Robotica (2024), **42**, pp. 4103–4119
doi:10.1017/S0263574724000778

RESEARCH ARTICLE

Trace inequalities and kinematic metrics

Yuwei Wu[1] and Gregory S. Chirikjian[1,2]

[1]Department of Mechanical Engineering, National University of Singapore, Singapore, Singapore
[2]University of Delaware, Newark, Delaware, United States
Corresponding author: Gregory S Chirikjian; Email: mpegre@nus.edu.sg

Received: 27 November 2023; **Revised:** 19 March 2024; **Accepted:** 5 April 2024; **First published online:** 12 September 2024

Keywords: triangle inequality; Golden-Thompson inequality

Abstract
Kinematics remains one of the cornerstones of robotics, and over the decade, Robotica has been one of the venues in which groundbreaking work in kinematics has always been welcome. A number of works in the kinematics community have addressed metrics for rigid-body motions in multiple different venues. An essential feature of any distance metric is the triangle inequality. Here, relationships between the triangle inequality for kinematic metrics and so-called trace inequalities are established. In particular, we show that the Golden-Thompson inequality (a particular trace inequality from the field of statistical mechanics) which holds for Hermitian matrices remarkably also holds for restricted classes of real skew-symmetric matrices. We then show that this is related to the triangle inequality for $SO(3)$ and $SO(4)$ metrics.

1. Introduction

In kinematics, it is natural to ask how large a rigid-body motion is and how to choose a meaningful weighting for the rotational and translational portions of the motion. For example, given the $(n+1) \times (n+1)$ homogeneous transformation matrix

$$H = \begin{pmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \tag{1}$$

that describes a rigid-body displacement in $\mathbb{R}^n$, how far is it from the $(n+1) \times (n+1)$ identity matrix $\mathbb{I}_{n+1}$ (which is the homogeneous transformation describing the null motion)? Having a kinematic distance metric $d(\cdot,\cdot)$ allows one to give a numerical answer: $d(H, \mathbb{I}_{n+1})$.

Then, for example, the problem of serial manipulator inverse kinematics which is usually stated as solving the homogeneous transformation equation

$$H_d = H_1(\theta_1)H_2(\theta_2)\cdots H_n(\theta_n)$$

for $\{\theta_i\}$ instead becomes a problem of minimizing the cost

$$C_0(\{\theta_i\}) \doteq d(H_d\,,\, H_1(\theta_1)H_2(\theta_2)\cdots H_n(\theta_n))\,.$$

Such reformulations of inverse kinematics can be particularly useful for binary-actuated systems where resolved rate methods can be difficult to apply given the discontinuous nature of binary actuators [1].

Another class of examples where metrics can be employed is in problems in sensor calibration such as solving $A_iX = YB_i$ for X and Y [2] and solving $A_iXB_i = YC_iZ$ for X, Y, Z [3] given sets of homogeneous transformations $\{A_i\}$, $\{B_i\}$, and $\{C_i\}$. Using metrics, these become problems of minimizing the cost functions

$$C_1(X,Y) = \sum_i d(A_i X,\ YB_i)$$

and

$$C_2(X,Y,Z) = \sum_i d(A_i X B_i,\ YC_i Z)\,.$$

Sometimes the sum of distances is replaced with sum of squares, to remove square roots from computations.

A number of metrics (or distance functions) have been proposed in the kinematics literature to address the sorts of problems described above. Whereas every metric must, by definition, be symmetric and satisfy the triangle inequality, additional invariance properties are also useful [4–6]. For a recent summary, see [7].

A seemingly unrelated body of literature in the field of statistical mechanics is concerned with the inequality

$$\text{trace}\,(\exp\,(A+B)) \le \text{trace}\,(\exp A \exp B)\,, \tag{2}$$

where $\exp(\cdot)$ is the matrix exponential and A and B are Hermitian matrices of any dimension. This is the *Golden-Thompson inequality* which was proved in 1965 independently in refs. [8] and [9]. In this article, we prove that the inequality (2) also holds when A and B are 3×3 or 4×4 skew-symmetric matrices of bounded norm. Though it has been stated in the literature that (2) extends to the case when A and B are Lie-algebra basis elements, with attribution often given to Kostant [10], in fact, it is not true unless certain caveats are observed, as will be discussed in Section 3.

2. Related work

2.1. SO(3) distance metrics and Euler's theorem

2.1.1. Upper bound from trace inequality

As will be shown in Section 3, (2) holds for skew-symmetric matrices with some caveats. This is relevant to the topic of $SO(3)$ matrices. It is well known that by Euler's theorem, every 3×3 rotation matrix can be written as

$$R = \exp\,(\theta\hat{\mathbf{n}})$$

where $\mathbf{n}$ is the unit vector in the direction of the rotation axis, $\hat{\mathbf{n}}$ is the unique skew-symmetric matrix such that

$$\hat{\mathbf{n}}\,\mathbf{v} = \mathbf{n}\times\mathbf{v}$$

for any $\mathbf{v}\in\mathbb{R}^3$, $\times$ is the cross product, and θ is the angle of the rotation. Letting $\mathbf{n}$ roam the whole sphere and restricting $\theta\in[0,\pi]$ covers all rotations, with redundancy at a set of measure zero. Since

$$\text{trace}(R) = 1 + 2\cos\theta,$$

from this equation, θ can be extracted from R as

$$\theta(R) = \cos^{-1}\left(\frac{\text{trace}(R)-1}{2}\right).$$

It can be shown that given two rotation matrices, then a valid distance metric is [11]

$$d(R_1,R_2) = \theta(R_1^T R_2)\,.$$

It is not difficult to show that this satisfies symmetry and positive definiteness. However, proving the triangle inequality is more challenging. But if the Golden-Thompson inequality (2) can be extended to the case of skew-symmetric matrices, it would provide a proof of the triangle inequality of the above $\theta(R_1^T R_2)$ distance metric. In order to see this, assume that

$$R_1^T R_2 = e^{\theta_1\hat{\mathbf{n}}_1} \text{ and } R_2^T R_3 = e^{\theta_2\hat{\mathbf{n}}_2},$$

with $\theta_1 \in [0, \pi]$ and $\theta_2 \in [0, \pi]$. It is not difficult to see that

$$\theta_1 + \theta_2 \ \geq \ \|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\|$$

since

$$\|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\|^2 = \theta_1^2 + \theta_2^2 + 2\theta_1\theta_2 \mathbf{n}_1 \cdot \mathbf{n}_2$$

and $\mathbf{n}_1 \cdot \mathbf{n}_2 \in [-1, 1]$. On one hand, if $\|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| \leq \pi$ and (2) does hold for skew-symmetric matrices, then computing

$$f_1 \doteq \text{trace}\left(e^{\theta_1 \hat{\mathbf{n}}_1 + \theta_2 \hat{\mathbf{n}}_2}\right) = 1 + 2\cos \|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\|$$

and

$$f_2 \doteq \text{trace}\left(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2}\right) = 1 + 2\cos\theta\left(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2}\right)$$

and observing (2) would give

$$f_1 \leq f_2 \,.$$

But the function $f(\theta) = 1 + 2\cos\theta$ is monotonically nonincreasing when $\theta \in [0, \pi]$, so $f(\theta) \leq f(\phi)$ implies $\theta \geq \phi$. Therefore, if (2) holds, then

$$\|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| \ \geq \ \theta(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2})\,.$$

Then

$$d(R_1, R_2) + d(R_2, R_3) = \theta_1 + \theta_2 \geq \|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| \geq \theta(R_1^T R_2 R_2^T R_3) = \theta(R_1^T R_3) = d(R_1, R_3).$$

On the other hand, if $\|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| > \pi$, then

$$d(R_1, R_2) + d(R_2, R_3) = \theta_1 + \theta_2 \geq \|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| > \pi \geq \theta(R_1^T R_3) = d(R_1, R_3).$$

Therefore, if the Golden-Thompson inequality can be generalized to the case of skew-symmetric matrices, the result will be stronger than the triangle inequality for the $SO(3)$ metric $\theta(R_1^T R_2)$ since the latter follows from the former.

2.1.2. Lower bound from quaternion sphere

Alternatively, unit quaternions provide a simple way to encode the axis-angle representation, that is,

$$\mathbf{q}(\mathbf{n}, \theta) = \left(\mathbf{n}\sin\frac{\theta}{2}, \cos\frac{\theta}{2}\right), \ \text{ where } \|\mathbf{q}(\mathbf{n}, \theta)\| = 1,$$

and we have the quaternion composition formula [12]

$$\cos\left(\frac{\theta(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2})}{2}\right) = \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} - \mathbf{n}_1 \cdot \mathbf{n}_2 \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}. \tag{3}$$

We can show that $\theta(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2})$ is bounded from below such that

$$\theta(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2}) \geq 2\|\mathbf{q}(\mathbf{n}_1, \theta_1) - \mathbf{q}(-\mathbf{n}_2, \theta_2)\|,$$

provided that

$$\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} - \mathbf{n}_1 \cdot \mathbf{n}_2 \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2} \geq 0.$$

To see this, let

$$W = \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} - \mathbf{n}_1 \cdot \mathbf{n}_2 \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2} \in [-1, 1],$$

and

$$\mathcal{Q} = 2\|\mathbf{q}(\mathbf{n}_1, \theta_1) - \mathbf{q}(-\mathbf{n}_2, \theta_2)\| = 2\sqrt{2 - 2\left(\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} - \mathbf{n}_1 \cdot \mathbf{n}_2 \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\right)} = 2\sqrt{2 - 2W}.$$

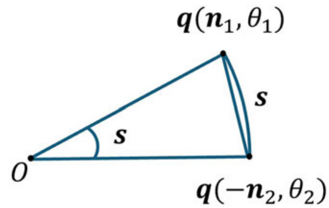

Figure 1. *Geometric interpretation of the lower bound inequality, where s is the arc length between $\boldsymbol{q}(\boldsymbol{n}_1, \theta_1)$ and $\boldsymbol{q}(-\boldsymbol{n}_2, \theta_2)$ on the quaternion sphere, as well as the angle between $\boldsymbol{Oq}(\boldsymbol{n}_1, \theta_1)$ and $\boldsymbol{Oq}(-\boldsymbol{n}_2, \theta_2)$.*

Let

$$f(h) = h^2 + 2\cos h - 2, h \in [0, +\infty).$$

It is easy to compute the derivative

$$f^{'}(h) = 2h - 2\sin h \geq 0.$$

Thus,

$$f(h_*) = h_*^2 + 2\cos h_* - 2 \geq f(0) = 0,$$

that is, $\cos h_* \geq \frac{2-h_*^2}{2}$ for any $h_* \geq 0$. Substituting h_* with $\sqrt{2-2W} \in [0, 2]$ gives

$$\cos\sqrt{2-2W} \geq W.$$

Let $\beta \in [0, \pi]$ such that $\cos\beta = W$, we have $\sqrt{2-2W} \leq \beta$, that is, $\mathcal{Q} = 2\sqrt{2-2W} \leq 2\beta$. If $W \in [0, 1]$, then $\beta \in [0, \frac{\pi}{2}]$ and $2\beta \in [0, \pi]$. So by (3),

$$\theta(e^{\theta_1\hat{\mathbf{n}}_1}e^{\theta_2\hat{\mathbf{n}}_2}) = 2\beta \geq 2\|\mathbf{q}(\mathbf{n}_1, \theta_1) - \mathbf{q}(-\mathbf{n}_2, \theta_2)\|.$$

On the other hand, if $W \in [-1, 0)$, then $\beta \in [\frac{\pi}{2}, \pi]$ and $2\beta \in [\pi, 2\pi]$. According to our definition of distance metric,

$$\theta(e^{\theta_1\hat{\mathbf{n}}_1}e^{\theta_2\hat{\mathbf{n}}_2}) = 2\pi - 2\beta,$$

which does not guarantee to be larger or equal than $2\|\mathbf{q}(\mathbf{n}_1, \theta_1) - \mathbf{q}(-\mathbf{n}_2, \theta_2)\|$. Geometrically speaking, $\theta(e^{\theta_1\hat{\mathbf{n}}_1}e^{\theta_2\hat{\mathbf{n}}_2})$ can be regarded as distance between two rotations $e^{\theta_1\hat{\mathbf{n}}_1}$ and $e^{-\theta_2\hat{\mathbf{n}}_2}$, which equals to the arc length between $\mathbf{q}(\mathbf{n}_1, \theta_1)$ and $\mathbf{q}(-\mathbf{n}_2, \theta_2)$ of the quaternion sphere. Furthermore, the arc length s is just the radian angle between $\mathbf{q}(\mathbf{n}_1, \theta_1)$ and $\mathbf{q}(-\mathbf{n}_2, \theta_2)$, that is,

$$\cos s = \mathbf{q}(\mathbf{n}_1, \theta_1) \cdot \mathbf{q}(-\mathbf{n}_2, \theta_2) = \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} - \mathbf{n}_1 \cdot \mathbf{n}_2 \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}.$$

But the arc length will always be larger than or equal to the Euclidean distance between $\mathbf{q}(\mathbf{n}_1, \theta_1)$ and $\mathbf{q}(-\mathbf{n}_2, \theta_2)$, that is,

$$s \geq \|\mathbf{q}(\mathbf{n}_1, \theta_1) - \mathbf{q}(-\mathbf{n}_2, \theta_2)\|,$$

which is equivalent to the lower bound discussed above (Fig. 1).

2.2. SO(4) distance metrics as an approximation for SE(3) using stereographic projection

It has been known in kinematics for decades that rigid-body motions in $\mathbb{R}^n$ can be approximated as rotations in $\mathbb{R}^{n+1}$ by identifying Euclidean space locally as the tangent plane to a sphere [13]. This has been used to generate approximately bi-invariant metrics for $SE(3)$ [14]. Related to this are approaches that use the singular value decomposition [15]. As with the $SO(3)$ case discussed above, if the Golden-Thompson inequality can be shown to hold for 4×4 skew-symmetric matrices, then a sharper version of the triangle inequality would hold for $SO(4)$ metrics.

 https://doi.org/10.1017/S0263574724000778

This is the subject of Section 3, which is the main contribution of this paper. In that section, it is shown that the Golden-Thompson inequality can be extended from Hermitian matrices to 4×4 skew-symmetric matrices and therefore to the 3×3 case as a special case. But before progressing to the main topic, some trace inequalities that arise naturally from other kinematic metrics are discussed. For example, the distance metric

$$d(R_1, R_2) \doteq \|R_1 - R_2\|_F$$

is a valid metric where the Frobenius norm of an arbitrary real matrix is

$$\|A\|_F \doteq \sqrt{\text{trace}(AA^T)}\,.$$

The triangle inequality for matrix norms then gives

$$\|R_1 - R_2\|_F + \|R_2 - R_3\|_F \geq \|R_1 - R_3\|_F\,.$$

Since the trace is invariant under similarity transformations, the above is equivalent to

$$\|\mathbb{I} - R_1^T R_2\|_F + \|\mathbb{I} - R_2^T R_3\|_F \geq \|\mathbb{I} - R_1^T R_3\|_F\,.$$

This is true in any dimension. But in the 3D case, we can go further using the same notation as in the previous section to get

$$\sqrt{3 - \text{trace}(e^{\theta_1 \hat{\mathbf{n}}_1})} + \sqrt{3 - \text{trace}(e^{\theta_2 \hat{\mathbf{n}}_2})} \geq \sqrt{3 - \text{trace}(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2})}\,. \tag{4}$$

This trace inequality is equivalently written as

$$\sqrt{1 - \cos\theta_1} + \sqrt{1 - \cos\theta_2} \geq \sqrt{1 - \cos\theta(e^{\theta_1 \hat{\mathbf{n}}_1} e^{\theta_2 \hat{\mathbf{n}}_2})}\,.$$

2.3. SE(3) metrics as matrix norms and resulting trace inequalities

Multiple metrics for $SE(3)$ have been proposed over the past decades, as summarized recently in ref. [7]. The purpose of this section is to review in more detail a specific metric that has been studied in refs. [16–18]. The concept of this metric for $SE(3)$ is to induce from the metric properties of the vector 2-norm

$$\|\mathbf{x}\|_2 \doteq \sqrt{\mathbf{x}^T \mathbf{x}}$$

in $\mathbb{R}^3$. Since Euclidean distance is by definition invariant to Euclidean transformations, given the pair $g = (R, \mathbf{t})$, which contains the same information as a homogeneous transformation, and given the group action $g \cdot \mathbf{x} \doteq R\mathbf{x} + \mathbf{t}$, then

$$\|g \cdot \mathbf{x} - g \cdot \mathbf{y}\|_2 = \|\mathbf{x} - \mathbf{y}\|_2\,.$$

Then, if a body with mass density $\rho(\mathbf{x})$ is moved from its original position and orientation, the total amount of motion can be quantified as

$$d(g, e) \doteq \sqrt{\int_{\mathbb{R}^3} \|g \cdot \mathbf{x} - \mathbf{x}\|_2^2\, \rho(\mathbf{x})\, d\mathbf{x}}\,. \tag{5}$$

This metric has the left-invariance property

$$d(h \circ g_1, h \circ g_2) = d(g_1, g_2)$$

where $h, g_1, g_2 \in SE(3)$. This is because if $h = (R, \mathbf{t}) \in SE(3)$, then

$$(h \circ g_i) \cdot \mathbf{x} = h \cdot (g_i \cdot \mathbf{x})$$

and

$$\|h\cdot(g_1\cdot\mathbf{x})-h\cdot(g_2\cdot\mathbf{x})\|_2=$$
$$\|R[g_1\cdot\mathbf{x}]+\mathbf{t}-R[g_2\cdot\mathbf{x}]-\mathbf{t}\|_2$$
$$=\|g_1\cdot\mathbf{x}-g_2\cdot\mathbf{x}\|_2.$$

It is also interesting to note that there is a relationship between this kind of metric for $SE(3)$ and the Frobenius matrix norm. That is, for $g=(R,\mathbf{t})$, and the corresponding homogeneous transformation $H(g)\in SE(3)$, the integral in (5) can be computed in closed form, resulting in a weighted norm

$$d(g,e)=\|H(g)-\mathbb{I}_4\|_W$$

where the weighted Frobenius norm is defined as

$$\|A\|_W \doteq \sqrt{\operatorname{tr}(A^TWA)}\,.$$

Here, with weighting matrix $W=W^T\in\mathbb{R}^{4\times 4}$ is $W=\begin{pmatrix} J & \mathbf{0}\\ \mathbf{0}^T & M\end{pmatrix}$. $M=\int_{\mathbb{R}^3}\rho(\mathbf{x})\,d\mathbf{x}$ is the mass, and $J=\int_{\mathbb{R}^3}\mathbf{x}\mathbf{x}^T\rho(\mathbf{x})\,d\mathbf{x}$ has a simple relationship with the moment of inertia matrix of the rigid body:

$$I=\int_{\mathbb{R}^3}\left((\mathbf{x}^T\mathbf{x})\mathbb{I}_3-\mathbf{x}\mathbf{x}^T\right)\rho(\mathbf{x})\,d\mathbf{x}=\operatorname{tr}(J)\mathbb{I}_3-J.$$

The metric in (5) can also be written as

$$d(g,e)=\sqrt{2\operatorname{tr}[(\mathbb{I}_3-R)J]+\mathbf{t}\cdot\mathbf{t}M}\,. \tag{6}$$

Furthermore, for $g_1,g_2\in SE(3)$

$$d(g_1,g_2)=\|H(g_1)-H(g_2)\|_W\,, \tag{7}$$

as explained in detail in ref. [5]. When evaluating the triangle inequality for this metric,

$$d(g_1,g_2)+d(g_2,g_3)\ \geq\ d(g_1,g_3)$$

gives another kind of trace inequality.

3. Extension of the Golden-Thompson inequality to SO(3) and SO(4)

Motivated by the arguments presented in earlier sections, in this section, the Golden-Thompson inequality is extended to $SO(3)$ and $SO(4)$. It is well known that the eigenvalues of a 4×4 skew-symmetric matrix are $\{\pm\psi_1 i,\pm\psi_2 i\}$ and eigenvalues of a 3×3 skew-symmetric matrix are $\{\pm\psi i,0\}$. In the following contents, we will prove that

$$\operatorname{trace}\left(\exp\left(A+B\right)\right)\ \leq\ \operatorname{trace}\left(\exp A\exp B\right),$$

for A and B being 4×4 skew-symmetric matrices provided that $|\psi_1|+|\psi_2|\leq\pi$, where $\{\pm\psi_1 i,\pm\psi_2 i\}$ are eigenvalues of $A+B$, or for A and B being 3×3 skew-symmetric matrices provided that $|\psi|\leq\pi$, where $\pm\psi i$ are eigenvalues of $A+B$.

3.1. 4D case

Let A be a 4×4 skew-symmetric matrix with its eigenvalues being $\{\pm\theta_1 i,\pm\theta_2 i\}$. Without loss of generality, we assume that $\theta_1\geq\theta_2\geq 0$. For every A, there exists an orthogonal matrix P such that [19]:

$$\Omega_A=P^{\intercal}AP,$$

where

$$\Omega_A = \begin{bmatrix} 0 & \theta_1 & 0 & 0 \\ -\theta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_2 \\ 0 & 0 & -\theta_2 & 0 \end{bmatrix}.$$

Let $\Omega_A = \theta_1 \Omega_1 + \theta_2 \Omega_2$, where

$$\Omega_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \Omega_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

Then

$$A = P\Omega_A P^\intercal = \sum_{i=1}^{2} \theta_i P \Omega_i P^\intercal = \sum_{i=1}^{2} \theta_i A_i,$$

where $A_i = P\Omega_i P^\intercal$. Notice that $\Omega_j^3 + \Omega_j = 0$ and $\Omega_1 \times \Omega_2 = 0 = \Omega_2 \times \Omega_1$. So,

$$A_j^3 + A_j = (P\Omega_i P^\intercal)^3 + P\Omega_i P^\intercal = P\left(\Omega_j^3 + \Omega_j\right) P^\intercal = 0,$$

and

$$A_1 \times A_2 = P\Omega_1 P^\intercal \times P\Omega_2 P^\intercal = 0 = P\Omega_2 P^\intercal \times P\Omega_1 P^\intercal = A_2 \times A_1.$$

In other words, A_1 and A_2 commute. Thus, we can expand the exponential of A as follows:

$$\exp A = \exp\left(\sum_{i=1}^{2} \theta_i A_i\right) = \prod_{i=1}^{2} \exp\left(\theta_i A_i\right) = \prod_{i=1}^{2} \left(I + \sin\theta_i A_i + (1 - \cos\theta_i) A_i^2\right). \tag{8}$$

The last equality comes from the fact that $A_j^3 + A_j = 0$. Expanding the above equation gives

$$\exp A = I + \sum_{i=1}^{2} \left(\sin\theta_i A_i + (1 - \cos\theta_i) A_i^2\right). \tag{9}$$

Given another 4×4 skew-symmetric matrix B whose eigenvalues are $\{\pm\phi_1 i, \pm\phi_2 i\}$ and $\phi_1 \geq \phi_2 \geq 0$, we have

$$\begin{aligned} \text{trace}\left(\exp A \exp B\right) &= \text{trace}\left(P^\intercal \exp(A) \exp(B) P\right) = \text{trace}\left(P^\intercal \exp(A) P P^\intercal \exp(B) P\right) \\ &= \text{trace}\left(\exp(P^\intercal A P) \exp(P^\intercal B P)\right) = \text{trace}\left(\exp \Omega_A \exp C\right), \end{aligned} \tag{10}$$

where $C = P^\intercal B P$. Notice

$$C^\intercal = P^\intercal B^\intercal P = -P^\intercal B P = -C,$$

so C is a skew-symmetric matrix as well, and C has exactly the same eigenvalues as B since conjugation does not change the eigenvalues of a matrix. A similar conclusion can be drawn:

$$C = Q\Omega_C Q^\intercal = \sum_{i=1}^{2} \phi_i Q \Omega_i Q^\intercal = \sum_{i=1}^{2} \phi_i C_i \text{ and } QQ^\intercal = \mathbb{I}_4.$$

Expanding (10) via (9) gives

$$\begin{aligned}
&\text{trace}\,(\exp A \exp B) = \text{trace}\,(\exp \Omega_A \exp C) = \\
&= \text{trace}\left(\left(I + \sum_{i=1}^{2}\left(\sin\theta_i\Omega_i + (1-\cos\theta_i)\Omega_i^2\right)\right)\left(I + \sum_{i=1}^{2}\left(\sin\phi_i C_i + (1-\cos\phi_i)C_i^2\right)\right)\right) \\
&= \text{trace}\left(I + \sum_{i=1}^{2}\left(\sin\theta_i\Omega_i + (1-\cos\theta_i)\Omega_i^2\right) + \sum_{i=1}^{2}\left(\sin\phi_i C_i + (1-\cos\phi_i)C_i^2\right)\right) \\
&+ \text{trace}\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\left(\sin\theta_i\Omega_i + (1-\cos\theta_i)\Omega_i^2\right)\left(\sin\phi_j C_j + (1-\cos\phi_j)C_j^2\right)\right). \qquad (11)
\end{aligned}$$

Divide Q into 2×2 blocks as follows:

$$Q = \left[\begin{array}{cc|cc} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ \hline q_{31} & q_{32} & q_{33} & q_{34} \\ q_{41} & q_{42} & q_{43} & q_{44} \end{array}\right] = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \qquad (12)$$

where

$$Q_{ij} = \begin{bmatrix} q_{2i-1,2j-1} & q_{2i-1,2j} \\ q_{2i,2j-1} & q_{2i,2j} \end{bmatrix}.$$

Denoting $\omega_{ij} = \det Q_{ij}$ and $\varepsilon_{ij} = \|Q_{ij}\|_F^2 = q_{2i-1,2j-1}^2 + q_{2i-1,2j}^2 + q_{2i,2j-1}^2 + q_{2i,2j}^2$, we have the following equities:

$$\begin{aligned}
&\text{trace}\left(\Omega_i C_j\right) = \text{trace}\,(\Omega_i Q\Omega_i Q^\intercal) = -2\omega_{ij} \\
&\text{trace}\left(\Omega_i^2 C_j^2\right) = \text{trace}\left(\Omega_i^2 Q\Omega_i^2 Q^\intercal\right) = \varepsilon_{ij} \\
&\text{trace}(C_i) = \text{trace}(Q\Omega_i Q^\intercal) = \text{trace}(\Omega_i) = 0 \\
&\text{trace}(C_i^2) = \text{trace}(Q\Omega_i^2 Q^\intercal) = \text{trace}(\Omega_i^2) = -2 \\
&\text{trace}\left(\Omega_i^2 C_j\right) = \text{trace}\left(\Omega_i C_j^2\right) = 0. \qquad (13)
\end{aligned}$$

Combining (11) with (13) gives

$$\begin{aligned}
\text{trace}\,(\exp A \exp B) &= 4 + 2\sum_{i=1}^{2}(\cos\theta_i - 1) + 2\sum_{i=1}^{2}(\cos\phi_i - 1) \\
&+ \sum_{i=1}^{2}\sum_{j=1}^{2}\left(-2\sin\theta_i\sin\phi_j\omega_{ij} + (1-\cos\theta_i)(1-\cos\phi_j)\varepsilon_{ij}\right). \qquad (14)
\end{aligned}$$

Using the fact $\varepsilon_{11} + \varepsilon_{12} = \varepsilon_{11} + \varepsilon_{21} = \varepsilon_{22} + \varepsilon_{12} = \varepsilon_{22} + \varepsilon_{21} = 2$ as Q is an orthogonal matrix, we can reduce (14) to

$$\text{trace}\,(\exp A \exp B) = \sum_{i=1}^{2}\sum_{j=1}^{2}\left(\cos\theta_i\cos\phi_j\varepsilon_{ij} - 2\sin\theta_i\sin\phi_j\omega_{ij}\right). \qquad (15)$$

For convenience, in the following, we will use L_1 to denote trace $(\exp A \exp B)$ and L_2 to denote trace $(\exp(A+B))$.

Lemma 3.1. Let $\omega_{11} = \det(Q_{\mathbf{11}}), \omega_{12} = \det(Q_{\mathbf{12}}),$ and $\varepsilon_{11} = \|Q_{11}\|_F^2$, where $Q_{\mathbf{11}}$ and $Q_{\mathbf{12}}$ are defined as in (12), then the following equity holds

$$\varepsilon_{11} = 1 + \omega_{11}^2 - \omega_{12}^2.$$

Proof. Since $q_{11}q_{21}+q_{12}q_{22}+q_{13}q_{23}+q_{14}q_{24}=0$ (by orthogonality), we have

$$\begin{aligned}(q_{11}q_{21}+q_{12}q_{22})^2 &= q_{11}^2q_{21}^2+q_{12}^2q_{22}^2+2q_{11}q_{12}q_{21}q_{22}\\ =(q_{13}q_{23}+q_{14}q_{24})^2 &= q_{13}^2q_{23}^2+q_{14}^2q_{24}^2+2q_{13}q_{14}q_{23}q_{24},\end{aligned}$$

that is,

$$q_{11}^2q_{21}^2+q_{12}^2q_{22}^2-q_{13}^2q_{23}^2-q_{14}^2q_{24}^2=-2q_{11}q_{12}q_{21}q_{22}+2q_{13}q_{14}q_{23}q_{24}.$$

Then,

$$\begin{aligned}RHS &= 1+\omega_{11}^2-\omega_{12}^2=1+(q_{11}q_{22}-q_{12}q_{21})^2-(q_{13}q_{24}-q_{14}q_{23})^2\\ &= 1+q_{11}^2q_{22}^2+q_{12}^2q_{21}^2-q_{13}^2q_{24}^2-q_{14}^2q_{23}^2-2q_{11}q_{12}q_{21}q_{22}+2q_{13}q_{14}q_{23}q_{24}\\ &= 1+q_{11}^2q_{22}^2+q_{12}^2q_{21}^2-q_{13}^2q_{24}^2-q_{14}^2q_{23}^2+q_{11}^2q_{21}^2+q_{12}^2q_{22}^2-q_{13}^2q_{23}^2-q_{14}^2q_{24}^2\\ &= 1+(q_{11}^2+q_{12}^2)(q_{21}^2+q_{22}^2)-(q_{13}^2+q_{14}^2)(q_{23}^2+q_{24}^2)\\ &= 1+(q_{11}^2+q_{12}^2)(q_{21}^2+q_{22}^2)-(1-q_{11}^2-q_{12}^2)(1-q_{21}^2-q_{22}^2)\\ &= q_{11}^2+q_{12}^2+q_{21}^2+q_{22}^2=\varepsilon_{11}=LHS,\end{aligned}$$

that is, $\varepsilon_{11}=1+\omega_{11}^2-\omega_{12}^2$. □

Lemma 3.2. Let ω_{11} and ω_{12} be the determinants of $\mathbf{Q_{11}}$ and $\mathbf{Q_{12}}$, where $\mathbf{Q_{11}}$ and $\mathbf{Q_{12}}$ are as defined in (12), then $|\omega_{11}+\omega_{12}|\le 1$ and $|\omega_{11}-\omega_{12}|\le 1$.

Proof. Since $\omega_{11}=q_{11}q_{22}-q_{12}q_{21}$, $\omega_{12}=q_{13}q_{24}-q_{14}q_{23}$, $q_{11}^2+q_{12}^2+q_{13}^2+q_{14}^2=1$, and $q_{21}^2+q_{22}^2+q_{23}^2+q_{24}^2=1$, we have

$$\begin{aligned}&(q_{11}^2+q_{12}^2+q_{13}^2+q_{14}^2)(q_{21}^2+q_{22}^2+q_{23}^2+q_{24}^2)-(\omega_{11}+\omega_{12})^2\\ &=(q_{11}q_{21}+q_{12}q_{22})^2+(q_{11}q_{23}+q_{14}q_{22})^2+(q_{11}q_{24}-q_{13}q_{22})^2\\ &+(q_{12}q_{23}-q_{14}q_{21})^2+(q_{12}q_{24}+q_{13}q_{21})^2+(q_{13}q_{23}+q_{14}q_{24})^2\ge 0,\end{aligned}$$

that is, $(\omega_{11}+\omega_{12})^2\le 1$. The same deduction gives $(\omega_{11}-\omega_{12})^2\le 1$. □

Let

$$m_1=\omega_{11}+\omega_{12} \text{ and } m_2=\omega_{11}-\omega_{12}.$$

Instantly by Lemma 3.2, we have $|m_1|\le 1$ and $|m_2|\le 1$. Recall that $\varepsilon_{11}+\varepsilon_{12}=2=\varepsilon_{12}+\varepsilon_{22}$, so $\varepsilon_{11}=\varepsilon_{22}$ and similarly $\varepsilon_{12}=\varepsilon_{21}$. By Lemme 3.1, we have

$$\varepsilon_{11}=\varepsilon_{22}=1+m_1m_2$$

and

$$\varepsilon_{12}=\varepsilon_{21}=2-\varepsilon_{11}=1-m_1m_2.$$

Since Q is an orthogonal matrix, $\det Q=\pm 1$ which is denoted as μ. In ref. [20], the author has shown that

$$\det Q_{11}=\det Q_{22}\det Q$$

and

$$\det Q_{12}=\det Q_{21}\det Q$$

if Q is an orthogonal matrix. Therefore, we have

$$\omega_{22}=\mu\omega_{11}=\frac{\mu(m1+m2)}{2}, \text{ and } \omega_{21}=\mu\omega_{12}=\frac{\mu(m1-m2)}{2}.$$

Substituting ω_{ij} and ε_{ij} with m_1 and m_2 into (15) gives

$$\begin{aligned} L_1 = &(1+m_1m_2)\cos\theta_1\cos\phi_1 + (1-m_1m_2)\cos\theta_1\cos\phi_2 \\ &+ (1-m_1m_2)\cos\theta_2\cos\phi_1 + (1+m_1m_2)\cos\theta_2\cos\phi_2 \\ &- (m_1+m_2)\sin\theta_1\sin\phi_1 - (m_1-m_2)\sin\theta_1\sin\phi_2 \\ &- \mu(m_1-m_2)\sin\theta_2\sin\phi_1 - \mu(m_1+m_2)\sin\theta_2\sin\phi_2. \end{aligned} \tag{16}$$

On the other hand,

$$L_2 = \text{trace}\,(\exp(A+B)) = \text{trace}\,(P^\intercal \exp(A+B)P) = \text{trace}\,(\exp(\Omega_A + C))\,. \tag{17}$$

Let $D = \Omega_A + C = \sum_{i=1}^{2}(\theta_i\Omega_i + \phi_i Q\Omega_i Q^\intercal)$. The characteristic polynomial $\mathcal{P}(\lambda)$ of D is

$$\mathcal{P}(\lambda) = \lambda^4 + \mathcal{P}_1\lambda^2 + \mathcal{P}_2,$$

where

$$\mathcal{P}_1 = \theta_1^2 + \theta_2^2 + \phi_1^2 + \phi_2^2 + 2\omega_{11}\theta_1\phi_1 + 2\omega_{12}\theta_1\phi_2 + 2\omega_{21}\theta_2\phi_1 + 2\omega_{22}\theta_2\phi_2,$$

and

$$\mathcal{P}_2 = (\theta_1\theta_2 + \phi_1\phi_2 + \omega_{12}\theta_1\phi_1 + \omega_{11}\theta_1\phi_2 + \omega_{22}\theta_2\phi_1 + \omega_{21}\theta_2\phi_2)^2\,.$$

Using the face that $\omega_{22} = \mu\omega_{11}$ and $\omega_{21} = \mu\omega_{12}$, we can solve the above quartic equation:

$$\lambda_{1,2} = \pm\left(\frac{\sqrt{f_1}+\sqrt{f_2}}{2}\right) i = \pm\psi_1 i,\ \ \lambda_{3,4} = \pm\left(\frac{|\sqrt{f_1}-\sqrt{f_2}|}{2}\right) i = \pm\psi_2 i, \tag{18}$$

where

$$f_1 = (\theta_1 + \mu\theta_2)^2 + (\phi_1+\phi_2)^2 + 2(\theta_1+\mu\theta_2)(\phi_1+\phi_2)(\omega_{11}+\omega_{12}),$$

and

$$f_2 = (\theta_1 - \mu\theta_2)^2 + (\phi_1-\phi_2)^2 + 2(\theta_1-\mu\theta_2)(\phi_1-\phi_2)(\omega_{11}-\omega_{12}).$$

Both f_1 and f_2 are guaranteed to be greater than or equal to 0 since $|\omega_{11}+\omega_{12}| \le 1$, $|\omega_{11}-\omega_{12}| \le 1$ (Lemme 3.2), $\theta_1 \pm \mu\theta_2 \ge 0$, and $\phi_1 \pm \phi_2 \ge 0$. So, expanding (17) by (9) gives

$$\begin{aligned} L_2 &= \text{trace}\left(I + \sin\psi_1 D_1 + (1-\cos\psi_1)D_1^2 + \sin\psi_2 D_2 + (1-\cos\psi_2)D_2^2\right) \\ &= 4 + 2(\cos\psi_1 - 1) + 2(\cos\psi_2 - 1) = 2\cos\psi_1 + 2\cos\psi_2 \\ &= 2\cos\left(\frac{\sqrt{f_1}+\sqrt{f_2}}{2}\right) + 2\cos\left(\frac{\sqrt{f_1}-\sqrt{f_2}}{2}\right) = 4\cos\frac{\sqrt{f_1}}{2}\cos\frac{\sqrt{f_2}}{2}. \end{aligned} \tag{19}$$

Perform the following coordinate transformations:

$$\begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5\mu & 0.5 & 0.5 \\ 0.5 & -0.5\mu & 0.5 & -0.5 \\ -0.5 & -0.5\mu & 0.5 & 0.5 \\ -0.5 & 0.5\mu & 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \phi_1 \\ \phi_2 \end{bmatrix}.$$

Applying the above transformation to (16) and (19) gives

$$L_1 = (\cos x_1 + \cos y_1 + m_1\cos x_1 - m_1\cos y_1)(\cos x_2 + \cos y_2 + m_2\cos x_2 - m_2\cos y_2), \tag{20}$$

and

$$L_2 = 2\cos\frac{K_1}{2}\cdot 2\cos\frac{K_2}{2}, \tag{21}$$

where

$$K_1 = \sqrt{2(1+m_1)x_1^2 + 2(1-m_1)y_1^2},$$

and

$$K_2 = \sqrt{2(1+m_2)x_2^2 + 2(1-m_2)y_2^2}.$$

Lemma 3.3. Let $a = \frac{K}{\sqrt{2(1+\zeta)}}$ and $b = \frac{K}{\sqrt{2(1-\zeta)}}$, where $\zeta \in (-1, 1)$ and $K \in (0, \pi]$. Then

$$(1+\zeta)\cos a + 1 - \zeta > 2\cos\frac{K}{2} \text{ and } (1-\zeta)\cos b + 1 + \zeta > 2\cos\frac{K}{2}.$$

Proof. Let

$$f(p) = -\frac{\sin^2(p)}{p^2}, \text{ where } p \in (0, +\infty),$$

and the derivative of $f(p)$ is

$$\frac{df}{dp} = \frac{2\sin^2(p) - 2\sin(p)\cos(x)p}{p^3}.$$

It is not difficult to conclude that $\frac{df}{dp} > 0$ when $p \in (0, \pi)$; that is, $f(p)$ is strictly increasing as $p \in (0, \pi]$. Assume that there exists a p_0 such that $f(p_0) < f(\frac{\pi}{2})$, then

$$-\frac{\sin^2 p_0}{p_0^2} < -\frac{1}{\frac{\pi^2}{4}}, \; i.e., \; \frac{4p_0^2}{\pi^2} < \sin^2 p_0 \leq 1.$$

So if such p_0 exists, it must satisfy $p_0 < \frac{\pi}{2}$, which means for any $p \geq \frac{\pi}{2}$, we have

$$f(p) \geq f(\frac{\pi}{2}) > f(\frac{\pi}{4}) \geq f(\frac{K}{4}).$$

Let $q = \sqrt{2(1+\zeta)} \in (0, 2)$, then $\frac{1}{q} > \frac{1}{2}$ and $\frac{K}{2q} > \frac{K}{4}$. If $\frac{K}{2q} < \frac{\pi}{2}$, then

$$f\left(\frac{\pi}{2}\right) > f\left(\frac{K}{2q}\right) > f\left(\frac{K}{4}\right)$$

since f is strictly increasing within that range. Otherwise if $\frac{K}{2q} \geq \frac{\pi}{2}$, we have $f\left(\frac{K}{2q}\right) > f(\frac{\pi}{4}) \geq f(\frac{K}{4})$. In other words, the following inequality is always valid:

$$f\left(\frac{K}{2q}\right) = \frac{4q^2(-\sin^2\frac{K}{2q})}{K^2} = \frac{2q^2(\cos\frac{K}{q} - 1)}{K^2} > f\left(\frac{K}{4}\right) = \frac{8(\cos\frac{K}{2} - 1)}{K^2}.$$

Multiplying both sides by $\frac{K^2}{4}$ gives

$$\frac{q^2}{2}(\cos\frac{K}{q} - 1) = (1+\zeta)\left(\cos\frac{K}{\sqrt{2(1+\zeta)}} - 1\right) > 2(\cos\frac{K}{2} - 1),$$

that is,

$$(1+\zeta)\cos a + 1 - \zeta > 2\cos\frac{K}{2}.$$

By letting $\zeta = -\zeta$, we have

$$(1-\zeta)\cos b + 1 + \zeta > 2\cos\frac{K}{2}.$$

□

Lemma 3.4. If $x > 0, y > 0, \zeta \in (-1, 1)$, and $K \in (0, \pi]$, then the only solution to the following equation is $x = y = K/2$

$$\begin{cases} \frac{\sin x}{x} = \frac{\sin y}{y} \\ (1+\zeta)x^2 + (1-\zeta)y^2 = \frac{K^2}{2} \end{cases}.$$

Proof. Let $h(x) = \frac{\sin x}{x}$. Assume that there exists $x_1 > x_2 > 0$ such that $h(x_1) = h(x_2)$. The derivative of $h(x)$ is

$$h'(x) = \frac{x\cos x - \sin x}{x^2}.$$

When $x \in (0, \frac{\pi}{2}]$, $x\cos x - \sin x$ will always be smaller than 0; that is, $h(x)$ is strictly decreasing. Therefore, to have $h(x_1) = h(x_2)$, x_1 must be greater than $\frac{\pi}{2}$. Assume $x_2 \le \frac{\pi}{2}$, then

$$h(x_1) = h(x_2) \ge h(\frac{\pi}{2}) = \frac{2}{\pi}.$$

Thus, we have

$$\frac{2}{\pi} \le h(x_1) \le \frac{\sin x_1}{x_1}, \ i.e.\ \frac{2x_1}{\pi} \le \sin x_1 \le 1,$$

which leads to $x_1 \le \frac{\pi}{2}$, contradicting what we previously stated. Thus, both x_1 and x_2 need to be larger than $\frac{\pi}{2}$. However,

$$\frac{K^2}{2} = (1+\zeta)x_1^2 + (1-\zeta)x_2^2 > (1+\zeta+1-\zeta)\cdot\frac{\pi^2}{4} = \frac{\pi^2}{2},$$

which causes a contradiction since $K \in (0, \pi]$. □

Theorem 3.5. If $(1+\zeta)x^2 + (1-\zeta)y^2 = \frac{K^2}{2}, \zeta \in [-1, 1]$, and $K \in [0, \pi]$, then

$$\cos x + \cos y + \zeta\cos x - \zeta\cos y \ge 2\cos\left(\frac{K}{2}\right) \ge 0.$$

Proof. If $\zeta = 1$, then $x = \pm\frac{K}{2}$. So, $LHS = 2\cos\left(\frac{K}{2}\right) = RHS$. Same for $\zeta = -1$. If $K = 0$, then $x = y = 0$. So, $LHS = 0 = RHS$. Now, let us restrict $\zeta \in (-1, 1)$ and $K \in (0, \pi]$. Let

$$f(x, y) = \cos x + \cos y + \zeta\cos x - \zeta\cos y - 2\cos\left(\frac{K}{2}\right),$$

and

$$g(x, y) = (1+\zeta)x^2 + (1-\zeta)y^2 - \frac{K^2}{2}.$$

To find the minimum of $f(x, y)$ subjected to the equality constraint $g(x, y) = 0$, we form the following Lagrangian function:

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y),$$

where λ is the Lagrange multiplier. Notice that $\mathcal{L}(\pm x, \pm y, \lambda) = f(\pm x, \pm y) + \lambda g(\pm x, \pm y) = f(x, y) + \lambda g(x, y) = \mathcal{L}(x, y, \lambda)$. Thus, the Lagrangian function is symmetric about $x = 0$ and $y = 0$. So, we only need to study how $\mathcal{L}(x, y, \lambda)$ behaves with $(x, y) \in [0, +\infty) \times [0, +\infty)$. To find stationary points of $\mathcal{L}$, we have

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial x} = [-(1+\zeta)\sin x] + 2\lambda(1+\zeta)x = 0 \\ \frac{\partial\mathcal{L}}{\partial y} = [-(1-\zeta)\sin y] + 2\lambda(1-\zeta)y = 0 \\ \frac{\partial\mathcal{L}}{\partial\lambda} = (1+\zeta)x^2 + (1-\zeta)y^2 - \frac{K^2}{2} = 0 \end{cases}.$$

We can readily obtain three sets of solutions to the above equation:

1. $x = 0, y = \sqrt{\frac{K^2}{2(1-\zeta)}}$ and $\lambda = \frac{\sin y}{2y}$;
2. $x = y, x = y = \frac{K}{2}$ and $\lambda = \frac{\sin\frac{K}{2}}{K}$;
3. $y = 0, x = \sqrt{\frac{K^2}{2(1+\zeta)}}$ and $\lambda = \frac{\sin x}{2x}$.

To have a fourth solution, we need to satisfy

$$\frac{(1+\zeta)\sin x}{2(1+\zeta)x} = \lambda = \frac{(1-\zeta)\sin y}{2(1-\zeta)y}$$

and

$$(1+\zeta)x^2 + (1-\zeta)y^2 - \frac{K^2}{2} = 0.$$

However, by Lemma 3.4, we conclude that there are no other solutions. Substituting those solutions back into $f(x, y)$ gives

$$f\left(0, \sqrt{\frac{K^2}{2(1-\zeta)}}\right) = (1-\zeta)\cos\left(\sqrt{\frac{K^2}{2(1-\zeta)}}\right) + 1 + \zeta - 2\cos\left(\frac{K}{2}\right)$$
$$f\left(\frac{K}{2}, \frac{K}{2}\right) = 2\cos\left(\frac{K}{2}\right) - 2\cos\left(\frac{K}{2}\right) = 0$$
$$f\left(\sqrt{\frac{K^2}{2(1+\zeta)}}, 0\right) = (1+\zeta)\cos\left(\sqrt{\frac{K^2}{2(1+\zeta)}}\right) + 1 - \zeta - 2\cos\left(\frac{K}{2}\right).$$

By Lemma 3.3, we have $f\left(0, \sqrt{\frac{K^2}{2(1-\zeta)}}\right) > 0$ and $f\left(\sqrt{\frac{K^2}{2(1+\zeta)}}, 0\right) > 0$. Therefore, we can conclude that the global minimum for $f(x, y)$ subjected to $g(x, y) = 0$ is zero, that is,

$$\cos x + \cos y + \zeta \cos x - \zeta \cos y \geq 2\cos\left(\frac{K}{2}\right).$$

□

Now recall that

$$L_1 = (\cos x_1 + \cos y_1 + m_1 \cos x_1 - m_1 \cos y_1)(\cos x_2 + \cos y_2 + m_2 \cos x_2 - m_2 \cos y_2),$$

and

$$L_2 = 2\cos\frac{K_1}{2} \cdot 2\cos\frac{K_2}{2}.$$

where

$$K_1 = \sqrt{2(1+m_1)x_1^2 + 2(1-m_1)y_1^2},$$

and

$$K_2 = \sqrt{2(1+m_2)x_2^2 + 2(1-m_2)y_2^2}.$$

By (18), we know

$$\psi_1 = \frac{K_1 + K_2}{2} \geq 0 \text{ and } \psi_2 = \frac{|K_1 - K_2|}{2} \geq 0,$$

where $\{\pm\psi_1 i, \pm\psi_2 i\}$ are eigenvalues of $A + B$. With the condition $\psi_1 + \psi_2 \leq \pi$, if $K_1 \geq K_2$, then $\psi_1 + \psi_2 = K_1 \leq \pi$ and so $K_2 \leq K_1 \leq \pi$; otherwise if $K_1 < K_2$, then $\psi_1 + \psi_2 = K_2 \leq \pi$ and $K_1 < K_2 \leq \pi$. In both cases, we have $K_1 \leq \pi$ and $K_2 \leq \pi$. By Theorem 3.5, we have

$$\cos x_1 + \cos y_1 + m_1 \cos x_1 - m_1 \cos y_1 \geq 2\cos\frac{K_1}{2} \geq 0,$$

and

$$\cos x_2 + \cos y_2 + m_2 \cos x_2 - m_2 \cos y_2 \geq 2\cos\frac{K_2}{2} \geq 0.$$

Therefore, we have $L_1 \geq L_2$, that is,

$$\text{trace}\,(\exp A \exp B) \geq \text{tr}\,(\exp(A + B)),$$

subjected to

$$\psi_1 + \psi_2 \leq \pi.$$

3.2. 3D case

Given two 3×3 skew-symmetric matrices A and B such that the eigenvalues of $A + B$ is $\{\pm\psi i, 0\}$ and $\psi \in [0, \pi]$, we can pad both A and B with zeros as follows:

$$\bar{A} = \begin{bmatrix} A & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & 0 \end{bmatrix} \text{ and } \bar{B} = \begin{bmatrix} B & \mathbf{0}_{3\times 1} \\ \mathbf{0}_{1\times 3} & 0 \end{bmatrix}.$$

Then,

$$\text{trace}\left(\exp \bar{A} \exp \bar{B}\right) = \text{trace}\,(\exp A \exp B) + 1,$$

and

$$\text{trace}\left(\exp (\bar{A} + \bar{B})\right) = \text{trace}\,(\exp (A + B)) + 1.$$

Notice that by padding zeros, the eigenvalues of $\bar{A} + \bar{B}$ become $\{\psi i, -\psi i, 0, 0\}$. As $\psi + 0 = \psi \leq \pi$, we have

$$\text{trace}\left(\exp \bar{A} \exp \bar{B}\right) \geq \text{trace}\left(\exp (\bar{A} + \bar{B})\right),$$

that is,

$$\text{trace}\,(\exp A \exp B) \geq \text{trace}\,(\exp (A + B)).$$

4. Applications

In this section, two very different applications of the trace inequality are illustrated.

4.1. BCH formula

The Baker-Campbell-Hausdorff (BCH) formula gives the value of Z that solves the following equation:

$$Z(X, Y) = \log\,(\exp\,(X) \exp\,(Y)) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}\left([X, [X, Y]] - [Y, [X, Y]]\right) + \cdots,$$

where $X, Y,$ and Z are in the Lie algebra of a Lie group, $[X, Y] = XY - YX$, and $\cdots$ indicates terms involving higher commutators of X and Y. The BCH formula is used for robot state estimation [21] and error propagation on the Euclidean motion group [22]. Let us denote all the terms after $X + Y$ as W and so

$$\exp\,(Z) = \exp\,(X) \exp\,(Y)$$
$$Z = \log\,(\exp\,(X) \exp\,(Y)) = X + Y + W.$$

Considering the case of $SO(3)$, we can write

$$X = \theta_1 \hat{\mathbf{n}}_1 \text{ and } Y = \theta_2 \hat{\mathbf{n}}_2,$$

where $\mathbf{n}_i$ is the unit vector in the direction of the rotation axis, $\hat{\mathbf{n}}_i$ is the unique skew-symmetric matrix such that

$$\hat{\mathbf{n}}_i \, \mathbf{v} = \mathbf{n}_i \times \mathbf{v}$$

for any $\mathbf{v} \in \mathbb{R}^3$, and $\boldsymbol{\theta_i} \in [\mathbf{0}, +\infty)$ is the angle of the rotation. Then,

$$d(\exp\,(Z), \mathbb{I}_3) = d(\exp\,(X + Y + W), \mathbb{I}_3) \leq \|\theta_1 \mathbf{n}_1 + \theta_2 \mathbf{n}_2\| = d(\exp\,(X + Y), \mathbb{I}_3),$$

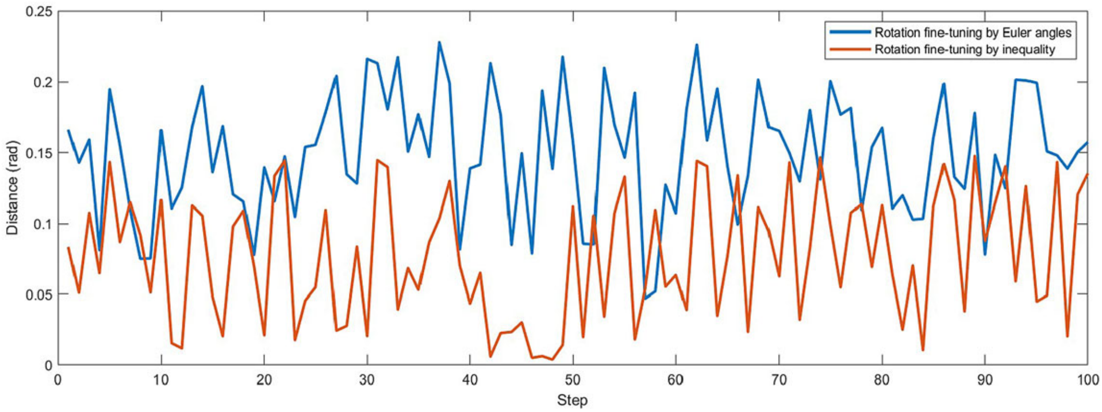

Figure 2. *Average radian distance between the target rotation and the actual rotation.*

provided that $\|\theta_1\mathbf{n}_1 + \theta_2\mathbf{n}_2\| \leq \pi$. So, we conclude that the existence of W will reduce the distance between $\exp(X+Y)$ and the identity $\mathbb{I}_3$ if $\|\theta_1\mathbf{n}_1 + \theta_2\mathbf{n}_2\| \leq \pi$.

4.2. Rotation fine-tuning

Euler angles are a powerful approach to decomposing rotation matrices into three sequential rotation matrices. Let us assume that a manipulator can rotate around the x, y, and z-axis, respectively. Therefore, to rotate the manipulator to a designated orientation R_d, we can compute the corresponding Euler angles α_1, β_1, and γ_1 such that

$$R_x(\alpha_1)R_y(\beta_1)R_z(\gamma_1) = R_d.$$

Assuming that whenever rotated, the device will incur some random noise to the input angle, that is,

$$R_1 = R_x(\alpha_1 + \delta\alpha_1)R_y(\beta_1 + \delta\beta_1)R_z(\gamma_1 + \delta\gamma_1) \neq R_d,$$

leading to deviations of the final orientation. To reduce the error, one can measure the actual rotation R_1 and compute another set of Euler angles $\{\alpha_2, \beta_2, \gamma_2\}$ such that

$$R_x(\alpha_2)R_y(\beta_2)R_z(\gamma_2) = R_d R_1^T.$$

But inevitably, noise will again be introduced, and the actual rotation will become

$$R_2R_1 = R_x(\alpha_2 + \delta\alpha_2)R_y(\beta_2 + \delta\beta_2)R_z(\gamma_2 + \delta\gamma_2)R_1 \neq R_d.$$

Therefore, one can repeat the above process until $d(\prod_{i=1}^{N} R_{N-i+1}, R_d)$ is within tolerance. Another approach to reducing the inaccuracy caused by the noise is applying the following inequality:

$$\|\theta_1\mathbf{n}_1 + \theta_2\mathbf{n}_2\| \geq \theta(e^{\theta_1\hat{\mathbf{n}}_1}e^{\theta_2\hat{\mathbf{n}}_2}).$$

To refine the current rotation by rotating the x-axis, that is, minimizing $d(R_x(\alpha)R_1, R_d) = \theta(R_x(\alpha)R_1R_d^T)$, we let $R_s = R_1R_d^T = \exp(\theta_s\hat{\mathbf{n}}_s)$, where $\theta_s \in [0, \pi]$. If $\alpha = \arg\min_\alpha \|\theta_s\mathbf{n}_s + \alpha\mathbf{e}_1\|$, that is, $\alpha = -(\mathbf{n}_s \cdot \mathbf{e}_1)\theta_s$, then

$$\theta(R_x(\alpha)R_1R_d^T) = \theta(R_x(\alpha)R_s) = \theta(e^{\alpha\hat{\mathbf{e}}_1}e^{\theta_s\hat{\mathbf{n}}_s}) \leq \|\alpha\mathbf{e}_1 + \theta_s\mathbf{n}_s\| = \theta_s\sqrt{1-(\mathbf{n}_s\cdot\mathbf{e}_1)^2} \leq \theta_s.$$

In other words, the inequality provides a simple way to reduce the angle of the resulting rotation by rotating around an axis with a specific angle. In practice, when given the R_s, we compute $|\mathbf{n}_s \cdot \mathbf{e}_1|$, $|\mathbf{n}_s \cdot \mathbf{e}_2|$, and $|\mathbf{n}_s \cdot \mathbf{e}_3|$ and choose the axis that has the largest dot value to rotate. The above process is repeated until the tolerance requirement is met.

To prove the effectiveness, we conduct the following experiment. The target rotation is chosen as $R_d = R_x(\alpha_*)R_y(\beta_*)R_z(\gamma_*)$, where α_*, β_*, and γ_* are all random numbers from $[0, 2\pi]$. We assume that whenever the device is rotated, there will be a noise, which is uniformly distributed within the range $[-0.15, 0.15]$, added to the input angle. In the first step, the manipulator is rotated according to the Euler angles for both methods. Then in the subsequent steps, it is refined three times either by Euler angles or by angles calculated from the inequality. For each approach, we refine the orientation to 100 steps, and at each step, the distance between the current rotation and the target rotation is measured. We conduct the above experiments 500 times, and the average distance is computed at each step for both approaches. The results are shown in Fig. 2. Overall, the radian distance is smaller if we refine the rotation by the inequality. In other words, the inequality provides a simple yet effective way to fine-tune the rotation in the presence of noise.

5. Conclusion

Kinematic metrics, that is, functions that measure the distance between two rigid-body displacements, are important in a number of applications in robotics ranging from inverse kinematics and mechanism design to sensor calibration. The triangle inequality is an essential feature of any distance metric. In this paper, it was shown how trace inequalities from statistical mechanics can be extended to the case of the Lie algebras $so(3)$ and $so(4)$ and how these are in turn related to the triangle inequality for metrics on the Lie groups $SO(3)$ and $SO(4)$. These previously unknown relationships may shed a new light on kinematic metrics for use in robotics.

Author contributions. Dr. Gregory Chirikjian made the conjecture that the trace inequality can be extended to the case of the Lie algebras $so(3)$ and $so(4)$ and proposed several potential applications. Yuwei Wu proved the conjecture. Both authors contributed to writing the article.

Financial support. This work was supported by NUS Startup grants A-0009059-02-00 and A-0009059-03-00, CDE Board Fund E-465-00-0009-01, and AME Programmatic Fund Project MARIO A-0008449-01-00.

Competing interests. The authors declare no conflicts of interest exist.

References

[1] J. Suthakorn and G. S. Chirikjian, "A new inverse kinematics algorithm for binary manipulators with many actuators," *Adv Robotics* **15**(2), 225–244 (2001).
[2] H. Li, Q. Ma, T. Wang and G. S. Chirikjian, "Simultaneous hand-eye and robot-world calibration by solving the $AX = YB$ problem without correspondence," *IEEE Robot Autom Lett* **1**(1), 145–152 (2015).
[3] Q. Ma, Z. Goh, S. Ruan and G. S. Chirikjian, "Probabilistic approaches to the AXB= YCZ AXB= YCZ calibration problem in multi-robot systems," *Auton Robot* **42**(7), 1497–1520 (2018).
[4] N. M. Amato, O. B. Bayazit, L. K. Dale, C. Jones and D. Vallejo, "Choosing Good Distance Metrics and Local Planners for Probabilistic Roadmap Methods," **In:** *Proceedings. 1998 IEEE International Conference on Robotics and Automation*, (IEEE, 1998) pp. 630–637.
[5] G. Chirikjian and S. Zhou, "Metrics on motion and deformation of solid models," *J Mech Design* **120**(2), 252–261 (1998).
[6] G. S. Chirikjian, "Partial bi-invariance of SE(3) metrics," *J Comput Inf Sci Eng* **15**(1), 011008 (2015).
[7] R. Di Gregorio, "Metrics proposed for measuring the distance between two rigid-body poses: Review, comparison, and combination," *Robotica* **42**(1), 302–318 (2024).
[8] S. Golden, "Lower bounds for the helmholtz function," *Phys Rev* **137**(4B), B1127), B1127–B1128 (1965).
[9] C. J. Thompson, "Inequality with applications in statistical mechanics," *J Math Phys* **6**(11), 1812–1813 (1965).
[10] B. Kostant, "On Convexity, the Weyl Group and the Iwasawa Decomposition," **In:** *Annales Scientifiques De l'École Normale Supérieure*, (1973) pp. 413–455.
[11] F. C. Park, "Distance metrics on the rigid-body motions with applications to mechanism design," *J Mech Design* **117**(1), 48–54 (1995).
[12] O. Rodrigues, "Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire," *J de mathématiq pures et appliqué* **5**, 380–440 (1840).
[13] J. M. McCarthy, "Planar and spatial rigid motion as special cases of spherical and 3-spherical motion," *J Mech Trans Autom Design* **105**(3), 569–575 (1983).

[14] K. R. Etzel and J. M. McCarthy, "Spatial Motion Interpolation in an Image Space of SO (4)," **In:** *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, (American Society of Mechanical Engineers, 1996) pp. V02BT02A015.
[15] P. M. Larochelle, A. P. Murray and J. Angeles, "A distance metric for finite sets of rigid-body displacements via the polar decomposition," *J Mech Design* **129**(8), 883–886 (2007).
[16] P. Fanghella and C. Galletti, "Metric relations and displacement groups in mechanism and robot kinematics," *J Mech Design* **117**(3), 470–478 (1995).
[17] K. Kazerounian and J. Rastegar, "Object Norms: A Class of Coordinate and Metric Independent Norms for Displacements," **In:** *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, (American Society of Mechanical Engineers, 1992) pp. 271–275.
[18] J. M. R. Martinez and J. Duffy, "On the metrics of rigid body displacements for infinite and finite bodies," *J Mech Design* **117**(3), 470–478 (1995).
[19] J. Gallier and D. Xu, "Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices," *Int J Robot Autom* **18**(1), 10–20 (2003).
[20] R. W. H. Hudson, *Kummer's Quartic Surface* (Cambridge Mathematical Library. Cambridge University Press, 1990).
[21] T. D. Barfoot, *State Estimation for Robotics* (Cambridge University Press, 2024).
[22] Y. Wang and G. S. Chirikjian, "Nonparametric second-order theory of error propagation on motion groups," *Int J Robot Res* **27**(11-12), 1258–1273 (2008).

Cite this article: Y. Wu and G. S. Chirikjian (2024). "Trace inequalities and kinematic metrics", Robotica **42**, 4103–4119. https://doi.org/10.1017/S0263574724000778