Psi((A, V))X = \Phi_A^{-1}(VXV^*)$ of $\mathcal{P}$ on $T_A\mathcal{M}^+ = \mathcal{M}^h$. Here we observe the associated tangent vector bundle

$$\mathcal{P} \times \mathcal{M}^h / \mathcal{U} = \mathcal{P} \underset{\rho}{\times} \mathcal{M}^h,$$

where each fiber is the hermitian matrices $\mathcal{M}^h$ with the right action $(A, V)W = (A, VW)$ of $W \in \mathcal{U}$ on $\pi^{-1}(A) \subset \mathcal{P}$ and the left action $\rho(W)X = WXW^*$ on the tangent space $T_A\mathcal{M}^+ = \mathcal{M}^h$. Then it can be identified with the tangent vector bundle $\mathcal{M}^h$ by the map $((A, V), X) \mapsto \Psi((A, V))(X)$ since

$$\begin{aligned} \Psi((A, V)W)\rho^{-1}(W)X &= \Psi((A, VW))W^*XW = \Phi_A^{-1}(VWW^*XWW^*V^*) \\ &= \Phi_A^{-1}(VXV^*) = \Psi((A, V))(X). \end{aligned}$$

This identification shows that we can determine the parallel displacement of tangent vectors along the curve γ by the connection of $\mathcal{P}$ and a horizontal lift of γ as in the below, see also [19]. Consider the *canonical flat connection*, that is, the horizontal vector in the tangent space $\mathcal{T}_{(A,V)}(\mathcal{P})$ is of the form (X, O) for some $X = X^*$. Then the *horizontal lift* Γ of a path γ is $\Gamma(t) = (\gamma(t), V)$ for any fixed $V \in \mathcal{U}$ (Once we adopt the starting point $\Gamma(0) = (\gamma(0), V)$ in the fiber $\pi^{-1}(\gamma(0))$, it is the integral curve under the tangent vectors $(\dot{\gamma}(t), O)$ of the required curve Γ for each t).

Since a tangent vector $Y \in \mathcal{M}^h$ also belongs to the associated bundle $\mathcal{M}^h$ of $\mathcal{P}$ and

$$\Psi((A, V))^{-1}Y = V^*\Phi_A(Y)V,$$

we have that the *parallel displacement* $P_t = P_t^0$ *from 0 to t along a path* γ of a tangent vector X on $\gamma(0)$ is obtained by

$$P_tX = \Psi((\gamma(t), V))(\Psi((\gamma(0), V))^{-1}X) = \Phi_{\gamma(t)}^{-1}(\Phi_{\gamma(0)}(X)).$$

Then the *covariant derivative* for a vector field $X(t)$ is

$$\begin{aligned} \nabla_{\dot{\gamma}}X &= \lim_{\varepsilon \to 0} \frac{P_t^{t+\varepsilon}X(t+\varepsilon) - X(t)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{\Phi_{\gamma(t)}^{-1}(\Phi_{\gamma(t+\varepsilon)}(X(t+\varepsilon))) - X(t)}{\varepsilon} \\ &= \Phi_{\gamma(t)}^{-1}\left(\lim_{\varepsilon \to 0} \frac{\Phi_{\gamma(t+\varepsilon)}(X(t+\varepsilon)) - \Phi_{\gamma(t)}(X(t))}{\varepsilon} \right) = \Phi_{\gamma(t)}^{-1}\left(\left(\Phi_{\gamma(t)}(X(t)) \right)' \right). \end{aligned}$$

Theorem 3.1. *The geodesic γ_F from A to B in $\mathcal{M}^+$ for the canonical flat connection in $\mathcal{P} = \mathcal{M}^+ \times \mathcal{U}$ is the path of quasi-arithmetic operator means*

$$\gamma_F(t) = A\, \mathfrak{m}_t B = F^{-1}((1-t)F(A) + tF(B)).$$

Proof. By the above argument, the geodesic equation $\nabla_{\dot\gamma}\dot\gamma = O$ implies

$$O = \Phi_{\gamma(t)}(\nabla_{\dot\gamma}\dot\gamma) = \left(\Phi_{\gamma(t)}(\dot\gamma(t))\right)' = (F(\gamma(t)))''.$$

So there exist $C_1 \in \mathcal{M}^h$ and $C_2 \in \mathcal{M}^+$ with $F(\gamma(t)) = tC_1 + C_2$. Since

$$F(A) = F(\gamma(0)) = C_2 \quad \text{and} \quad F(B) = F(\gamma(1)) = C_1 + C_2,$$

we have $C_2 = F(A)$ and $C_1 = F(B) - F(A)$, so that $\gamma(t) = A\,\mathfrak{m}_{r,t}B$. $\square$

Now we define a Finsler metric by the Minkovski norm in the sense of Cartan [21,22], that is, it is equivalent to the original norm and satisfies the Finsler condition $L(X;\gamma(0)) = L(P_tX;\gamma(t))$ for all paths γ:

Theorem 3.2. *For any unitarily invariant norm $\||\ \||$, the norm of $X \in \mathcal{M}^h$ defined as*

$$L(X;A) \equiv L_{F,QAM}(X;A) \equiv \||\Phi_A(X)\|| = \||(F^{[1]}(d_i,d_j)) \circ U^*XU\||$$

is a Finsler metric and the geodesic length $d(A,B)$ makes $\mathcal{M}^+$ a metric space;

$$d(A,B) = \int_0^1 L(\dot\gamma_F(t);\gamma_F(t))dt = \||F(B) - F(A)\||.$$

Proof. The equivalence for norms is clear since the space is finite dimensional. The translation

$$\Phi_{\gamma(t)}(P_tX) = \Phi_{\gamma(t)}(\Phi^{-1}_{\gamma(t)}(\Phi_{\gamma(0)}(X))) = \Phi_{\gamma(0)}(X)$$

implies the Finsler condition of $L(X;A)$. Here it follows from (2) that

$$L(\dot\gamma(t);\gamma(t)) = \||\Phi_{\gamma(t)}(\dot\gamma(t))\|| = \||F(\gamma(t))'\||$$

in general. Since the geodesic γ_F from A to B is auto-parallel, we have

$$\begin{aligned}\int_0^1 L(\dot\gamma_F(t);\gamma_F(t))dt &= L(\dot\gamma_F(0);\gamma_F(0))\int_0^1 dt = L(\dot\gamma_F(0);\gamma_F(0)) \\ &= \||\Phi_A(\dot\gamma_F(0))\|| = \||F(\gamma_F(t))'\big|_{t=0}\|| = \||F(B) - F(A)\||.\end{aligned}$$

It is clear that d is a metric function. $\square$

Finally in this section, we mention a new class of relative entropy. By the differential formula (2′), we have

$$\dot\gamma_F(t) = U_t\left(\left(\frac{1}{F^{[1]}(d_i(t),d_j(t))}\right) \circ U_t^*(F(B) - F(A))U_t\right)U_t^*.$$

So we can define the *arithmetic relative operator entropy for F* as

$$\mathfrak{S}_F(A|B) = \dot\gamma_F(0) = As_{\mathfrak{m}}B = U\left(\left(\frac{1}{F^{[1]}(d_i,d_j)}\right) \circ U^*(F(B) - F(A))U\right)U^*.$$

Also we define the *arithmetic relative entropy* as

$$\mathfrak{s}_F(A|B) = \mathrm{tr}\ \mathfrak{S}_F(A|B) = \mathrm{tr}\ \left(F'(A)^{-1}(F(B) - F(A))\right).$$

In case $F = \log$, we have the minus quantity of the Umegaki entropy $s_U(A|B)$ [25] for matrices;

$$\mathfrak{s}_F(A|B) = \mathrm{tr}\ \mathfrak{S}_F(A|B) = \mathrm{tr}\ (A(\log B - \log A)) = -s_U(A|B).$$

4. Path of quasi-geometric means

Let F be an increasing positive smooth function F with $F' > 0$. Based on one of the Hiai–Petz geometries whose geodesic is $(A^\alpha \,\#_t\, B^\alpha)^{1/\alpha}$, we define the *quasi-geometric operator mean* $\#_{F,t}$ for positive invertible operators A and B on a Hilbert space as

$$A \,\#_{F,t}\, B = F^{-1}(F(A) \,\#_t\, F(B)). \tag{3}$$

This mean is invariant for a positive scalar multiplication $F \mapsto \alpha F$, so that we may assume $F(1) = 1$. Then we have a bijective correspondence between these means and such functions:

Lemma 4.1. *A map $\#_{F,t} \mapsto F$ from the quasi-geometric operator means to the smooth increasing positive function F with $F' > 0$ and $F(1) = 1$ is one-to-one.*

Proof. For strictly increasing positive smooth functions F and G with $F(1) = G(1) = 1$, assume $A \,\#_{F,t}\, B = A \,\#_{G,t}\, B$. Here we take commuting positive definite matrices A and B whose spectra are greater than 1 and put $C = \log A$ and $D = \log B$. A function $f(x) = \log F(e^x) - \log F(e)$ is smooth and satisfies $f(1) = 0$ and $f' > 0$. Then

$$F^{-1}(F(A) \,\#_t\, F(B)) = F^{-1}(F(e^C)^{1-t} F(e^D)^t) = e^{f^{-1}((1-t)f(C)+tf(D))}$$

and taking $g(x) = \log G(e^x) - \log G(e)$, we have $g(1) = 0, g' > 0$ and

$$f^{-1}((1-t)f(C) + tf(D)) = g^{-1}((1-t)g(C) + tg(D)).$$

Thus we have the two numerical quasi-arithmetic means are the same. Then, by Lemma 2.1, we have

$$gf^{-1}(ty) = tgf^{-1}(y)$$

and hence gf^{-1} is linear. Thus there exists $s > 0$ with $g(x) = sf(x)$, that is, $G(z) = F(z)^s$. Since we can take positive definite matrices X and Y such that

$$(X^s \,\#_t\, Y^s)^{\frac{1}{s}} \neq X \,\#_t\, Y$$

for $s \neq 1$, then s must be 1 and hence $f = g$. Since $F(1) = G(1) = 1$, we have

$$-\log F(e) = f(0) = g(0) = -\log G(e),$$

so that we have $F(x) = G(x)$ for all $x > 1$. Since the case $x \leqq 1$ is similarly shown, we have $F = G$. $\square$

We give a simple example for positive semi-definite matrices, but the continuity guarantees the example for $(X^s \,\#_t\, Y^s)^{\frac{1}{s}} \neq X \,\#_t\, Y$ in the above proof:

Example 1. For $\alpha, x > 0$, put

$$A = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}, \quad B = P = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad P_x = \frac{x}{x+1}\begin{pmatrix} 1 & \frac{1}{\sqrt{x}} \\ \frac{1}{\sqrt{x}} & \frac{1}{x} \end{pmatrix}.$$

Since P and P_x are projections, we have

$$\begin{aligned} A \,\#\, B &= \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} \sqrt{\frac{1}{2}\begin{pmatrix} 1 & \frac{1}{\sqrt{x}} \\ \frac{1}{\sqrt{x}} & \frac{1}{x} \end{pmatrix}} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} \sqrt{\frac{1+x}{2x} P_x} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} = \sqrt{\frac{1+x}{2x}} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} P_x \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{x} \end{pmatrix} \\ &= \sqrt{\frac{1+x}{2x}} \frac{x}{x+1} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{\sqrt{2x}}{\sqrt{x+1}} P, \end{aligned}$$

and hence

$$(A \# B)^\alpha = \frac{\sqrt{2x}^\alpha}{\sqrt{x+1}^\alpha} P.$$

On the other hand, the formula

$$A^\alpha \# B^\alpha = \frac{\sqrt{2x}^\alpha}{\sqrt{x^\alpha + 1}} P$$

shows the above is not equal if $x \neq 1$ and $\alpha > 0$.

The derivative at $t = 0$ is given by

$$\left.\frac{\partial(1 \#_{F,t} x)}{\partial t}\right|_{t=0} = \lim_{t\to 0} \frac{F^{-1}(F(x)^t)) - 1}{t} = \frac{1}{F'(F^{-1}(1))} \log F(x) = \frac{\log F(x)}{F'(1)}.$$

The derivative $\frac{d(A \#_{F,t} B)}{dt}$ are discussed in the next section.

5. QGM geometry

As a geometry of a type of the CPR one [7,8] or the Bhatia–Holbrook one [6,5], we will discuss a geometry for a manifold $\mathcal{M}^+$ where the geodesic is the quasi-geometric mean. It is a generalization of the other type of the Hiai–Petz geometry (the case $F(x) = x^\alpha$) and is called the *QGM geometry*.

For a smooth and positive function F with $F' > 0$, consider a principal fiber bundle $\mathcal{P}_F = \{\mathcal{G}, \mathcal{M}^+, \mathcal{U}, \pi_F\}$ where $\mathcal{G}$ is the regular matrices, π_F is the projection $\pi_F(G) = F^{-1}(GG^*)$ and $\mathcal{U}$ is the unitary matrices as the structure group which acts naturally from the right. Then each fiber is of the form $\pi_F^{-1}(F^{-1}(GG^*)) = G\mathcal{U}$. Define a connection by the *horizontal subspace* $G\mathcal{M}^h$ of the tangent space $T_G\mathcal{P}_F$ of $\mathcal{G}$ at G. Note that the vertical subspace is $iG\mathcal{M}^h$ which is the tangent space of the fiber $G\mathcal{U}$ at G.

Let Γ be the horizontal lift of the path γ. Then

$$\gamma = \pi_F(\Gamma) = F^{-1}(\Gamma\Gamma^*) \quad \text{and} \quad \Gamma^{-1}\dot{\Gamma} \in \mathcal{M}^h$$

and consequently

$$\Gamma\Gamma^* = F(\gamma) \quad \text{and} \quad \Gamma^{-1}\dot{\Gamma} = (\Gamma^{-1}\dot{\Gamma})^* = \dot{\Gamma}^*(\Gamma^*)^{-1}.$$

Thereby $\dot{\Gamma}\Gamma^* = \Gamma\dot{\Gamma}^*$ and hence $(F(\gamma))' = \dot{\Gamma}\Gamma^* + \Gamma\dot{\Gamma}^* = 2\dot{\Gamma}\Gamma^*$, so that we have

$$\dot{\Gamma} = \frac{1}{2}(F(\gamma))'(\Gamma^*)^{-1} = \frac{1}{2}(F(\gamma))'(\Gamma\Gamma^*)^{-1}\Gamma = \frac{1}{2}(F(\gamma))'(F(\gamma))^{-1}\Gamma,$$

which is called the *transport equation*. Thus the horizontal lift Γ is the solution of the transport differential equation though it is hard to express it explicitly.

We also use a linear transformation Φ_A on the tangent space $\mathcal{T}_A(\mathcal{M}^+) = \mathcal{M}^h$

$$\Phi_A(X) = U\left[(F^{[1]}(d_i, d_j)) \circ U^*XU\right]U^*.$$

For $G \in \mathcal{G}$, its action on the tangent vector X of $\mathcal{M}^+$ at A is defined by

$$\Theta(G)X \equiv \Theta_F(G)X = \Phi_A^{-1}(GXG^*).$$

Note that the point G can be identified with the linear map $\Theta(G)$. Then the inverse action is

$$\Theta(G)^{-1}X = G^{-1}\Phi_A(X)(G^*)^{-1}.$$

Here the associated bundle $\mathcal{P}_F \times \mathcal{M}^h/\mathcal{U} = \mathcal{P}_F \times_\rho \mathcal{M}^h$ with the fiber $\mathcal{M}^h$ and the left action $\rho(V)X = VXV^*$ of $V \in \mathcal{U}$ can be identified with the tangent vector bundle $\mathcal{M}^h$ since the equivalence relation in $\mathcal{P}_F \times \mathcal{M}^h$

$$(G, X) \sim (G, X)V = (GV, \rho(V^*)X) = (GV, V^*XV)$$

is compatible with the identification $G \mapsto \Theta(G)$;

$$\Theta(GV)V^*XV = \Phi_A^{-1}(GV(V^*XV)(GV)^*) = \Phi_A^{-1}(GXG^*) = \Theta(G)X.$$

Therefore the *parallel displacement* P_t^0 of a tangent field X along the path γ from 0 to t is defined as

$$\begin{aligned} P_t^0 X &= \Theta(\Gamma(t))(\Theta(\Gamma(0))^{-1}X(0)) \\ &= \Phi_{\gamma(t)}^{-1}\left(\Gamma(t)\Gamma(0)^{-1}\Phi_{\gamma(0)}(X(0))(\Gamma(0)^*)^{-1}\Gamma(t)^*\right). \end{aligned}$$

Then we obtain the *covariant derivative* $\nabla_{\dot\gamma}$:

$$\nabla_{\dot\gamma}X = \Phi_\gamma^{-1}\left((\Phi_\gamma(X))' - \frac{(F(\gamma))'(F(\gamma))^{-1}\Phi_\gamma(X) + \Phi_\gamma(X)(F(\gamma))^{-1}(F(\gamma))'}{2}\right).$$

Theorem 5.1. *The connection in the principal fiber bundle $\mathcal{P}_F$ in the above yields the geodesic equation*

$$(F(\gamma))'' = (F(\gamma))'(F(\gamma))^{-1}(F(\gamma))'$$

and the geodesic from A to B

$$\delta_F(t) = F^{-1}(F(A)\,\#_t\,F(B)).$$

Proof. Since the geodesic δ is auto-parallel, we have the geodesic equation;

$$\begin{aligned} O &= \Phi_\delta(\nabla_{\dot\delta}\delta) = (\Phi_\delta(\dot\delta))' - (F(\delta))'(F(\delta))^{-1}\Phi_\delta(\dot\delta) \\ &= (F(\delta))'' - (F(\delta))'(F(\delta))^{-1}(F(\delta))'. \end{aligned}$$

As in the CPR geometry, we have

$$F(\delta(t)) = F(\delta(0))\,\#_t\,F(\delta(1)) = F(A)\,\#_t\,F(B)$$

and hence the required geodesic is obtained. $\square$

Remark 5.1. Here we give another proof to solve the geodesic equation. Let

$$h(t) = F(\delta(0))^{-1/2}F(\delta(t))F(\delta(0))^{-1/2} = F(A)^{-1/2}F(\delta(t))F(A)^{-1/2}.$$

Then h is a smooth path from I to $F(A)^{-1/2}F(B)F(A)^{-1/2}$ and also satisfies $h'' = h'h^{-1}h'$. By

$$(h'h^{-1})' = h''h^{-1} - h'h^{-1}h'h^{-1} = (h'' - h'h^{-1}h')h^{-1} = O,$$

we have $h'(t)h(t)^{-1} = C$ for some matrix C. Since $C = h'(0)$ is hermitian, $h'(t)$ and $h(t)$ are commuting for each t. Then, using the Cauchy integral for a curve Ω with $\mathrm{Re}\,\Omega > 0$ as the boundary of an open region including all the spectra $\sigma(h(t))$, we have

$$\log h(t) = \frac{1}{2\pi i}\int_\Omega (\log z)(z - h(t))^{-1}dz$$

and the commutativity shows

$$\begin{aligned} \frac{d\log h(t)}{dt} &= \frac{1}{2\pi i}\int_\Omega (\log z)(z - h(t))^{-1}h'(t)(z - h(t))^{-1}dz \\ &= \frac{h'(t)}{2\pi i}\int_C (\log z)(z - h(t))^{-2}dz = h'(t)\log' h(t) = h'(t)h(t)^{-1} = C. \end{aligned}$$

Thus we have $\log h(t) = tC$, or $h(t) = \exp(tC)$. Moreover h is determined as

$$h(t) = (F(A)^{-1/2}F(B)F(A)^{-1/2})^t$$

by $h(1) = F(A)^{-1/2}F(B)F(A)^{-1/2}$ and hence

$$\delta_F(t) = F^{-1}(F(A)\,\#_t\,F(B)).$$

Similarly to the preceding case, we also show that $\mathcal{M}^+$ is a Finsler space:

Theorem 5.2. *For a unitarily invariant norm $|||\ |||$, the following Minkovski norm*

$$L(X; A) = L_{F,QGM}(X; A) = |||F(A)^{-1/2}\Phi_A(X)F(A)^{-1/2}|||$$

defines a Finsler metric and the geodesic length $d(A, B)$ from A to B makes $\mathcal{M}^+$ a metric space:

$$d(A, B) = \int_0^1 L(\dot{\delta}_F(t); \delta_F(t))dt = |||\log F(A)^{-1/2}F(B)F(A)^{-1/2}|||.$$

Proof. It follows from $F(\gamma) = \Gamma\Gamma^*$ that

$$F(\gamma(t))^{-1/2}\Gamma(t)\Gamma(0)^{-1}F(\gamma(0))(\Gamma(0)^*)^{-1}\Gamma(t)^*F(\gamma(t))^{-1/2} = I,$$

so that we have $W = F(\gamma(t))^{-1/2}\Gamma(t)\Gamma(0)^{-1}F(\gamma(0))^{1/2}$ is unitary and hence

$$\begin{aligned}
L(P_t^0X(0); \gamma(t)) &= \left|\left|\left| F(\gamma(t))^{-1/2}\Phi_{\gamma(t)}(P_t^0X(0))F(\gamma(t))^{-1/2} \right|\right|\right| \\
&= \left|\left|\left| F(\gamma(t))^{-1/2}\left(\Gamma(t)\Gamma(0)^{-1}\Phi_{\gamma(0)}(X(0))(\Gamma(0)^*)^{-1}\Gamma(t)^*\right)F(\gamma(t))^{-1/2} \right|\right|\right| \\
&= \left|\left|\left| WF(\gamma(0))^{-1/2}\Phi_{\gamma(0)}(X(0))F(\gamma(0))^{-1/2}W^* \right|\right|\right| \\
&= \left|\left|\left| F(\gamma(0))^{-1/2}\Phi_{\gamma(0)}(X(0))F(\gamma(0))^{-1/2} \right|\right|\right| = L(X(0); \gamma(0)).
\end{aligned}$$

Thus it is a Finsler metric. By the differential formula (2),

$$\begin{aligned}
L(\dot{\gamma}(t); \gamma(t)) &= |||F(\gamma(t))^{-1/2}\Phi_{\gamma(t)}(\dot{\gamma}(t))F(\gamma(t))^{-1/2}||| \\
&= |||F(\gamma(t))^{-1/2}(F(\gamma))'(t)F(\gamma(t))^{-1/2}|||.
\end{aligned}$$

Since the geodesic δ_F from A to B is auto-parallel, we have $d(A, B)$:

$$\begin{aligned}
\int_0^1 L(\dot{\delta}_F(t); \delta_F(t))dt &= L(\dot{\delta}_F(0); \delta_F(0))\int_0^1 dt = L(\dot{\delta}_F(0); \delta_F(0)) \\
&= |||F(A)^{-1/2}\Phi_A(\dot{\delta}_F(0))F(A)^{-1/2}||| = |||F(A)^{-1/2}(F(\delta_F))'(0)F(A)^{-1/2}||| \\
&= |||F(A)^{-1/2}S(F(A)|F(B))F(A)^{-1/2}||| = |||\log F(A)^{-1/2}F(B)F(A)^{-1/2}|||.
\end{aligned}$$

Moreover the translation invariance of $L(\dot{\delta}_F(t); \delta_F(t))$ implies the symmetric law of d:

$$d(A, B) = L(\dot{\delta}_F(1); \delta_F(1)) = |||F(B)^{-1/2}(F(\delta_F))'(1)F(B)^{-1/2}||| = d(B, A),$$

so that d is a metric function. □

Remark 5.2. To make this metric L an exact extension in the Hiai–Petz geometry, it should be

$$L(X; A) = \frac{1}{F'(1)}|||F(A)^{-1/2}\Phi_A(X)F(A)^{-1/2}|||.$$

But in this paper, we use the above definition for simplicity.

By the differential formula (2') again, we have

$$\dot{\delta}_F(t) = U_t\left(\left(\frac{1}{F^{[1]}(d_i(t), d_j(t))}\right) \circ U_t^* \frac{d\,F(A)\,\#_t\,F(B)}{dt} U_t\right) U_t^*.$$

As the derivative at 0 of the geodesic, the *geometric relative operator entropy for F* is defined as

$$\begin{aligned} S_F(A|B) &= \dot{\delta}_F(0) = \Phi_A^{-1}\,(S\,(F(A)|F(B))) \\ &= U\left(\left(\frac{1}{F^{[1]}(d_i, d_j)}\right) \circ U^* S(F(A)|F(B)) U\right) U^* \end{aligned}$$

and also its *geometric relative entropy* is

$$s_F(A|B) = \mathrm{tr}\ F'(A)^{-1} S(F(A)|F(B)).$$

In case $F(x) = x$, it is the minus quantity of the Belavkin–Staszewski relative entropy [3].

For a path γ and an invertible matrix Y, define the path γ_Y by

$$\gamma_Y(t) \equiv F^{-1}(Y^* F(\gamma(t)) Y).$$

Then we have an invariance property, cf. [13]:

Lemma 5.3. *$L_{[F]}(\dot{\gamma}_Y; \gamma_Y) = L_{[F]}(\dot{\gamma}; \gamma)$ and $d(\gamma_Y, \delta_Y) = d(\gamma, \delta)$.*

Proof. Since $|||Z||| = |||\,|Z|\,||| = |||\sqrt{Z^*Z}||| = |||\sqrt{ZZ^*}|||$, we have

$$\begin{aligned} L_{[F]}(\dot{\gamma}_Y; \gamma_Y) &= \left|\left|\left| F(\gamma_Y)^{-\frac{1}{2}} (F(\gamma_Y))' F(\gamma_Y)^{-\frac{1}{2}} \right|\right|\right| \\ &= \left|\left|\left| (Y^*F(\gamma)Y)^{-\frac{1}{2}} (Y^*F(\gamma)Y)' (Y^*F(\gamma)Y)^{-\frac{1}{2}} \right|\right|\right| \\ &= \left|\left|\left| (Y^*F(\gamma)Y)^{-\frac{1}{2}} Y^*(F(\gamma))' Y (Y^*F(\gamma)Y)^{-\frac{1}{2}} \right|\right|\right| \\ &= \left|\left|\left| \sqrt{(Y^*F(\gamma)Y)^{-\frac{1}{2}} Y^*(F(\gamma))' Y (Y^*F(\gamma)Y)^{-1} Y^*(F(\gamma))' Y (Y^*F(\gamma)Y)^{-\frac{1}{2}}} \right|\right|\right| \\ &= \left|\left|\left| \sqrt{(Y^*F(\gamma)Y)^{-\frac{1}{2}} Y (F(\gamma))' F(\gamma)^{-1} (F(\gamma))' Y^* (Y^*F(\gamma)Y)^{-\frac{1}{2}}} \right|\right|\right| \\ &= \left|\left|\left| \sqrt{F(\gamma)^{-\frac{1}{2}} (F(\gamma))' Y (Y^*F(\gamma)Y)^{-1} Y^*(F(\gamma))' F(\gamma)^{-\frac{1}{2}}} \right|\right|\right| \\ &= \left|\left|\left| \sqrt{F(\gamma)^{-\frac{1}{2}} (F(\gamma))' F(\gamma)^{-1} (F(\gamma))' F(\gamma)^{-\frac{1}{2}}} \right|\right|\right| \\ &= \left|\left|\left| F(\gamma)^{-\frac{1}{2}} (F(\gamma))' F(\gamma)^{-\frac{1}{2}} \right|\right|\right| = L_{[F]}(\dot{\gamma}; \gamma). \end{aligned}$$

Let γ and δ be geodesics. By the relation for all $s, t \in [0, 1]$

$$F(\gamma_Y(s)\,\#_{F,t}\,\delta_Y(s)) = Y^* F(\gamma(s)\,\#_{F,t}\,\delta(s)) Y,$$

similarly we have

$$L_{[F]}\left(\frac{d\,\gamma_Y(s)\,\#_{F,t}\,\delta_Y(s)}{dt}; \gamma_Y(s)\,\#_{F,t}\,\delta_Y(s)\right) = L_{[F]}\left(\frac{d\,\gamma(s)\,\#_{F,t}\,\delta(s)}{dt}; \gamma(s)\,\#_{F,t}\,\delta(s)\right),$$

which implies $d(\gamma_Y, \delta_Y) = d(\gamma, \delta)$. $\square$

So we have the following property like the CPR geometry which suggests the curvature would be negative, see [2,5]:

Theorem 5.4. *For geodesics γ and δ in the QGM geometry, the followings hold and they are equivalent:*
(i) $J(t) = d(\gamma(t), \delta(t)) = \left\| \log F\left(\gamma(t)^{-\frac{1}{2}}\right) F(\delta(t)) F\left(\gamma(t)^{-\frac{1}{2}}\right) \right\|$ *is convex.*
(ii) $d(F^{-1}(F(A)^t), F^{-1}(F(B)^t)) \leqq td(A,B)$.
(iii) $d(\gamma(t), \delta(t)) \leqq (1-t)d(\gamma(0), \delta(0)) + td(\gamma(1), \delta(1))$.

Proof. Based on Araki's inequality in [1] for positive semidefinite X and Y

$$\prod_{j=1}^{k} \lambda_j(Y^{-t/2}X^tY^{-t/2}) \leqq \prod_{j=1}^{k} \lambda_j^t\left(Y^{-\frac{1}{2}}XY^{-\frac{1}{2}}\right)$$

for $0 \leqq t \leqq 1$ and $1 \leqq k \leqq n$ where λ_j is the jth eigenvalue (singular value in this case) under the decreasing order, we have the majorization

$$\log Y^{-t/2}X^tY^{-t/2} \prec t \log Y^{-\frac{1}{2}}XY^{-\frac{1}{2}}.$$

Since positive semidefinite matrices X and Y are arbitrary, we have

$$\log F(B)^{-t/2}F(A)^tF(B)^{-t/2} \prec t \log F(B)^{-\frac{1}{2}}F(A)F(B)^{-\frac{1}{2}}$$

and consequently we have (ii) holds by the convexity of $f(x) = |x|$:

$$\begin{aligned} d(F^{-1}(F(A)^t), F^{-1}(F(B)^t)) &= \left\| \log F(B)^{-t/2}F(A)^tF(B)^{-t/2} \right\| \\ &\leqq t \left\| \log F(B)^{-\frac{1}{2}}F(A)F(B)^{-\frac{1}{2}} \right\| = td(A,B). \end{aligned}$$

Thus (ii) holds.

Next, we show that (ii) implies (iii): For $\gamma(t) = A \#_{F,t} B$, $\delta(t) = C \#_{F,t} D$ and $\zeta(t) = C \#_{F,t} B$, the triangle inequality and Lemma 5.3 show (iii) by

$$\begin{aligned} & d(\gamma(t), \delta(t)) \leqq d(\gamma(t), \zeta(t)) + d(\zeta(t), \delta(t)) \\ &= d\left(\gamma_{F(B)^{-\frac{1}{2}}}(t), \zeta_{F(B)^{-\frac{1}{2}}}(t)\right) + d\left(\zeta_{F(C)^{-\frac{1}{2}}}(t), \delta_{F(C)^{-\frac{1}{2}}}(t)\right) \\ &= d\left(F^{-1}\left(\left(F(B)^{-\frac{1}{2}}F(A)F(B)^{-\frac{1}{2}}\right)^{1-t}\right), F^{-1}\left(\left(F(B)^{-\frac{1}{2}}F(C)F(B)^{-\frac{1}{2}}\right)^{1-t}\right)\right) \\ &\quad + d\left(F^{-1}\left(\left(F(C)^{-\frac{1}{2}}F(B)F(C)^{-\frac{1}{2}}\right)^{t}\right), F^{-1}\left(\left(F(C)^{-\frac{1}{2}}F(D)F(C)^{-\frac{1}{2}}\right)^{t}\right)\right) \\ &\leqq (1-t)d\left(F^{-1}\left(F(B)^{-\frac{1}{2}}F(A)F(B)^{-\frac{1}{2}}\right), F^{-1}\left(F(B)^{-\frac{1}{2}}F(C)F(B)^{-\frac{1}{2}}\right)\right) \\ &\quad + td\left(F^{-1}\left(F(C)^{-\frac{1}{2}}F(B)F(C)^{-\frac{1}{2}}\right), F^{-1}\left(F(C)^{-\frac{1}{2}}F(D)F(C)^{-\frac{1}{2}}\right)\right) \\ &= (1-t)d\left(\gamma_{F(B)^{-\frac{1}{2}}}(0), \zeta_{F(B)^{-\frac{1}{2}}}(0)\right) + td\left(\zeta_{F(C)^{-\frac{1}{2}}}(1), \delta_{F(C)^{-\frac{1}{2}}}(1)\right) \\ &= (1-t)d(\gamma(0), \zeta(0)) + td(\zeta(1), \delta(1)) = (1-t)d(\gamma(0), \delta(0)) + td(\gamma(1), \delta(1)). \end{aligned}$$

Next we show (iii) implies (i). Let $\Gamma(t) = (A \#_{F,p} B) \#_t (A \#_{F,q} B)$ and $\Delta(t) = (C \#_{F,p} D) \#_t (C \#_{F,q} D)$. Then, by the interpolationality, we have

$$\Gamma(t) = A \#_{F,(1-t)p+tq} B \quad \text{and} \quad \Delta(t) = C \#_{F,(1-t)p+tq} D$$

and then

$$\Gamma(0) = A \#_{F,p} B, \quad \Gamma(1) = A \#_{F,q} B, \quad \Delta(0) = C \#_{F,p} D \quad \text{and} \quad \Delta(1) = C \#_{F,q} D.$$

Thus, we have (iii) for Γ and Δ implies (i) for γ and δ. Considering geodesics $\gamma(t) = F^{-1}(F(A)^t)$ and $\delta(t) = F^{-1}(F(B)^t)$, we have (i) implies (ii);

$$\begin{aligned} d(F^{-1}(F(A)^t), F^{-1}(F(B)^t)) &= J(t) = J((1-t)0+t) \leqq (1-t)J(0) + tJ(1) \\ &= (1-t)d(I,I) + td(A,B) = td(A,B). \end{aligned}$$

Thus all the conditions hold and they are equivalent. □

6. Shortest paths

Recall that the norm $\|\ \|$ is strictly convex if $\|x+y\| < 2$ for distinct unit vectors x and y. This is equivalent to the condition that $\|x+y\| = \|x\| + \|y\|$ if and only if one vector is the positive scalar multiple of the other.

By the differential formula (2), we have

$$\int_0^1 L_F(\dot\gamma(t);\gamma(t))dt = \int_0^1 |||(F(\gamma(t))'|||dt \geq q|||\int_0^t (F(\gamma(t))'dt||| = |||F(B)-F(A)|||,$$

so that the geodesic $A\,\mathfrak{m}_{F,t}B$ attains the shortest length. Moreover only the geodesic is the shortest path if the norm is strictly convex:

Theorem 6.1. *Let F be a smooth function on $(0,\infty)$ with $F' > 0$. If a unitarily invariant norm is strictly convex, the geodesic*

$$\gamma_F(t) = A\,\mathfrak{m}_{F,t}B = F^{-1}((1-t)F(A) + tF(B))$$

is the unique shortest path for the metric

$$L_F(X;A) = L_{F,QAM}(X;A) = |||(F^{[1]}(d_i,d_j)) \circ U^*XU|||$$

where the shortest length is $|||F(B)-F(A)|||$.

Proof. Suppose γ attains the shortest length. Since

$$L_F(\dot\gamma;\gamma) = |||(F^{[1]}(d_i(t),d_j(t))) \circ U_t^*\dot\gamma(t)U_t||| = \left|\left|\left|\frac{dF(\gamma(t))}{dt}\right|\right|\right|$$

for a parametrized diagonalization $U_t^*\gamma(t)U_t = \mathrm{diag}(d_j(t))$, the length $\ell(\gamma)$ satisfies

$$|||F(B)-F(A)||| = \int_0^1 L_F(\dot\gamma;\gamma)dt \geq q\left|\left|\left|\int_0^1 \frac{dF(\gamma(t))}{dt}dt\right|\right|\right| = |||F(B)-F(A)|||.$$

For $H(t) = F(\gamma(t))$, it must satisfy

$$\int_0^1 |||H'(t)|||dt = \left|\left|\left|\int_0^1 H'(t)dt\right|\right|\right| = |||F(B)-F(A)|||.$$

Here we use the broken line approximation to obtain the length of $H(t)$:

$$\int_0^1 |||H'(t)|||dt = \lim_{|\Delta|\to 0} \sum_{t_n\in\Delta} |||H(t_{n+1}) - H(t_n)|||.$$

Take the following monotone increasing sequence converging to $\int_0^1 \|H'(t)\|dt$:

$$\sum_{k=1}^{2^n} \left|\left|\left|H\left(\frac{k}{2^n}\right) - H\left(\frac{k-1}{2^n}\right)\right|\right|\right| \quad \uparrow \quad \int_0^1 |||H'(t)|||dt.$$

Then all the triangle inequalities

$$\left\| H\left(\frac{k}{2^n}\right) - H\left(\frac{k-1}{2^n}\right) \right\| \leqq \left\| H\left(\frac{2k}{2^{n+1}}\right) - H\left(\frac{2k-1}{2^{n+1}}\right) \right\| + \left\| H\left(\frac{2k-1}{2^{n+1}}\right) - H\left(\frac{2(k-1)}{2^{n+1}}\right) \right\|$$

are equal, so that there exists $s_{\frac{2k-1}{2^{n+1}}} > 0$ with

$$H\left(\frac{2k}{2^{n+1}}\right) - H\left(\frac{2k-1}{2^{n+1}}\right) = s_{\frac{2k-1}{2^{n+1}}} \left(H\left(\frac{2k-1}{2^{n+1}}\right) - H\left(\frac{2(k-1)}{2^{n+1}}\right) \right),$$

that is, $H\left(\frac{2k-1}{2^{n+1}}\right)$ at each binary fraction $\frac{2k-1}{2^{n+1}}$ in [0,1] is the convex combination for $H\left(\frac{k-1}{2^n}\right)$ and $H\left(\frac{k}{2^n}\right)$;

$$H\left(\frac{2k-1}{2^{n+1}}\right) = \frac{s_{\frac{2k-1}{2^{n+1}}} H\left(\frac{k-1}{2^n}\right) + H\left(\frac{k}{2^n}\right)}{s_{\frac{2k-1}{2^{n+1}}} + 1}$$

holds for all n and $k = 1, \ldots, 2^n$. Thus, all the constants $s_{\frac{2k-1}{2^{n+1}}}$ are defined from the terminal points $H(0) = F(A)$ and $H(1) = F(B)$ with $s_0 = 0$ and $s_1 = 1$. Therefore we can define a function w on the binary fractions in [0,1] inductively with the relation

$$H\left(\frac{2k-1}{2^{n+1}}\right) = \left(1 - w\left(\frac{2k-1}{2^{n+1}}\right)\right) F(A) + w\left(\frac{2k-1}{2^{n+1}}\right) F(B).$$

In fact, paying attention to the coefficient of $F(B)$, we have the recurrence equation

$$w\left(\frac{2k-1}{2^n}\right) = \frac{s_{\frac{2k-1}{2^n}} w\left(\frac{k-1}{2^{n-1}}\right) + w\left(\frac{k}{2^{n-1}}\right)}{s_{\frac{2k-1}{2^n}} + 1}.$$

By the initial conditions $w(0) = 0$ and $w(1) = 1$, the function w is monotone increasing on the above fractions. The smoothness of γ implies that w is smoothly extended to a function on [0,1] which is monotone increasing and satisfies

$$F(\gamma(t)) = H(t) = (1 - w(t))F(A) + w(t)F(B).$$

Thus

$$\gamma(t) = F^{-1}((1 - w(t))F(A) + w(t)F(B)) = A\, \mathfrak{m}_{F,w(t)} B,$$

that is, γ can be identified with $A\, \mathfrak{m}_{F,t} B$. $\square$

Example 2. Similarly to the Hiai–Petz geometry shown in [13], each Ky Fan k-norm gives another path attaining the shortest length. We may assume that $F(x)$ is a non-affine fundamental function since we have already shown the case $F(x) = x^\alpha$ in [13]. Then there exists an interval (a, b) with $b > a > 0$ such that

$$F''(1 + ty)ty + F'(1 + ty) > 0$$

holds for all $y \in (a, b)$. This condition implies $F'(1 + t(x-1))(x-1)$ is monotone increasing for $x \in (a+1, b+1)$. Take monotone decreasing numbers $b_j \in (a+1, b+1)$ and put the diagonal matrix $B = \mathrm{diag}(b_j) > 1$. Recall that the Ky Fan k-norm of $f(B)$ for a increasing function f is

$$\|f(B)\|_{(k)} = \sum_{j=1}^{k} f(b_j).$$

Then we can see that $\alpha(t) = (1-t)I + tB = I + t(B-I)$ has the shortest length: since $\dot{\alpha}(t) = B - I$,

$$\begin{aligned} L_{F,\|\ \|_{(k)}}(\dot{\alpha};\alpha) &= \|F'(\alpha(t)) \circ (B-I)\|_{(k)} \\ &= \|\mathrm{diag}(F'(1+t(b_j-1))(b_j-1))\|_{(k)} \\ &= \sum_{j=1}^{k} F'(1+t(b_j-1))(b_j-1) = \sum_{j=1}^{k}(F(1+t(b_j-1)))', \end{aligned}$$

so that,

$$\begin{aligned} \ell_{\|\ \|_{(k)}}(\alpha) &= \int_0^1 \sum_{j=1}^{k}(F(1+t(b-1)))'dt = \sum_{j=1}^{k}\left[F(1+t(b_j-1))\right]_0^1 \\ &= \sum_{j=1}^{k} F(b_j) = \|F(B)\|_{(k)} = \|F(B) - F(I)\|_{(k)}. \end{aligned}$$

Also, in the QGM geometry, the geodesic $\delta_F(t) = F^{-1}(F(A) \#_t F(B))$ is the shortest if the norm is strictly convex:

Theorem 6.2. *Let F be an increasing smooth nonnegative function on $(0,\infty)$ with $F(1) = 1$ and $F' > 0$. If a unitarily invariant norm is strictly convex, the geodesic*

$$\delta_F(t) = A \#_{F,t} B = F^{-1}(F(A) \#_t F(B))$$

is the unique shortest path for the metric $L_{F,QGM}(X;A) = L_{[F]}$ defined by

$$L_{[F]}(X;A) = \left|\!\left|\!\left| F(A)^{-\frac{1}{2}} U \left[(F^{[1]}(d_i,d_j)) \circ U^*XU\right] U^*F(A)^{-\frac{1}{2}} \right|\!\right|\!\right|,$$

where the shortest length is

$$\left|\!\left|\!\left| \log(F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}}) \right|\!\right|\!\right|.$$

To show this theorem, we use *Dyson's expansion* in [4] for $H(t) = \log F(\gamma(t))$;

$$(F(\gamma(t)))' = \frac{d}{dt}e^{H(t)} = \int_0^1 e^{uH(t)}H'(t)e^{(1-u)H(t)}du.$$

Proof. The Hiai–Kosaki logarithmic–geometric mean inequality in [16] implies

$$\begin{aligned} L_{[F]}(\dot{\gamma};\gamma) &= \left|\!\left|\!\left| F(\gamma)^{-\frac{1}{2}}(F(\gamma))'F(\gamma)^{-\frac{1}{2}} \right|\!\right|\!\right| \\ &= \left|\!\left|\!\left| e^{-\frac{H(t)}{2}}\left(\int_0^1 e^{uH(t)}H'(t)e^{(1-u)H(t)}du\right)e^{-\frac{H(t)}{2}} \right|\!\right|\!\right| \\ &= \left|\!\left|\!\left| \int_0^1 e^{uH(t)}e^{-\frac{H(t)}{2}}H'(t)e^{-\frac{H(t)}{2}}e^{(1-u)H(t)}du \right|\!\right|\!\right| \\ &\geq q \left|\!\left|\!\left| e^{\frac{H(t)}{2}}e^{-\frac{H(t)}{2}}H'(t)e^{-\frac{H(t)}{2}}e^{\frac{H(t)}{2}} \right|\!\right|\!\right| = |\!|\!| H'(t) |\!|\!|. \end{aligned}$$

For each path γ from A to B, consider $\delta(t) \equiv \gamma_{F(A)^{-1/2}}$ is a path from I to $F^{-1}(F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}})$ and $H(t) = \log F(\delta(t))$. In this case,

$$H(0) = O \quad \text{and} \quad H(1) = \log F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}}.$$

Then the length $\ell(\delta)$ is estimated in the below:

$$\ell(\delta) \equiv \int_0^1 L_{[F]}(\dot{\delta};\delta)dt \geq q\int_0^1 |||H'(t)|||dt \geq q\left|\left|\left|\int_0^1 H'(t)dt\right|\right|\right|$$
$$= |||H(1)-H(0)||| = \left|\left|\left|\log(F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}})\right|\right|\right|,$$

so that, the geodesic δ_F attains the shortest length by Lemma 5.3.

Now suppose a path γ from A to B attains the shortest length. Then it must satisfy

$$\int_0^1 |||H'(t)|||dt = \left|\left|\left|\int_0^1 H'(t)dt\right|\right|\right| = \left|\left|\left|\log F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}}\right|\right|\right|.$$

Therefore, similarly to the preceding proof, we have

$$H(t) = (1-w(t))H(0) + w(t)H(1) = w(t)\log F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}},$$

that is,

$$\delta(t) = F^{-1}\left(\left(F(A)^{-\frac{1}{2}}F(B)F(A)^{-\frac{1}{2}}\right)^{w(t)}\right)$$

and hence $\gamma(t)$ equals $A\,\#_{\alpha,t}\,B$ as paths. $\square$

We also give another path with the shortest length for the Ky Fan k-norm:

Example 3. The path $\beta(t) = F^{-1}((1-t)F(I)+tF(B))$ attains the shortest length for the matrix B in the preceding example. In fact,

$$L_{[F],k}(\dot{\beta};\beta) = \|F(\beta)^{-1/2}(F(\beta))'F(\beta)^{-1/2}\|_{(k)}$$
$$= \|(F(B)-F(I))((1-t)F(1)+tF(B))^{-1}\|_{(k)}.$$

Putting

$$G(x) = \frac{F(x)-F(1)}{(1-t)F(1)+tF(x)},$$

we have

$$G'(x) = \frac{F(1)F'(x)}{((1-t)F(1)+tF(x))^2} > 0$$

and hence G is monotone increasing. Therefore $L_{[F],k}(\dot{\beta};\beta) = \sum_{j=1}^k G(b_j)$ and hence

$$\ell(\beta) = \int_0^1 \sum_{j=1}^k G(b_j)dt = \sum_{j=1}^k \int_0^1 \frac{F(b_j)-F(1)}{(1-t)F(1)+tF(b_j)}dt$$
$$= \sum_{j=1}^k \left[\log((1-t)F(1)+tF(b_j))\right]_0^1 = \sum_{j=1}^k (\log F(b_j) - \log F(1))$$
$$= \|\log F(B) - \log F(I)\|_{(k)} = \left\|\log F(I)^{-\frac{1}{2}}F(B)F(I)^{-\frac{1}{2}}\right\|_{(k)},$$

so that β is one of the shortest paths.

Acknowledgement

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 20540166, 2008.

References

[1] H. Araki, On an inequality of Lieb and Thirring, Lett. Math. Phys. 19 (1990) 167–170.
[2] E. Andruchow, G. Corach, D. Stojanoff, Geometrical significance of Lowner–Heinz inequality, Proc. Amer. Math. Soc. 128 (2000) 1031–1037.
[3] V.P. Belavkin, P. Staszewski, C*-algebraic generalization of relative entropy and entropy, Ann. Inst. H. Poincaré Sect. A 37 (1982) 51–58.
[4] R. Bhatia, Matrix Analysis, Springer, 1997.
[5] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2007.
[6] R. Bhatia, J.A.R. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 423 (2006) 594–618.
[7] G. Corach, H. Porta, L. Recht, Geodesics and operator means in the space of positive operators, Int. J. Math. 4 (1993) 193–202.
[8] G. Corach, A.L. Maestripieri, Differential and metrical structure of positive operators, Positivity 3 (1999) 297–315.
[9] J.I. Fujii, E. Kamei, Relative operator entropy in noncommutative information theory, Math. Japon. 34 (1989) 341–348.
[10] J.I. Fujii, E. Kamei, Uhlmann's interpolational method for operator means, Math. Japon. 34 (1989) 541–547.
[11] J.I. Fujii, E. Kamei, Interpolational paths and their derivatives, Math. Japon. 39 (1993) 557–560.
[12] J.I. Fujii, M. Nakamura, S.-E. Takahasi, Cooper's approach to chaotic operator means, Sci. Math. Japon. 63 (2006) 319–324.
[13] J.I. Fujii, The Hiai–Petz geodesic for strongly convex norm is the unique shortest path, Sci. Math. Japon. Online e-2009 (2009) 579–586.
[14] J.I. Fujii, Structure of Hiai–Petz parametrized geometry for positive definite matrices, Linear Algebra Appl. 432 (2010) 318–326.
[15] J.I. Fujii, M. Fujii, T. Miura, H. Takagi, S.-E. Takahasi, Continuously differentiable means, J. Inequal. Appl. 2006 (2006) (Art. ID 75941).
[16] F. Hiai, H. Kosaki, Means of Hilbert space operators, Lecture Notes in Mathematics, vol. 1820, Springer-Verlag, 2003.
[17] F. Hiai, D. Petz, Riemannian metrics on positive definite matrices related to means, Linear Algebra Appl. 430 (2009) 3105–3130.
[18] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis', Cambridge University Press, 1991.
[19] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, vol. 1, Wiley Interscience, New York, 1963.
[20] F. Kubo, T. Ando, Means of positive linear operators, Math. Ann. 246 (1980) 205–224.
[21] M. Matsumoto, Metric Differential Geometry, Shokabo, 1975 (in Japanese).
[22] H. Ohmori, Theory of Infinite-dimensional Lie Group, Kinokuniya, Tokyo, 1978 (in Japanese).
[23] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat. Phys. 52 (1988) 479–487.
[24] A. Uhlmann, Relative entropy and the Wigner–Yamase–Dyson–Lieb concavity in an interpolation theory, Commun. Math. Phys. 54 (1977) 22–32.
[25] H. Umegaki, Conditional expectation in an operator algebra IV, Kodai Math. Sem. Rep. 14 (1962) 59–85.
[26] K. Yanagi, K. Kuriyama, S. Furuichi, Generalized Shannon inequalities based on Tsallis relative operator entropy, Linear Algebra Appl. 394 (2005) 109–118.