HYPERBOLICITY OF THE KARCHER MEAN

EDUARDO GHIGLIONI AND YONGDO LIM

ABSTRACT. The main concern of this paper is the Karcher mean of linearly independent triples (A, B, C) on the hyperboic manifold of 2×2 positive definite matrices of determinant 1. We show that the Karcher mean is of the form

$$\Lambda(A, B, C) = xA + y(B + C), \qquad 0 < x, y \text{ and } x + 2y < 1$$

under the trace condition $\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1})$. We further find an invertible hyperbolic matrix M depending only on the trace values $\mathrm{tr}(AB^{-1})$ and $\mathrm{tr}(BC^{-1})$ such that $\begin{bmatrix} x \\ y \end{bmatrix} = M \begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix}$ for some (unique) $\theta \in \mathbb{R}$.

2010 Mathematics Subject Classification. 15A18, 15A42, 47A64, 53C20.
Key words and phrases. Positive definite matrix, Karcher mean, CS-decomposition, Hyperboloid of two sheets, Hyperbolic matrix, Positively stable matrix.

1. INTRODUCTION AND MAIN RESULTS

The present paper is a continuation of the previous work regarding a closed form of the Karcher mean for 2×2 positive matrices triples [6, 8, 9].

Let $\mathbb{P}$ be the set of 2×2 positive definite Hermitian matrices, which is a Cartan-Hadamard Riemannian manifold equipped with the trace metric $\delta(A, B) = \|\log A^{-1}B\|_2$, where $\|X\|_2 = \sqrt{\mathrm{tr}\, X^*X}$ is the Frobenius norm. The Karcher mean of a triple $(A, B, C) \in \mathbb{P}^3$ is defined as

$$\Lambda(A, B, C) := \underset{X \in \mathbb{P}}{\arg\min}\, (\delta^2(X, A) + \delta^2(X, B) + \delta^2(X, C))$$

and also as a unique positive definite solution of the Karcher equation

$$\log(X^{-1/2}AX^{-1/2}) + \log(X^{-1/2}BX^{-1/2}) + \log(X^{-1/2}CX^{-1/2}) = 0. \tag{1.1}$$

See [4, 5, 11, 10, 12, 13, 14, 15, 16, 17, 18] for more on the Riemannian trace metric and the Karcher mean.

Of special interest to us here is to find a closed-form expression of the Karcher mean $\Lambda(A,B,C)$. By joint homogeneity of the Karcher mean, this problem can be reduced to the hyperbolic manifold $\mathbf{H}_2$ of positive definite matrices of determinant 1. For $A,B \in \mathbf{H}_2$, the matrix geometric mean A and B defined by $A\#B := A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$ coincides with the geometric middle $\Lambda(A,B)$ between A and B and has the following linear form [1, 3]:

$$A\#B = \frac{A+B}{\sqrt{\det(A+B)}}. \tag{1.2}$$

This together with Sturm's SLLN [19] and Holbrook's no dice theorem [10] provides a remarkable *linear* property of the Karcher mean on $\mathbf{H}_2$:

$$\Lambda(A,B,C) = xA + yB + zC, \qquad x,y,z \geq 0. \tag{1.3}$$

The coefficients x,y,z are unique when the triple (A,B,C) is linearly independent in the Eucliden space of 2×2 Hermitian matrices. It is shown in [9] that

$$x = y = z = \frac{1}{\sqrt{\det(A+B+C)}} \tag{1.4}$$

under the trace condition $\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1}) = \mathrm{tr}(BC^{-1})$.

The main concern of this paper is the case of linearly independent triples $\{A,B,C\} \subset \mathbf{H}_2$ satisfying

$$\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1}), \tag{1.5}$$

which is equivalent to the isosceles property of the triangle of A,B,C in the hyperbolic manifold $\mathbf{H}_2$; $\overline{AB} = \overline{AC}$ (see Proposition 2.1).

The first main result in this paper is the following theorem.

Theorem 1.1. *There exist unique positive real numbers x and y such that*

$$\Lambda(A,B,C) = xA + yB + yC. \tag{1.6}$$

The key tool in the proof of Theorem 1.1 is the following unitary diagonalization of 2×2 positive definite matrix A :

$$A = U \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix} U^*, \tag{1.7}$$

where λ, β are eigenvalues of A and U is a unitary *diagonal* matrix. The unitary diagonal matrix U and the CS matrix in (1.7), which is *orhogonal and involutory*, are obtained from the CS decomposition of 2×2 unitary matrices.

By invariance of the Karcher mean under permutations, (1.4) is immediate from Theorem 1.1. However, finding a closed form of x and y respectively is still problematic in general case. We will show that the coefficient vector $(x, y, z = y)$ lies in a hyperboloid of two sheets determined by a and c;

$$x^2 + y^2 + z^2 + 2axy + 2bxz + 2cyz = 1, \tag{1.8}$$

where $a = b := (1/2)\,\mathrm{tr}(AB^{-1})$ and $c := (1/2)\,\mathrm{tr}(BC^{-1})$, and that

$$c + 1 < \sqrt{c^2 + 8a^2} - 1 < 2a^2, \tag{1.9}$$

from which the following matrix depending on a and c

$$M := \frac{1}{\sqrt{\xi}} \begin{bmatrix} \sqrt{\frac{\xi - c}{\xi + c + 2}} & -\sqrt{\frac{\xi + c}{\xi - c - 2}} \\ \frac{1}{\sqrt{2}}\sqrt{\frac{\xi + c}{\xi + c + 2}} & \frac{1}{\sqrt{2}}\sqrt{\frac{\xi - c}{\xi - c - 2}} \end{bmatrix}, \qquad \xi := \sqrt{c^2 + 8a^2} \tag{1.10}$$

is real with $\det M = \frac{1}{\sqrt{2(2a^2 - c - 1)}} > 0$ and $\mathrm{tr}(M) > 0$. In particular, all eigenvalues of M have positive real parts, that is, M is positively stable and hence is a hyperbolic matrix in the sense that all eigenvalues have non-zero real parts.

The second main result is the following.

Theorem 1.2. *There exists a unique $\theta \in \mathbb{R}$ such that*

$$\begin{bmatrix} x \\ y \end{bmatrix} = M \begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix}. \tag{1.11}$$

Moreover,

$$\tanh\left(-\sqrt{\frac{(\xi + c)(\xi - c - 2)}{(\xi - c)(\xi + c + 2)}}\right) < \theta < \tanh\left(\sqrt{\frac{(\xi - c)(\xi - c - 2)}{(\xi + c)(\xi + c + 2)}}\right) \tag{1.12}$$

and

$$e^{\theta} = \frac{\sqrt{\xi}}{2} \frac{(2a^2 - c - 1)\left(\sqrt{x^2 + \frac{c+1}{2a^2 - c - 1}} - x\right)}{\sqrt{(\xi + c)(\xi + c + 2)} - \sqrt{(\xi - c)(\xi - c - 2)}}.$$

Invertibility of M ensures uniqueness of θ, called the *hyperbolic angle* of the Karcher mean $\Lambda(A, B, C)$. An alternative expression for M is presented in Section 3.

2. Proof of Theorem 1.1

We list some basic properties of matrices in $\mathbf{H}_2$.

Proposition 2.1. *Let $A, B, C \in \mathbf{H}_2$.*

(i) $\operatorname{tr}(A) = \operatorname{tr}(A^{-1})$, *and* $\operatorname{tr}(A) = \operatorname{tr}(B)$ *if and only if A and B have the same spectrum;*

(ii) $\det(A+B) = 2 + \operatorname{tr}(AB^{-1})$;

(iii) $\operatorname{tr}(AB^{-1}) \geq 2$, *and* $\operatorname{tr}(AB^{-1}) = 2$ *if and only if* $A = B$;

(iv) *For* $x, y, z > 0$,

$$\begin{aligned}(x+y+z)^2 &\leq \det(xA + yB + zC) \\ &= x^2 + y^2 + z^2 + xy\operatorname{tr}(AB^{-1}) + xz\operatorname{tr}(AC^{-1}) + yz\operatorname{tr}(BC^{-1}),\end{aligned}$$

and equality holds only when $A = B = C$.

Furthermore, the following are equivalent:

(a) $\operatorname{tr}(AB^{-1}) = \operatorname{tr}(AC^{-1})$;

(b) $A^{-1/2}BA^{-1/2}$ *and* $A^{-1/2}CA^{-1/2}$ *have the same spectrum;*

(c) $A^{-1/2}BA^{-1/2}$ *and* $A^{-1/2}CA^{-1/2}$ *are unitary similar; and*

(d) $\delta(A, B) = \delta(A, C)$.

Proof. (i) is immediate from the fact that the map $x \mapsto x + x^{-1}$ is injective on $[1, \infty)$.

(ii) By a direct computation, $\det(I + X) = 1 + \det(X) + \operatorname{tr}(X)$ for every Hermitian matrix X. Then for $A, B, C \in \mathbb{P}$,

$$\begin{aligned}\det(A+B) &= \det(A^{1/2}(I + A^{-1/2}BA^{-1/2})A^{1/2}) = \det(A)\det(I + A^{-1/2}BA^{-1/2}) \\ &= \det(A)\left(1 + \det(A^{-1/2}BA^{-1/2}) + \operatorname{tr}(A^{-1/2}BA^{-1/2})\right) \\ &= \det(A) + \det(B) + \det(A)\operatorname{tr}(A^{-1}B).\end{aligned}$$

Hence (ii) follows immediately.

(iii) Let λ be an eigenvalue of $A^{1/2}B^{-1}A^{1/2}$. Since $\det(A^{1/2}B^{-1}A^{1/2}) = 1$, λ^{-1} is an eigenvalue of $A^{1/2}BA^{1/2} = I$ and hence $\operatorname{tr}(AB^{-1}) = \operatorname{tr}(A^{1/2}B^{-1}A^{1/2}) = \lambda + \lambda^{-1}$. This implies that $\operatorname{tr}(AB^{-1}) = \lambda + \lambda^{-1} \geq 2$, and that $\operatorname{tr}(AB^{-1}) = 2$ if and only if $\lambda = 1$ if and only if $A = B$.

(iv) Using the determinant formula $\det(A+B)$ in the proof of (ii),

$$\begin{aligned}\det(A+B+C) &= \det(A+B)+\det(C)+\det(C)\operatorname{tr}(C^{-1}(A+B))\\ &= \det(A)+\det(B)+\det(C)+\det(A)\operatorname{tr}(A^{-1}B)\\ &\quad +\det(C)(\operatorname{tr}(C^{-1}A)+\operatorname{tr}(C^{-1}B)).\end{aligned}$$

This together with (i) implies that for $A, B, C \in \mathbf{H}_2$,

$$\begin{aligned}\det(xA+yB+zC) &= x^2+y^2+z^2+xy\operatorname{tr}(A^{-1}B)+xz\operatorname{tr}(C^{-1}A)+yz\operatorname{tr}(C^{-1}B)\\ &= x^2+y^2+z^2+xy\operatorname{tr}(AB^{-1})+xz\operatorname{tr}(AC^{-1})+yz\operatorname{tr}(BC^{-1}).\end{aligned}$$

By the Minkowski's determinantal inequality,

$$\begin{aligned}\sqrt{\det(xA+yB+zC)} &\geq \sqrt{\det(xA+yB)}+\sqrt{\det(zC)}\\ &\geq \sqrt{\det(xA)}+\sqrt{\det(yB)}+\sqrt{\det(zC)}\\ &= x+y+z.\end{aligned}$$

Equality holds if and only if $zC = \alpha(xA+yB)$ and $yB = \beta(xA)$ for some positive α and β if and only if $A = B = C$ from $\det(A) = \det(B) = \det(C) = 1$.

The equivalences between (a), (b) and (c) are straightforward from (i) and

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

The equivalence between (a) and (d) appears in [9]. □

In the following, we let $\{A, B, C\} \subset \mathbf{H}_2$ be linearly independent. We have seen that

$$\Lambda(A,B,C) = xA+yB+zC \tag{2.13}$$

for some $x, y, z \geq 0$. We will first show that x, y, z are strictly positive by introducing a useful reduction on the triple $\{A, B, C\}$. Note that $A \neq B$ from linear independence. Let $d > 1$ be an eigenvalue of $A^{-1/2}BA^{-1/2}$. Pick a unitary matrix U such that

$$A^{-1/2}BA^{-1/2} = UDU^*, \qquad D := \operatorname{diag}(d, d^{-1}) \tag{2.14}$$

and set

$$W = U^*A^{-1/2}CA^{-1/2}U. \tag{2.15}$$

Then the triple $\{I, D, W\}$ is linearly independent because the congruence transformation $X \mapsto A^{1/2}UXU^*A^{1/2}$ maps I, D, W to A, B, C respectively. In particular, W is not diagonal. Moreover,

$$\text{(2.16)} \quad \mathrm{tr}(AB^{-1}) = \mathrm{tr}(D), \qquad \mathrm{tr}(AC^{-1}) = \mathrm{tr}(W), \qquad \mathrm{tr}(BC^{-1}) = \mathrm{tr}(DW^{-1}).$$

By invariancy under congruence transformations of the Karcher mean,

$$\begin{aligned} xA + yB + zC &= \Lambda(A, B, C) \\ &= \Lambda(A^{1/2}IA^{1/2}, A^{1/2}(A^{-1/2}BA^{-1/2})A^{1/2}, A^{1/2}(A^{-1/2}CA^{-1/2})A^{1/2}) \\ &= A^{1/2}\Lambda(I, UDU^*, A^{-1/2}CA^{-1/2})A^{1/2} \\ &= A^{1/2}\Lambda(I, UDU^*, UWU^*)A^{1/2} \\ &= A^{1/2}U\Lambda(I, D, W)U^*A^{1/2}. \end{aligned}$$

This implies that

$$\text{(2.17)} \qquad \Lambda(I, D, W) = xI + yD + zW.$$

Linear independence of A, B, C plays a key role for positivity of x, y, z.

Proposition 2.2. *The coefficients x, y, z are positive such that $x + y + z < 1$ and*

$$\text{(2.18)} \qquad \Lambda(A^{-1}, B^{-1}, C^{-1}) = xA^{-1} + yB^{-1} + zC^{-1}.$$

Proof. By Proposition 2.1 (iv), $x + y + z < 1$. To prove $x, y, z > 0$, it suffices to show $z \neq 0$ by permutation invariance of the Karcher mean. We may assume that $A = I$ and $B = D$, a diagonal matrix, by the previous reduction process. Suppose that $z = 0$. By the Karcher equation,

$$\begin{aligned} 0 &= \log(xI + yD) + \log(xI + yD)D^{-1} + \log(xI + yD)^{1/2}C^{-1}(xI + yD)^{1/2} \\ &= \log(xI + yD)^2D^{-1} + \log(xI + yD)^{1/2}C^{-1}(xI + yD)^{1/2}. \end{aligned}$$

That is,

$$(xI + yD)^2D^{-1} = (xI + yD)^{-1/2}C(xI + yD)^{-1/2}$$

and hence $C = (xI + yD)^3D^{-1}$ is a diagonal matrix. This contradicts to linear independence of $\{I, D, C\}$.

To prove (2.18) we let $Z := xA^{-1} + yB^{-1} + zC^{-1}$. By Proposition 2.1 (iv),

$$\begin{aligned}\det(Z) &= x^2 + y^2 + z^2 + xy\,\mathrm{tr}(A^{-1}B) + xz\,\mathrm{tr}(A^{-1}C) + yz\,\mathrm{tr}(B^{-1}C)\\ &= x^2 + y^2 + z^2 + xy\,\mathrm{tr}(AB^{-1}) + xz\,\mathrm{tr}(AC^{-1}) + yz\,\mathrm{tr}(BC^{-1})\\ &= \det(xA + yB + zC) = \det(\Lambda(A,B,C)) = 1.\end{aligned}$$

By computing trace

$$\begin{aligned}\mathrm{tr}[\Lambda(A,B,C)Z] &= \mathrm{tr}(xA + yB + zC)(xA^{-1} + yB^{-1} + zC^{-1})\\ &= 2(x^2 + y^2 + z^2 + xy\,\mathrm{tr}(AB^{-1}) + xz\,\mathrm{tr}(AC^{-1}) + yz\,\mathrm{tr}(BC^{-1}))\\ &= 2.\end{aligned}$$

By Proposition 2.1 (iii), $Z = \Lambda(A,B,C)^{-1}$ and hence

$$xA^{-1} + yB^{-1} + zC^{-1} = Z = \Lambda(A,B,C)^{-1} = \Lambda(A^{-1}, B^{-1}, C^{-1}).$$

□

Next, we shall show that $y = z$ under the trace condition

$$\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1}),$$

equivalently $d + d^{-1} = \mathrm{tr}(D) = \mathrm{tr}(W)$, from (2.16). An important fact is that $d > 1$ is also an eigenvalue of the non-diagonal matrix W, which follows from the equivalence between (a) and (b) of Proposition 2.1. Let V be a unitary matrix such that $W = VDV^*$, unitary diagonalization of W. By CS-decomposition of V (cf. Theorem VII. 1. 6 of [2]),

$$V = V_1 \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} V_2$$

for some unitary diagonal matrices V_1, V_2 and $0 \le \varphi < 2\pi$. It follows from $V_2\, D\, V_2^* = D$ that

$$W = V_1 \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} D \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} V_1^*. \tag{2.19}$$

Set

$$R_\varphi := \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}.$$

Then R_φ is an *orthogonal involutory* matrix, that is,

$$R_\varphi = R_\varphi^T = R_\varphi^{-1}.$$

Since

$$\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix},$$

we have from (2.19) that

$$W = D_1 R_\varphi D R_\varphi D_1^*, \tag{2.20}$$

where $D_1 := V_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is a unitary diagonal matrix.

Let

$$W_0 := R_\varphi D R_\varphi.$$

Then W_0 is not diagonal with determinant one, because W is not diagonal, and hence the triple $\{I, D, W_0\}$ is linearly independent. By Proposition 2.2, there exist positive reals x', y', z' such that

$$\Lambda(I, D, W_0) = x'I + y'D + z'W_0.$$

From the involutive property of the orthogonal matrix R_ϑ,

$$\begin{aligned} x'I + y'D + z'W_0 &= \Lambda(I, D, W_0) = \Lambda(I, D, R_\varphi D R_\varphi) = \Lambda(I, R_\varphi D R_\varphi, D) \\ &= R_\varphi \Lambda(I, D, R_\varphi^T D R_\varphi^T) R_\varphi = R_\varphi \Lambda(I, D, R_\varphi D R_\varphi) R_\varphi \\ &= R_\varphi \Lambda(I, D, W_0) R_\varphi \\ &= R_\varphi (x'I + y'D + z'W_0) R_\varphi = x' R_\varphi^2 + y' R_\varphi D R_\varphi + z' R_\varphi W_0 R_\varphi \\ &= x'I + y'W_0 + z' R_\varphi^2 D R_\varphi^2 = x'I + y'W_0 + z'D. \end{aligned}$$

By linearly independence, $y' = z'$.

Finally, by invariancy under permutations and congruence transformations of the Karcher mean and by the fact that D_1 is unitary diagonal,

$$\begin{aligned} xI + yD + zW &= \Lambda\left(I, D, W\right) = \Lambda\left(I, D, D_1 R_\varphi D R_\varphi D_1^*\right) \\ &= D_1 \Lambda\left(D_1^* D_1, D_1^* D D_1, R_\varphi D R_\varphi\right) D_1^* \\ &= D_1 \Lambda\left(I, D, W_0\right) D_1^* \\ &= D_1(x'I + y'D + y'W_0)D_1^* \\ &= x'D_1D_1^* + y'D_1DD_1^* + y'D_1W_0D_1^* \\ &= x'I + y'D + y'W. \end{aligned}$$

Again by linear independence,

$$x = x', \qquad y = y', \qquad z = y'.$$

In particular, $y = z$. This completes the proof of Theorem 1.1.

The following is immediate by Theorem 1.1 and by permutation invariancy of the Karcher mean.

Corollary 2.3 ([9])**.** *If* $\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1}) = \mathrm{tr}(BC^{-1})$, *then*

$$\Lambda(A, B, C) = \frac{A + B + C}{\sqrt{\det(A + B + C)}}. \tag{2.21}$$

For a triple $\{x, y, z\}$ in an inner product space,

$$||x - y|| = ||x - z|| \Longleftrightarrow ||m - y|| = ||m - z||, \qquad m := \frac{x + y + z}{3} \tag{2.22}$$

from $||x - 2y + z||^2 - ||x - 2z + y||^2 = 3\left(||x - y||^2 - ||x - z||^2\right).$

The following is an appropriate version of (2.22); a property of the centroid of isosceles triangles on the hyperbolic manifold $\mathbf{H}_2$.

Corollary 2.4. *Let* $A, B, C \in \mathbf{H}_2$ *be linearly independent. If* $\delta(A, B) = \delta(A, C)$, *then* $\delta(X, B) = \delta(X, C)$ *for* $X = \Lambda(A, B, C)$.

Proof. Suppose that $\delta(A,B)=\delta(A,C)$, that is, $\mathrm{tr}(AB^{-1})=\mathrm{tr}(AC^{-1})$ by Proposition 2.1. By Theorem 1.1, $\Lambda(A,B,C)=xA+yB+yC$ for some positive x,y. Since

$$\begin{aligned}\mathrm{tr}(XB^{-1}) &= \mathrm{tr}(xAB^{-1}+yI+yCB^{-1})=x\,\mathrm{tr}(AB^{-1})+2y+y\,\mathrm{tr}(CB^{-1})\\ &= x\,\mathrm{tr}(AC^{-1})+y\,\mathrm{tr}(BC^{-1})+2y=\mathrm{tr}(xAC^{-1}+yBC^{-1}+yI)\\ &= \mathrm{tr}(XC^{-1}),\end{aligned}$$

we have form Proposition 2.1 that $\delta(X,B)=\delta(X,C)$. □

3. Proof of Theorem 1.2

Let $A,B,C\in\mathbf{H}_2$ be linearly independent and let $x,y,z>0$ such that

$$\Lambda(A,B,C)=xA+yB+zC.$$

We shall introduce a quadratic surface containing the coefficient vector (x,y,z). By the determinantal identity of the Karcher mean,

$$1=\det\Lambda(A,B,C)=\det(xA+yB+zC)$$

and by Proposition 2.1 (iv), $x+y+z<1$ and

$$x^2+y^2+z^2+xy\,\mathrm{tr}(AB^{-1})+xz\,\mathrm{tr}(AC^{-1})+yz\,\mathrm{tr}(BC^{-1})=1.$$

Set

$$a=\frac{1}{2}\mathrm{tr}(AB^{-1}),\quad b=\frac{1}{2}\mathrm{tr}(AC^{-1}),\quad c=\frac{1}{2}\mathrm{tr}(BC^{-1}). \tag{3.23}$$

Then

$$x^2+y^2+z^2+2axy+2bxz+2cyz=1. \tag{3.24}$$

That is, the vector $(x,y,z)\in(0,1)^3$ lies in the quadric surface determined by a,b,c. By Proposition 2.1 (iii),

$$a,b,c>1. \tag{3.25}$$

In terms of quadratic form, (3.24) can be written as

$$\mathbf{v}^TQ\mathbf{v}=1,\qquad \mathbf{v}^T:=(x,y,z) \tag{3.26}$$

where

$$Q = Q(A, B, C) := \begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}.$$

We note that the symmetric matrix Q depends only a, b, c, but the coefficient vector $\mathbf{v}^T = (x, y, z)$ depends on the Karcher mean $\Lambda(A, B, C)$. To determine the type of quadratic surface, we need to compute the inertia of Q. The characteristic polynomial $\det(Q - \lambda I)$ of Q is

$$-\lambda^3 + 3\lambda^2 + (a^2 + b^2 + c^2 - 3)\lambda - (a^2 + b^2 + c^2) + 2abc + 1. \tag{3.27}$$

Let $\lambda_1 \leq \lambda_2 \leq \lambda_3$ be eigenvalues of $Q = Q(A, B, C)$. Note that

$$\lambda_1 + \lambda_2 + \lambda_3 = 3. \tag{3.28}$$

It is shown in [8] that by linear independency of $\{A, B, C\}$,

$$\det Q(A, B, C) = -(a^2 + b^2 + c^2) + 2abc + 1 > 0. \tag{3.29}$$

(One can restrict (A, B, C) to a linearly independent triple $(I, D, R_\varphi D R_\varphi)$, where $D \neq I$ is diagonal, via the reduction process) By (3.25) and Sylvester's criterion on principal minors, Q is not positive semidefinite, that is, one of eigenvalues of Q is strictly negative. This together with (3.29) leads to

$$\lambda_1 \leq \lambda_2 < 0 < \lambda_3. \tag{3.30}$$

We then conclude that the quadratic surface (3.24) determined by a, b, c is a hyperboloid of two sheets.

Consider an orthogonal diagonalization of the symmetric matrix Q;

$$Q(A, B, C) = U^T \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) U.$$

Letting

$$\mathbf{w} = [w_1, w_2, w_3]^T := U\mathbf{v}$$

leads to

$$\mathbf{w}^T \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) \mathbf{w} = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2 = 1. \tag{3.31}$$

Using the parametric representation of a hyperboloid of two sheets,

$$\begin{aligned} w_1 &= \frac{\sinh(u)\cosh(v)}{\sqrt{-\lambda_1}}, \\ w_2 &= \frac{\sinh(v)}{\sqrt{-\lambda_2}}, \\ w_3 &= \frac{\cosh(u)\cosh(v)}{\sqrt{\lambda_3}}, \end{aligned}$$

where $u, v \in \mathbb{R}$, one can see that the 3×3 real nonsingular matrix

$$K := U^T \mathrm{diag}\left(\frac{1}{\sqrt{-\lambda_1}}, \frac{1}{\sqrt{-\lambda_2}}, \frac{1}{\sqrt{\lambda_3}}\right) \tag{3.32}$$

satisfies

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = K \begin{bmatrix} \sinh(u)\cosh(v) \\ \sinh(v) \\ \cosh(u)\cosh(v) \end{bmatrix} = K \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cosh(u)\cosh(v) \\ \sinh(v) \\ \sinh(u)\cosh(v) \end{bmatrix} \tag{3.33}$$

for some $u, v \in \mathbb{R}$. Finding an explicit form of K in terms of a, b, c is equivalent to that of $\lambda_j, j = 1, 2, 3$.

It follows from $\lambda_1 + \lambda_2 + \lambda_3 = 3$ and $\det Q = \lambda_1\lambda_2\lambda_3$ that

$$\lambda_3 - 3 = -(\lambda_1 + \lambda_2) > 0$$

and

$$\lambda_1 = \frac{3 - \lambda_3 - \sqrt{(\lambda_3 - 3)^2 - 4\lambda_3^{-1}[-(a^2 + b^2 + c^2) + 2abc + 1]}}{2}, \tag{3.34}$$

$$\lambda_2 = \frac{3 - \lambda_3 + \sqrt{(\lambda_3 - 3)^2 - 4\lambda_3^{-1}[-(a^2 + b^2 + c^2) + 2abc + 1]}}{2}. \tag{3.35}$$

Moreover,

$$\lambda_1 = \lambda_2 \Longleftrightarrow \lambda_3^3 - 6\lambda_3^2 + 9\lambda_3 - 4[-(a^2 + b^2 + c^2) + 2abc + 1] = 0. \tag{3.36}$$

It is not easy to find an explicit form of λ_3 in terms of a, b, c. However, if two of a, b, c are equal, which is the case of our main concern, then we have an explicit form of the eigenvalues.

In the following we assume that $a = b$, that is, $\mathrm{tr}(AB^{-1}) = \mathrm{tr}(AC^{-1})$. Then

$$\det Q = (c - 1)(2a^2 - (1 + c)). \tag{3.37}$$

It then follows from $c > 1$ and $\det Q > 0$ that

$$2a^2 > c + 1. \tag{3.38}$$

One can show directly that

$$\lambda_3 = \frac{c + 2 + \sqrt{c^2 + 8a^2}}{2}$$

and

$$\{\lambda_1, \lambda_2\} = \left\{1 - c, \frac{c + 2 - \sqrt{c^2 + 8a^2}}{2}\right\}.$$

Furthermore,

$$\lambda_1 = \lambda_2 \Longleftrightarrow a(= b) = c, \tag{3.39}$$

and

$$\lambda_1 = \begin{cases} 1 - c, & \text{if } a = b \leq c \\ \frac{c+2-\sqrt{c^2+8a^2}}{2}, & \text{if } a = b > c. \end{cases} \tag{3.40}$$

Set

$$\xi = \xi(A, B, C) := \sqrt{c^2 + 8a^2}.$$

By (3.38),

$$\xi > c + 2 \tag{3.41}$$

and (1.9) holds true. By a direct computation, the real matrix M in Theorem 1.2

$$M := \begin{bmatrix} \sqrt{\frac{\xi-c}{\xi(\xi+c+2)}} & -\sqrt{\frac{\xi+c}{\xi(\xi-c-2)}} \\ \sqrt{\frac{\xi+c}{2\xi(\xi+c+2)}} & \sqrt{\frac{\xi-c}{2\xi(\xi-c-2)}} \end{bmatrix} \tag{3.42}$$

is nonsingular with

$$\det M = \frac{1}{\sqrt{2(2a^2 - c - 1)}} = \frac{\sqrt{c-1}}{\sqrt{2 \det Q}} > 0.$$

Proof of Theorem 1.2.

By Theorem 1.1, $y = z$. That is, $\mathbf{v} = (x, y, y)^T$.

Set

$$\alpha := \sqrt{1 - \frac{c}{\sqrt{c^2 + 8a^2}}}, \qquad \beta := \sqrt{1 + \frac{c}{\sqrt{c^2 + 8a^2}}}. \tag{3.43}$$

Then

$$\alpha^2 + \beta^2 = 2, \qquad \alpha\beta = \sqrt{\frac{8a^2}{c^2 + 8a^2}}. \tag{3.44}$$

Case 1: $a = b \leq c$. In this case,

$$\lambda_1 = 1 - c \leq \lambda_2 = \frac{c + 2 - \sqrt{c^2 + 8a^2}}{2} \leq \lambda_3 = \frac{c + 2 + \sqrt{c^2 + 8a^2}}{2}.$$

From $\lambda_3 - \lambda_2 = \sqrt{c^2 + 8a^2}$ and $\lambda_1 + \lambda_2 + \lambda_3 = 3$, we have

$$\begin{aligned} \alpha &= \sqrt{1 - \frac{1 - \lambda_1}{\lambda_3 - \lambda_2}} = \sqrt{\frac{\lambda_3 - \lambda_2 + \lambda_1 - 1}{\lambda_3 - \lambda_2}} = \sqrt{\frac{2(1 - \lambda_2)}{\lambda_3 - \lambda_2}}, \\ \beta &= \sqrt{1 + \frac{1 - \lambda_1}{\lambda_3 - \lambda_2}} = \sqrt{\frac{\lambda_3 - \lambda_2 - \lambda_1 + 1}{\lambda_3 - \lambda_2}} = \sqrt{\frac{2(\lambda_3 - 1)}{\lambda_3 - \lambda_2}}. \end{aligned}$$

Let

$$U := \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}\beta & \frac{1}{2}\alpha & \frac{1}{2}\alpha \\ \frac{1}{\sqrt{2}}\alpha & \frac{1}{2}\beta & \frac{1}{2}\beta \end{bmatrix}.$$

By (3.43),

$$\begin{aligned} UU^T &= \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}\beta & \frac{1}{2}\alpha & \frac{1}{2}\alpha \\ \frac{1}{\sqrt{2}}\alpha & \frac{1}{2}\beta & \frac{1}{2}\beta \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}}\beta & \frac{1}{\sqrt{2}}\alpha \\ -\frac{1}{\sqrt{2}} & \frac{1}{2}\alpha & \frac{1}{2}\beta \\ \frac{1}{\sqrt{2}} & \frac{1}{2}\alpha & \frac{1}{2}\beta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}(\alpha^2 + \beta^2) & 0 \\ 0 & 0 & \frac{1}{2}(\alpha^2 + \beta^2) \end{bmatrix} = I. \end{aligned}$$

It follows from

$$\alpha\beta = \sqrt{\frac{8a^2}{c^2 + 8a^2}} = \frac{2\sqrt{2}a}{\lambda_3 - \lambda_2} \tag{3.45}$$

and

$$\lambda_2\beta^2 + \lambda_3\alpha^2 = 2, \qquad \lambda_2\alpha^2 + \lambda_3\beta^2 = 4 - 2\lambda_1 > 0 \tag{3.46}$$

that for $T := \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$,

$$U^T T U = \begin{bmatrix} \frac{\lambda_2\beta^2+\lambda_3\alpha^2}{2} & \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_2) & \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_2) \\ \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_2) & \frac{\lambda_2\alpha^2+\lambda_3\beta^2}{4}+\frac{\lambda_1}{2} & \frac{\lambda_2\alpha^2+\lambda_3\beta^2}{4}-\frac{\lambda_1}{2} \\ \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_2) & \frac{\lambda_2\alpha^2+\lambda_3\beta^2}{4}-\frac{\lambda_1}{2} & \frac{\lambda_2\alpha^2+\lambda_3\beta^2}{4}+\frac{\lambda_1}{2} \end{bmatrix} = Q.$$

Letting

$$\mathbf{w} = [w_1, w_2, w_3]^T := U\mathbf{v}$$

leads to

$$\mathbf{w}^T T \mathbf{w} = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2 = 1. \tag{3.47}$$

From

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ y \end{bmatrix} = U^T\mathbf{w} = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}}\beta & \frac{1}{\sqrt{2}}\alpha \\ -\frac{1}{\sqrt{2}} & \frac{1}{2}\alpha & \frac{1}{2}\beta \\ \frac{1}{\sqrt{2}} & \frac{1}{2}\alpha & \frac{1}{2}\beta \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix},$$

we have that

$$w_1 = 0$$

and

$$x = \frac{1}{\sqrt{2}}(\alpha w_3 - \beta w_2), \qquad y = \frac{1}{2}(\alpha w_2 + \beta w_3). \tag{3.48}$$

By (3.47),

$$w_2 = \frac{\sinh\theta}{\sqrt{-\lambda_2}}, \quad w_3 = \frac{\cosh\theta}{\sqrt{\lambda_3}} \tag{3.49}$$

for some $\theta \in \mathbb{R}$. Hence

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{\beta}{\sqrt{-2\lambda_2}} & \frac{\alpha}{\sqrt{2\lambda_3}} \\ \frac{\alpha}{2\sqrt{-\lambda_2}} & \frac{\beta}{2\sqrt{\lambda_3}} \end{bmatrix} \begin{bmatrix} \sinh\theta \\ \cosh\theta \end{bmatrix} = \begin{bmatrix} \frac{\alpha}{\sqrt{2\lambda_3}} & -\frac{\beta}{\sqrt{-2\lambda_2}} \\ \frac{\beta}{2\sqrt{\lambda_3}} & \frac{\alpha}{2\sqrt{-\lambda_2}} \end{bmatrix} \begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix}. \tag{3.50}$$

One can see that

$$\begin{bmatrix} \frac{\alpha}{\sqrt{2\lambda_3}} & -\frac{\beta}{\sqrt{-2\lambda_2}} \\ \frac{\beta}{2\sqrt{\lambda_3}} & \frac{\alpha}{2\sqrt{-\lambda_2}} \end{bmatrix} = M.$$

This shows existence of θ. Uniqueness is direct from invertibility of M.

Since $x, y > 0$, we get from (3.50) that

$$\tanh^{-1}\left(\frac{\alpha}{\beta}\frac{\sqrt{-\lambda_2}}{\sqrt{\lambda_3}}\right) > \theta$$

and

$$\theta > \tanh^{-1}\left(-\frac{\beta}{\alpha}\frac{\sqrt{-\lambda_2}}{\sqrt{\lambda_3}}\right).$$

That is,

$$\tanh^{-1}\left(-\frac{\beta}{\alpha}\frac{\sqrt{-\lambda_2}}{\sqrt{\lambda_3}}\right) < \theta < \tanh^{-1}\left(\frac{\alpha}{\beta}\frac{\sqrt{-\lambda_2}}{\sqrt{\lambda_3}}\right). \tag{3.51}$$

The upper and lower bounds are same with that of (1.12).

Note that

$$x = \frac{\alpha}{\sqrt{2\lambda_3}}\cosh\theta - \frac{\beta}{\sqrt{-2\lambda_2}}\sinh\theta,$$

equivalently,

$$e^{2\theta}\left(\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}\right) - 2xe^{\theta} + \frac{\alpha}{\sqrt{2\lambda_3}} + \frac{\beta}{\sqrt{-2\lambda_2}} = 0.$$

From (3.46), we have that

$$\frac{\alpha}{\sqrt{2\lambda_3}} < \frac{\beta}{\sqrt{-2\lambda_2}},$$

and hence

$$\begin{aligned} e^{\theta} &= \frac{x \pm \sqrt{x^2 - \left(\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}\right)\left(\frac{\alpha}{\sqrt{2\lambda_3}} + \frac{\beta}{\sqrt{-2\lambda_2}}\right)}}{\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}} \\ &= \frac{x \pm \sqrt{x^2 - \left(\frac{\alpha^2}{2\lambda_3} + \frac{\beta^2}{2\lambda_2}\right)}}{\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}} = \frac{x \pm \sqrt{x^2 + \frac{\lambda_1 - 2}{\lambda_2\lambda_3}}}{\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}}, \end{aligned}$$

where the last equality follows from (3.46). As $\lambda_1 - 2$ is negative and $\lambda_2\lambda_3$ is negative,

$$\sqrt{x^2 + \frac{\lambda_1 - 2}{\lambda_2\lambda_3}} > \sqrt{x^2} = x$$

and hence

$$x + \sqrt{x^2 + \frac{\lambda_1 - 2}{\lambda_2\lambda_3}} > 0, \quad x - \sqrt{x^2 + \frac{\lambda_1 - 2}{\lambda_2\lambda_3}} < 0.$$

This implies that

$$e^{\theta} = \frac{x - \sqrt{x^2 + \frac{\lambda_1 - 2}{\lambda_2\lambda_3}}}{\frac{\alpha}{\sqrt{2\lambda_3}} - \frac{\beta}{\sqrt{-2\lambda_2}}} = \frac{\sqrt{\xi}}{2}\frac{(2a^2 - c - 1)\left(\sqrt{x^2 + \frac{c+1}{2a^2 - c - 1}} - x\right)}{\sqrt{(\xi + c)(\xi + c + 2)} - \sqrt{(\xi - c)(\xi - c - 2)}}.$$

Case 2. $a = b > c$. In this case,

$$\lambda_1 = \frac{c+2-\sqrt{c^2+8a^2}}{2} \le \lambda_2 = 1-c \le \lambda_3 = \frac{c+2+\sqrt{c^2+8a^2}}{2}.$$

Set

$$V := \begin{bmatrix} -\frac{1}{\sqrt{2}}\beta & \frac{1}{2}\alpha & \frac{1}{2}\alpha \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\alpha & \frac{1}{2}\beta & \frac{1}{2}\beta \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} U.$$

Then $V^TV = VV^T = I$ and for $T := \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$,

$$V^TTV = \begin{bmatrix} \frac{\lambda_1\beta^2+\lambda_3\alpha^2}{2} & \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_1) & \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_1) \\ \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_1) & \frac{\lambda_1\alpha^2+\lambda_3\beta^2}{4}+\frac{\lambda_2}{2} & \frac{\lambda_1\alpha^2+\lambda_3\beta^2}{4}-\frac{\lambda_2}{2} \\ \frac{\alpha\beta}{2\sqrt{2}}(\lambda_3-\lambda_1) & \frac{\lambda_1\alpha^2+\lambda_3\beta^2}{4}-\frac{\lambda_2}{2} & \frac{\lambda_1\alpha^2+\lambda_3\beta^2}{4}+\frac{\lambda_2}{2} \end{bmatrix} = Q.$$

Set $\mathbf{u} = (u_1, u_2, u_3)^T =: V\mathbf{v}$. Then $\mathbf{u}^T T\mathbf{u} = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \lambda_3 u_3^2 = 1$ and

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ y \end{bmatrix} = V^T\mathbf{u} = \begin{bmatrix} -\frac{1}{\sqrt{2}}\beta & 0 & \frac{1}{\sqrt{2}}\alpha \\ \frac{1}{2}\alpha & -\frac{1}{\sqrt{2}} & \frac{1}{2}\beta \\ \frac{1}{2}\alpha & \frac{1}{\sqrt{2}} & \frac{1}{2}\beta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Hence $u_2 = 0$ and

$$x = \frac{1}{\sqrt{2}}(\alpha u_3 - \beta u_1), \qquad y = \frac{1}{2}(\alpha u_1 + \beta u_3).$$

There exists θ^* such that

$$u_1 = \frac{\sinh\theta^*}{\sqrt{-\lambda_1}}, \quad u_3 = \frac{\cosh\theta^*}{\sqrt{\lambda_3}}.$$

As Case 1, we get that

$$\tanh^{-1}\left(-\frac{\beta}{\alpha}\frac{\sqrt{-\lambda_1}}{\sqrt{\lambda_3}}\right) < \theta^* < \tanh^{-1}\left(\frac{\alpha}{\beta}\frac{\sqrt{-\lambda_1}}{\sqrt{\lambda_3}}\right) \tag{3.52}$$

and

$$\begin{cases} x = \frac{\alpha}{\sqrt{2\lambda_3}}\cosh\theta^* - \frac{\beta}{\sqrt{-2\lambda_1}}\sinh\theta^*, \\ y = \frac{\beta}{2\sqrt{\lambda_3}}\cosh\theta^* + \frac{\alpha}{2\sqrt{-\lambda_1}}\sinh\theta^*. \end{cases} \tag{3.53}$$

The remaining part of proof is similar to Case 1. This completes the proof of Theorem 1.2.

In terms of eigenvalues $\lambda_1 \leq \lambda_2 < \lambda_3$ of $Q = Q(A, B, C)$, we have the following alternative expression of x and y.

Corollary 3.1. *If $a = b \leq c$, there exists unique $\theta \in \mathbb{R}$ such that*

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{\lambda_3 - \lambda_2}} \begin{bmatrix} \sqrt{\frac{1-\lambda_2}{\lambda_3}} & -\sqrt{\frac{1-\lambda_3}{\lambda_2}} \\ \sqrt{\frac{\lambda_3-1}{2\lambda_3}} & \sqrt{\frac{\lambda_2-1}{2\lambda_2}} \end{bmatrix} \begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix}. \tag{3.54}$$

In particular,

$$\tanh^{-1}\left(-\sqrt{\frac{\lambda_2(1-\lambda_3)}{\lambda_3(1-\lambda_2)}}\right) < \theta < \tanh^{-1}\left(\sqrt{\frac{\lambda_2(\lambda_2-1)}{\lambda_3(\lambda_3-1)}}\right).$$

If $a = b > c$, then there exists unique θ such that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{\lambda_3 - \lambda_1}} \begin{bmatrix} \sqrt{\frac{1-\lambda_1}{\lambda_3}} & -\sqrt{\frac{1-\lambda_3}{\lambda_1}} \\ \sqrt{\frac{\lambda_3-1}{2\lambda_3}} & \sqrt{\frac{\lambda_1-1}{2\lambda_1}} \end{bmatrix} \begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix}. \tag{3.55}$$

In particular,

$$\tanh^{-1}\left(-\sqrt{\frac{\lambda_1(1-\lambda_3)}{\lambda_3(1-\lambda_1)}}\right) < \theta < \tanh^{-1}\left(\sqrt{\frac{\lambda_1(\lambda_1-1)}{\lambda_3(\lambda_3-1)}}\right).$$

Remark 3.2. Note from [9] and (3.39) that $x = y$ if and only if $a = b = c$ if and only if $\lambda_1 = \lambda_2$, in which case, $x = y = \frac{1}{\sqrt{\det(A+B+C)}}$.

Suppose that $a = b = c$. Then $\xi = 3a$ and

$$\alpha = \sqrt{\frac{2}{3}}, \qquad \beta = \frac{2}{\sqrt{3}}, \qquad \lambda_1 = \lambda_2 = 1 - a, \qquad \lambda_3 = 2a + 1.$$

It then follows from (3.50) that $\theta = 0$ and

$$x = y = \frac{1}{\sqrt{3\lambda_3}} = \frac{1}{\sqrt{3(2a+1)}} = \frac{1}{\sqrt{\det(A+B+C)}}.$$

Moreover,

$$M = \frac{1}{\sqrt{3}} \begin{bmatrix} \frac{1}{\sqrt{2a+1}} & -\frac{\sqrt{2}}{\sqrt{a-1}} \\ \frac{1}{\sqrt{2a+1}} & \frac{1}{\sqrt{2(a-1)}} \end{bmatrix}.$$

Question. Does $\theta = 0$ a necessary condition for $a = b = c$?

Remark 3.3. Note that for $a = b$, (3.24) becomes

$$x^2 + 2y^2 + 4axy + 2cy^2 = 1, \tag{3.56}$$

alternatively, $x^2 + (y^{'})^2 + kxy^{'} = 1$, where $y^{'} := \sqrt{2(1+c)}y$ and $k := \frac{4a}{\sqrt{2(1+c)}} > 2$. Setting $\begin{bmatrix} u \\ v \end{bmatrix} := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y' \end{bmatrix}$ yields $\left(\frac{k}{2}+1\right)u^2 - \left(\frac{k}{2}-1\right)v^2 = 1$. It follows from $k > 2$ that

$$u = \frac{1}{\sqrt{\frac{k}{2}+1}} \cosh t, \qquad v = \frac{1}{\sqrt{\frac{k}{2}-1}} \sinh t$$

for some $t \in \mathbb{R}$. That is,

$$\begin{aligned} x &= \frac{1}{\sqrt{k+2}} \cosh(t) - \frac{1}{\sqrt{k-2}} \sinh(t), \\ y^{'} &= \frac{1}{\sqrt{k+2}} \cosh(t) + \frac{1}{\sqrt{k-2}} \sinh(t) \end{aligned}$$

and hence

$$\begin{bmatrix} x \\ y \end{bmatrix} = N \begin{bmatrix} \cosh(t) \\ \sinh(t) \end{bmatrix}, \qquad N := \begin{bmatrix} \frac{1}{\sqrt{k+2}} & -\frac{1}{\sqrt{k-2}} \\ \frac{1}{\sqrt{2(1+c)(k+2)}} & \frac{1}{\sqrt{2(1+c)(k-2)}} \end{bmatrix}.$$

The N is also hyperbolic. The following shows that the matrix M is much better than N. Suppose that $a = b = c$. By the preceding remark, $\theta = 0$. Solving

$$\begin{bmatrix} \cosh\theta \\ \sinh\theta \end{bmatrix} = M^{-1}N \begin{bmatrix} \cosh(t) \\ \sinh(t) \end{bmatrix}$$

yields

$$t = \tanh^{-1}\left(-\frac{-\frac{1}{\sqrt{2}}\sqrt{\frac{4a}{4a+2}}\frac{1}{\sqrt{k+2}} + \sqrt{\frac{2a}{4a+2}}\frac{1}{\sqrt{2(1+a)(k+2)}}}{\frac{1}{\sqrt{2}}\sqrt{\frac{4a}{4a+2}}\frac{1}{\sqrt{k-2}} + \sqrt{\frac{2a}{4a+2}}\frac{1}{\sqrt{2(1+a)(k-2)}}} \right).$$

Example 3.4. Let

$$D = \frac{1}{100}\begin{bmatrix} 101+\sqrt{201} & 0 \\ 0 & 101-\sqrt{201} \end{bmatrix}, \quad W = \frac{101}{20100}\begin{bmatrix} 201+\sqrt{201} & \sqrt{\frac{151}{10050}\frac{20100}{101}} \\ \sqrt{\frac{151}{10050}\frac{20100}{101}} & 201-\sqrt{201} \end{bmatrix}.$$

Note that $\text{tr}(D) = \text{tr}(W) = \text{tr}(WD^{-1}) = 2.02$ and $\det(D) = \det(W) = 1$, so that $a = 1.01$. By the previous remark,

$$M = \frac{1}{\sqrt{3}} \begin{bmatrix} \frac{1}{\sqrt{2.02+1}} & -\frac{\sqrt{2}}{\sqrt{1.01-1}} \\ \frac{1}{\sqrt{2.02+1}} & \frac{1}{\sqrt{2(1.01-1)}} \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} \frac{1}{\sqrt{3.02}} & -\frac{\sqrt{2}}{\sqrt{0.01}} \\ \frac{1}{\sqrt{3.02}} & \frac{1}{\sqrt{0.02}} \end{bmatrix}.$$

The matrix M has positive real eigenvalues; $\lambda_1(M) \approx 1.31099 < \lambda_2(M) \approx 3.10372$. Using the Matrix Means Toolbox by Bini and Iannazzo for MATLAB we can see numerically that $\Lambda(I, D, W) \approx \begin{bmatrix} 1.0723 & 0.0409 \\ 0.0409 & 0.9310 \end{bmatrix}$ and $x \approx 0.3273, y \approx 0.3336, \theta \approx 0.000603531$.

Example 3.5. Let

$$D = \begin{bmatrix} 18 + \sqrt{323} & 0 \\ 0 & \frac{1}{18+\sqrt{323}} \end{bmatrix}, \quad W = \begin{bmatrix} 18 + \frac{321}{\sqrt{323}} & 2\sqrt{\frac{322}{323}} \\ 2\sqrt{\frac{322}{323}} & 18 - \frac{321}{\sqrt{323}} \end{bmatrix}.$$

Then $a = b = \frac{\text{tr}(D)}{2} = \frac{\text{tr}(W)}{2} = 18, c = \frac{\text{tr}(DW^{-1})}{2} = 3, \xi = \sqrt{c^2 + 8a^2} = 51$, and

$$M = \frac{1}{\sqrt{51}} \begin{bmatrix} \sqrt{\frac{48}{56}} & -\sqrt{\frac{54}{46}} \\ \frac{1}{\sqrt{2}}\sqrt{\frac{54}{56}} & \frac{1}{\sqrt{2}}\sqrt{\frac{48}{46}} \end{bmatrix} = \frac{1}{\sqrt{51}} \begin{bmatrix} \sqrt{\frac{6}{7}} & -\sqrt{\frac{27}{23}} \\ \frac{1}{\sqrt{2}}\sqrt{\frac{27}{28}} & \frac{1}{\sqrt{2}}\sqrt{\frac{24}{23}} \end{bmatrix}.$$

It has complex eigenvalues; $\lambda(M) \approx 0.115393 \pm 0.120617i$. Numerically, $\Lambda(I, D, W) \approx \begin{bmatrix} 9.8980 & 0.2728 \\ 0.2728 & 0.1085 \end{bmatrix}$, and $x \approx 0.0855, y \approx 0.1366, \theta \approx 0.332439$.

We closed this section with a sufficient condition on the linearly independent triples (A, B, C) in $\mathbb{P}$ satisfying

$$\Lambda(A, B, C) = xA + y(B + C).$$

Let (A, B, C) be a linearly independent triple in $\mathbb{P}$. Set

$$A_0 = \frac{A}{\sqrt{\det(A)}}, \quad B_0 = \frac{B}{\sqrt{\det(B)}}, \quad C_0 = \frac{C}{\sqrt{\det(C)}}.$$

Assume that $\text{tr}(A_0 B_0^{-1}) = \text{tr}(A_0 C_0^{-1})$, equivalently

$$\frac{\text{tr}(AB^{-1})}{\text{tr}(AC^{-1})} = \sqrt{\det(B^{-1}C)}. \tag{3.57}$$

By Theorem 1.1, there exist unique positive real numbers x_0, y_0 such that $\Lambda(A_0, B_0, C_0) = x_0A_0 + y_0B_0 + y_0C_0$ and hence by the determinantal identity for the Karcher mean

$$\begin{aligned}\Lambda(A, B, C) &= \det(ABC)^{\frac{1}{6}}\Lambda(A_0, B_0, C_0) = \det(ABC)^{\frac{1}{6}}(x_0A_0 + y_0B_0 + y_0C_0) \\ &= \det(A^{-2}BC)^{\frac{1}{6}}x_0A + \det(AB^{-2}C)^{\frac{1}{6}}y_0B + \det(ABC^{-2})^{\frac{1}{6}}y_0C.\end{aligned}$$

Setting

$$x := \det(A^{-2}BC)^{\frac{1}{6}}x_0, \qquad y := \det(AB^{-2}C)^{\frac{1}{6}}y_0, \qquad z := \det(ABC^{-2})^{\frac{1}{6}}y_0$$

yields $\Lambda(A, B, C) = xA+yB+zC$. We note that $y = z$ if and only if $\det(B) = \det(C)$, in which case $\operatorname{tr}(AB^{-1}) = \operatorname{tr}(AC^{-1})$ from (3.57).

Corollary 3.6. *For every linearly independent triple* (A, B, C) *in* $\mathbb{P}$ *satisfying* $\det(B) = \det(C)$ *and* $\operatorname{tr}(AB^{-1}) = \operatorname{tr}(AC^{-1})$,

$$\Lambda(A, B, C) = xA + y(B + C)$$

for some positive real numbers x, y.

4. Final remarks

Let A, B, C be linearly independent 2×2 positive definite matrices of determinant 1, and let $a = \frac{1}{2}\operatorname{tr}(AB^{-1}), b = \frac{1}{2}\operatorname{tr}(AC^{-1}), c = \frac{1}{2}\operatorname{tr}(BC^{-1})$. By Proposition 2.1, there exist unique positive real numbers x, y, z such that $\Lambda(A, B, C) = xA + yB + zC$. One can see that the ALM mean [1] is also linearly representable in the sense that $\operatorname{Alm}(A, B, C) = x'A + y'B + z'C$ for some unique *nonnegative* real numbers x', y', z'. Similarly for the BMP mean [7]. This raises the question whether every matrix geometric mean admits a linear representation or not. It is known only by numerical tests that the coefficient vectors (x, y, z) and (x', y', z') are different, alternatively $\Lambda(A, B, C) \neq \operatorname{Alm}(A, B, C)$ for general (A, B, C), a standing open problem after the Karcher mean was appeared [4]. We list some related open problems including that of finding a closed-form of x and y, under $a = b$, and of finding a geometric meaning of the hyperbolic angle θ of the Karcher mean.

1. By [9], $a = b = c$ if and only if $x = y = z$. Does the trace condition $a = b$ a necessary and sufficient condition for $y = z$?

2. The Karcher equation in this paper is involved only for positivity of x, y, z. One can see that $y' = z'$ when $a = b$. Does the coefficients x', y', z' are all positive?
3. The invertible matrix M in (1.11) depends only on $a = b, c$. One can show that Theorem 1.2 holds for the ALM mean:
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = M \begin{bmatrix} \cosh\theta' \\ \sinh\theta' \end{bmatrix}$$
for some (unique) θ'. If $\theta \neq \theta'$, then $\Lambda(A, B, C) \neq \mathrm{Alm}(A, B, C)$. Let $D = \mathrm{diag}(2, 1/2)$ and $W = \begin{bmatrix} \frac{3}{4} & \frac{\sqrt{5}}{4} \\ \frac{\sqrt{5}}{4} & \frac{7}{4} \end{bmatrix}$. Then $\mathrm{tr}(D) = \mathrm{tr}(W) = \frac{5}{2}, \mathrm{tr}(DW^{-1}) = \frac{31}{8}$ and
$$\mathrm{Alm}(I, D, W) = \begin{bmatrix} 1.0865 & 0.1595 \\ 0.1595 & 0.9438 \end{bmatrix} \neq \Lambda(I, D, W) = \begin{bmatrix} 1.0863 & 0.1592 \\ 0.1592 & 0.9439 \end{bmatrix}.$$
In this case
$$M \approx \begin{bmatrix} 0.2553 & -2.3494 \\ 0.3047 & 1.6613 \end{bmatrix},$$
and $\theta \approx -0.0197, \theta' \approx -0.0203$.
4. Some properties of the real matrix K in (3.32), like hyperbolicity or negative stability, are clearly of interest for further work.

Acknowledgments

The authors are grateful to Professor Hiai for pointing out the CS decomposition that shortens the original proof. The work of E. Ghiglioni is partially supported by Grants CONICET-PIP 152, PICT-2015-1505 and UNLP11X829. The work of Y. Lim was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) No. 2015R1A3A2031159.

References

[1] T. Ando, C. K. Li and R. Mathias, Geometric means, Linear Algebra Appl. **385** (2004), 305–334.
[2] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[3] R. Bhatia, *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007.

[4] R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. **413** (2006), 594–618.

[5] R. Bhatia and R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. **353** (2012), 1453–1467.

[6] R. Bhatia and T. Jain, The geometric mean of exponentials of Pauli matrices, Journal of Ramanujan Mathematical Society **30** (2015), 199–204.

[7] D. Bini, B. Meini, and F. Poloni, An effective matrix geometric mean satisfying the Ando-Li-Mathias properties, Math. Comp. **79** (2010), 437–452.

[8] H. Choi, E. Ghiglioni, and Y. Lim, The Karcher mean of three variables and quadric surfaces, J. Math. Anal. Appl. **490** (2020), 124321.

[9] E. Ghiglioni, Y. Lim and M. Pálfia, The Karcher mean of linearly independent triples, Linear Algebra Appl. **610** (2021), 203-221.

[10] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, J. Ramanujan Math. Soc. **27** (2012), 509–521.

[11] H. Karcher, Riemannian center of mass and mollifier smoothing, Comm. Pure Appl. Math. **30** (1977), 509–541.

[12] S. Lang, *Fundamentals of Differential Geometry*, Graduate Texts in Mathematics, Springer-Verlag, 1999.

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

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

[15] J. Lawson and Y. Lim, Karcher means and Karcher equations of positive operators, Trans. Amer. Math. Soc. Series B **1** (2014), 1–22.

[16] J. Lawson and Y. Lim, Following the trail of the operator geometric mean, Proceedings of the XXXVIII Workshop on Geometric Methods in Physics 2019, to appear.

[17] Y. Lim and M. Pálfia, The matrix power means and the Karcher mean, J. Funct. Anal. **262** (2012), 1498–1514.

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

[19] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature. Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), 357–390, Contemp. Math., 338, Amer. Math. Soc., Providence, RI, 2003.

Centro de Matematica, FCE-UNLP, La Plata, Argentina and Instituto Argentino de Matematica "Alberto P. Calderon" (IAM-CONICET)

Email address: `eghiglioni@mate.unlp.edu.ar`

Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Korea
Email address: ylim@skku.edu