

Minimal geodesics on $GL(n)$ for left-invariant, right- $O(n)$ -invariant Riemannian metrics

Robert Martin* and Patrizio Neff†

September 30, 2014

We provide an easy approach to the geodesic distance on the general linear group $GL(n)$ for left-invariant Riemannian metrics which are also right- $O(n)$ -invariant. The parametrization of geodesic curves and the global existence of length minimizing geodesics are deduced using simple methods based on the calculus of variations and classical analysis only. The geodesic distance is discussed for some special cases and applications towards the theory of nonlinear elasticity are indicated.

1. Introduction and preliminaries

The interpretation of the general linear group $GL(n)$ as a Riemannian manifold instead of a simple subset of the linear matrix space $\mathbb{R}^{n \times n}$ has recently been motivated by results in the theory of nonlinear elasticity [22, 21], showing a connection between the *logarithmic strain measure* $\log \sqrt{F^T F}$ of a deformation gradient $F \in GL^+(n)$ and the *geodesic distance* of F to the special orthogonal group $SO(n)$. Since the requirements of objectivity and isotropy strongly suggest a distance measure on $GL^+(n)$ which is right invariant under rotations and left invariant with respect to action of $GL^+(n)$, we restrict our considerations to Riemannian metrics on $GL(n)$ which are left- $GL(n)$ -invariant as well as right $O(n)$ -invariant.

Although the theory of Lie groups is obviously applicable to $GL(n)$ with such a metric, this general approach utilizes many intricate results from the abstract theory of differential geometry and is therefore not easily accessible to readers not sufficiently familiar with these subjects. Furthermore, while the explicit parametrization of geodesic curves has been given for the canonical left-invariant metric on $GL^+(n)$ [1] as well as for left-invariant, right- $SO(n)$ -invariant¹ metrics on $SL(n)$ [17], analogous results are not found in the literature for the more general case on $GL(n)$.

The aim of this paper is therefore to provide a more accessible approach to this type of Riemannian metrics on $GL(n)$ and the induced geodesic distances as well as to deduce the parametrization of geodesic curves using only basic methods from the calculus of variations and classical analysis. In order to keep this article as self-contained as possible, we will begin by stating (and proving) some very basic facts on Riemannian metrics for the special case of $GL(n)$.

*Corresponding author, Lehrstuhl für Nichtlineare Analysis und Modellierung, Fakultät für Mathematik, Universität Duisburg-Essen, Thea-Leymann Str. 9, 45127 Essen, Germany, email: robert.martin@uni-due.de

†Head of Lehrstuhl für Nichtlineare Analysis und Modellierung, Fakultät für Mathematik, Universität Duisburg-Essen, Thea-Leymann Str. 9, 45127 Essen, Germany, email: patrizio.neff@uni-due.de

¹Note that Mielke only considers metrics induced by an inner product of the form (1.38), although the special case $n = 4$ admits a different, non-canonical inner product which is also bi-invariant under $SO(n)$; c.f. Remark 1.13.

1.1. Basic Definitions

Let $\mathbb{R}^{n \times n}$ denote the set of all $n \times n$ real matrices and let $\mathbb{1}$ denote the identity matrix in $\mathbb{R}^{n \times n}$. We define the groups

$$\begin{aligned} \mathrm{GL}(n) &= \{X \in \mathbb{R}^{n \times n} \mid \det(X) \neq 0\} && (\text{general linear group}), \\ \mathrm{GL}^+(n) &= \{X \in \mathbb{R}^{n \times n} \mid \det(X) > 0\}, \\ \mathrm{GL}^-(n) &= \{X \in \mathbb{R}^{n \times n} \mid \det(X) < 0\}, \\ \mathrm{SL}(n) &= \{X \in \mathbb{R}^{n \times n} \mid \det(X) = 1\} && (\text{special linear group}), \\ \mathrm{O}(n) &= \{X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1}\} && (\text{orthogonal group}), \\ \mathrm{SO}(n) &= \{X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1} \wedge \det(X) = 1\} && (\text{orthogonal group}). \end{aligned}$$

Furthermore, we define the set of *symmetric matrices* $\mathrm{Sym}(n) = \{X \in \mathbb{R}^{n \times n} \mid X^T = X\}$, the set of *positive definite symmetric matrices* $\mathrm{PSym}(n) = \{X \in \mathbb{R}^{n \times n} \mid X^T = X \wedge \forall v \in \mathbb{R}^n \setminus \{0\} : v^T X v > 0\}$ and the set of *skew symmetric matrices* $\mathfrak{so}(n) = \{X \in \mathbb{R}^{n \times n} \mid X^T = -X\}$.

1.2. Distance functions on $\mathbb{R}^{n \times n}$

A *distance* on a set M is a function $\mathrm{dist} : M \times M \rightarrow [0, \infty]$ with

$$\begin{aligned} \mathrm{dist}(A, A) &= 0, & \mathrm{dist}(A, B) &= \mathrm{dist}(B, A) \\ \text{and } \mathrm{dist}(A, C) &\leq \mathrm{dist}(A, B) + \mathrm{dist}(B, C) && (\text{triangle inequality}) \end{aligned} \quad (1.1)$$

for all $A, B, C \in M$. Note that dist is a *metric* on $\mathbb{R}^{n \times n}$ if and only if, additionally,

$$0 < \mathrm{dist}(A, B) < \infty \quad \text{for all } A \neq B.$$

The most common distance on $\mathbb{R}^{n \times n}$ is the *Euclidean distance*: The canonical inner product

$$\langle M, N \rangle = \mathrm{tr}(M^T N) = \sum_{i,j=1}^n M_{i,j} N_{i,j}, \quad M, N \in \mathbb{R}^{n \times n}, \quad (1.2)$$

where $\mathrm{tr} M = \sum_{i=1}^n M_{i,i}$ denotes the *trace* of M , induces the Euclidean norm (or *Frobenius norm*)

$$\|M\| = \sqrt{\langle M, M \rangle} = \sqrt{\sum_{i,j=1}^n M_{i,j}^2}, \quad (1.3)$$

and the Euclidean distance is the metric given by

$$\mathrm{dist}_{\mathrm{euclid}}(A, B) = \|A - B\|. \quad (1.4)$$

While $\mathrm{dist}_{\mathrm{euclid}}$ induces a distance function on $\mathrm{GL}(n)$ as well, it does not appear as a “natural” inner property of the general linear group: since $\mathrm{GL}(n)$ is not a linear space, the term $\|A - B\|$ depends on the underlying algebraic structure of the vector space $\mathbb{R}^{n \times n}$. Furthermore, because $\mathrm{GL}(n)$ is not a closed subset of $\mathbb{R}^{n \times n}$, it is not complete under the Euclidean distance.

A more proper distance measure should take into account the algebraic properties of $\mathrm{GL}(n)$ as a group. To find such a function we interpret $\mathrm{GL}(n)$ as a Riemannian manifold, more specifically as a submanifold of the vector space $\mathbb{R}^{n \times n}$: as an open subset of $\mathbb{R}^{n \times n}$, the tangent space $T_A \mathrm{GL}(n)$ at an arbitrary point $A \in \mathrm{GL}(n)$ is given by²

$$T_A \mathrm{GL}(n) = \mathbb{R}^{n \times n} =: \mathfrak{gl}(n). \quad (1.5)$$

To obtain a distance function respecting this structure on $\mathrm{GL}(n)$, we will consider a measurement along connecting curves.

²Note that in the theory of Lie Groups $\mathfrak{gl}(n)$ usually denotes the tangent space $T_{\mathbb{1}} \mathrm{GL}(n)$ at the identity only. From the perspective of classical analysis, all tangent spaces $T_A \mathrm{GL}(n)$ can simply be identified with $\mathbb{R}^{n \times n}$, allowing us to employ a much simpler notion of “smoothness”.

1.3. The geodesic distance

A Riemannian metric on $\mathrm{GL}(n)$ is a smooth (i.e. infinitely differentiable) function

$$g : \mathrm{GL}(n) \times \mathfrak{gl}(n) \times \mathfrak{gl}(n) \rightarrow \mathbb{R}, \quad (A, M, N) \mapsto g_A(M, N) \quad (1.6)$$

such that for every fixed $A \in \mathrm{GL}(n)$ the function $g_A(\cdot, \cdot) : \mathfrak{gl}(n) \times \mathfrak{gl}(n) \rightarrow \mathbb{R}$ is a positive definite symmetric bilinear form, i.e.

$$g_A(\lambda_1 M_1 + \lambda_2 M_2, N) = \lambda_1 g_A(M_1, N) + \lambda_2 g_A(M_2, N), \quad g_A(M, N) = g_A(N, M), \quad g_A(T, T) > 0$$

for all $M, N, M_1, M_2, T \in \mathfrak{gl}(n)$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $T \neq 0$. A Riemannian metric allows for a measurement of sufficiently smooth curves in $\mathrm{GL}(n)$: the *length* L and *energy* E of a curve $X \in C^1([a, b]; \mathrm{GL}(n))$ are given by

$$L(X) := \int_0^1 \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))} dt \quad \text{and} \quad E(X) := \int_0^1 g_{X(t)}(\dot{X}(t), \dot{X}(t)) dt, \quad (1.7)$$

where we employ the notation $\dot{X}(t) = \frac{d}{dt} X(t)$.

Note that the length of X is defined similarly to the length of curves in Euclidean spaces. Thus many well-known properties, like invariance under reparameterization, still hold in the Riemannian case. Some such properties will be discussed further in the following section.

The geodesic distance $\mathrm{dist}_{\mathrm{geod}}(A, B)$ between $A, B \in \mathrm{GL}(n)$ can now be defined as the infimum over the length of curves connecting A and B . For this we need an exact definition of the admissible sets of curves.

Definitions 1.1. We denote by

$$C_r^k([a, b]; \mathrm{GL}(n)) := \{X \in C^k([a, b]; \mathrm{GL}(n)) \mid \dot{X}(t) \neq 0 \forall t \in [a, b]\} \quad (1.8)$$

the set of *regular k -times differentiable* curves in $\mathrm{GL}(n)$ over the interval $[a, b]$. Note that, by the usual definition of differentiability on closed intervals as the restriction of differentiable functions on the i th derivative \mathbb{R} , $X^{(i)}(a)$ and $X^{(i)}(b)$ at the boundaries are well-defined as the one-sided limits $\lim_{t \searrow a} X^{(i)}(t) \neq 0$, $\lim_{t \nearrow b} X^{(i)}(t) \neq 0$. We furthermore define the set of *piecewise k -times differentiable curves* in $\mathrm{GL}(n)$ over the interval $[a, b]$ as

$$\begin{aligned} \mathcal{A}^k([a, b]) := \{X \in C^0([a, b]; \mathrm{GL}(n)) \mid \exists a = a_0 < a_1 < \dots < a_{m+1} = b : \\ X|_{[a_k, a_{k+1}]} \in C_r^k([a_k, a_{k+1}]; \mathrm{GL}(n)) \forall k = 0, \dots, m\}. \end{aligned} \quad (1.9)$$

Note that, by this definition, partially differentiable curves are continuous everywhere. Finally, for $A, B \in \mathrm{GL}(n)$, the *admissible set of curves connecting A and B* is

$$\mathcal{A}_A^B := \{X \in \mathcal{A}^1([a, b]) \mid a, b \in \mathbb{R}, X(a) = A, X(b) = B\}, \quad (1.10)$$

the admissible set over the fixed interval $[a, b]$ is denoted by

$$\mathcal{A}_A^B([a, b]) := \{X : [a, b] \rightarrow \mathrm{GL}(n) : X \in \mathcal{A}_A^B\} \quad (1.11)$$

and the general *admissible set of curves* is

$$\mathcal{A} := \bigcup_{A, B \in \mathrm{GL}(n)} \mathcal{A}_A^B. \quad (1.12)$$

Remark 1.2. While the notion of piecewise differentiability is often found in the literature, a specific definition is often omitted. The definition used here guarantees the existence of one-sided limits $\lim_{t \searrow t_0} X^{(i)}(t) \neq 0$, $\lim_{t \nearrow t_0} X^{(i)}(t) \neq 0$ everywhere and thus, in particular, that their length $L(X) < \infty$ is well-defined.

We can now properly define the geodesic distance function:

Definition 1.3. Let $A, B \in \text{GL}(n)$. Then

$$\text{dist}_{\text{geod}}(A, B) := \inf_{X \in \mathcal{A}_A^B} L(X) \quad (1.13)$$

is called the *geodesic distance* between two matrices A and B .

Remark 1.4. Note that $\text{dist}_{\text{geod}}$ is indeed a distance function: to see that it fulfils the triangle inequality (1.1), choose a curve $X \in \mathcal{A}_A^B$ connecting A and B with $L(X) \leq \text{dist}_{\text{geod}}(A, B) + \varepsilon$ as well as a curve $Y \in \mathcal{A}_B^C$ connecting B and C with $L(Y) \leq \text{dist}_{\text{geod}}(B, C) + \varepsilon$. We assume (without loss of generality, as we will see in Lemma 1.5) that both X and Y are defined on the interval $[0, 1]$. Let Z denote the curve obtained by “attaching” Y to X , i.e.

$$Z : [0, 2] \rightarrow \text{GL}(n) : \quad Z(t) = \begin{cases} X(t) & : t \in [0, 1] \\ Y(t-1) & : t \in [1, 2] \end{cases}. \quad (1.14)$$

Then Z is piecewise differentiable, $Z(0) = X(0) = A$ and $Z(2) = Y(1) = C$, and thus $Z \in \mathcal{A}_A^C$. We find³

$$\begin{aligned} L(Z) &= \int_0^2 \sqrt{g_{Z(t)}(\dot{Z}, \dot{Z})} dt = \int_0^1 \sqrt{g_{Z(t)}(\dot{Z}, \dot{Z})} dt + \int_1^2 \sqrt{g_{Z(t)}(\dot{Z}, \dot{Z})} dt \\ &= \int_0^1 \sqrt{g_{X(t)}(\dot{X}, \dot{X})} dt + \int_0^1 \sqrt{g_{Y(t)}(\dot{Y}, \dot{Y})} dt \\ &= L(X) + L(Y) \leq \text{dist}_{\text{geod}}(A, B) + \text{dist}_{\text{geod}}(B, C) + 2\varepsilon \end{aligned} \quad (1.15)$$

and thus $\text{dist}_{\text{geod}}(A, C) \leq L(Z) \leq \text{dist}_{\text{geod}}(A, B) + \text{dist}_{\text{geod}}(B, C) + 2\varepsilon$ for all $\varepsilon > 0$, which shows the triangle inequality.

The geodesic distance can be considered a generalization of the Euclidean distance: if we measure the length of a curve $\gamma : [a, b] \rightarrow \mathbb{R}^{n \times n}$ by $L(\gamma) = \int_a^b \sqrt{\langle \dot{\gamma}, \dot{\gamma} \rangle} dt$, then the shortest curve connecting $A, B \in \mathbb{R}^{n \times n}$ is a straight line with length $\|A - B\|$. Thus the Euclidean distance can be interpreted as the infimum over the length of connecting curves as well.

Furthermore, note that $\text{GL}(n)$ is not connected but has two connected components $\text{GL}^+(n) = \{A \in \text{GL}(n) \mid \det A > 0\}$ and $\text{GL}^-(n) = \{A \in \text{GL}(n) \mid \det A < 0\}$ [6]. Thus $\text{dist}_{\text{geod}}(A, B) < \infty$ if and only if A and B are in the same connected component, i.e. iff $\det(AB) > 0$. As we will see later on, for left-invariant Riemannian metrics we can focus on the case $A, B \in \text{GL}^+(n)$ without loss of generality.

1.4. Length and energy of curves

This section will give a brief overview of basic properties of curves in the Riemannian manifold $\text{GL}(n)$ as well as the length and energy functional. Most of these properties directly correspond to the case of curves in the Euclidean space, and thus most of the proofs are adapted from these cases as well (see e.g. [13, chapter 12]).

Consider the integrand $\sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))}$ in the definition of the length of a curve X . In analogy to the Euclidean space, we call this term the *speed* of X . Since, by definition of a Riemannian metric, $g_A(\cdot, \cdot)$ defines an inner product on the tangent space $T_A \text{GL}(n) = \mathfrak{gl}(n)$, it induces a norm on $\mathfrak{gl}(n)$. We will therefore simplify notation by writing $\|M\|_A = \sqrt{g_A(M, M)}$ for $A \in \text{GL}(n)$, $M \in \mathfrak{gl}(n)$. Then a piecewise differentiable curve $X \in \mathcal{A}([a, b])$ has *constant speed* with regard to the Riemannian metric g if $\|\dot{X}(t)\|_{X(t)} = \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))}$ is constant on $[a, b]$.

³Here and henceforth we will often omit the integration variable and write e.g. \dot{Z} instead of $\dot{Z}(t)$.

Lemma 1.5. *Let $X \in \mathcal{A}^1([a, b])$. Then there exists a unique piecewise differentiable $\varphi \in C^0([a, b]; [a, b])$, $\varphi(a) = a$, $\varphi(b) = b$, $\varphi'(t) > 0$ such that $X \circ \varphi$ has constant speed.*

Proof. Choose $a = a_0 < \dots < a_{m+1} = b$ such that X is continuously differentiable on $[a, b] \setminus \{a_0, \dots, a_{m+1}\}$. Then the function

$$l : [a, b] \rightarrow [0, L(X)], \quad l(t) := \int_a^t \|\dot{X}(t)\|_{X(t)} dt = L(X|_{[a,t]}) \quad (1.16)$$

is continuously differentiable on $[a, b] \setminus \{a_0, \dots, a_{m+1}\}$ as well with $l'(t) = \|\dot{X}(t)\|_{X(t)}$. By definition of regular curves (Definition 1.1), $\dot{X}(t) \neq 0$ for all $t \in [a, b] \setminus \{a_0, \dots, a_{m+1}\}$ and, because of the positive definiteness of Riemannian metrics, $\|\dot{X}(t)\|_{X(t)} = \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))} > 0$. Thus l is a bijection and

$$\tilde{\varphi} : [0, L(X)] \rightarrow [a, b], \quad \tilde{\varphi}(s) = l^{-1}(s) \quad (1.17)$$

is well-defined, continuous on $[0, L(X)]$ and continuously differentiable on $[0, L(X)] \setminus \{l(a_0), \dots, l(a_{m+1})\}$. Let $\tilde{X} := X \circ \tilde{\varphi}$. Then

$$\dot{\tilde{X}}(t) = \frac{d}{dt}(X \circ \tilde{\varphi})(t) = \tilde{\varphi}'(t)\dot{X}(\tilde{\varphi}(t)) = \frac{1}{l'(\tilde{\varphi}(t))}\dot{X}(\tilde{\varphi}(t)) = \frac{\dot{X}(\tilde{\varphi}(t))}{\|\dot{X}(\tilde{\varphi}(t))\|_{X(\tilde{\varphi}(t))}} \quad (1.18)$$

if \tilde{X} is differentiable at t and thus $\|\dot{\tilde{X}}(\tilde{\varphi}(t))\|_{X(\tilde{\varphi}(t))} \equiv 1$, hence the transformation

$$\varphi : [a, b] \rightarrow [a, b], \quad \varphi(t) := \tilde{\varphi}\left((t-a)\frac{L(X)}{b-a}\right) \quad (1.19)$$

has the desired properties.

To see the uniqueness of φ , note that if both X and $X \circ \hat{\varphi}$ have constant speed, then

$$\begin{aligned} \left\| \frac{d}{dt}(X \circ \hat{\varphi})(t) \right\|_{(X \circ \hat{\varphi})(t)} \equiv c &\implies \|\hat{\varphi}'(t) \cdot \dot{X}(\hat{\varphi}(t))\|_{X(\hat{\varphi}(t))} \equiv c \\ &\implies |\hat{\varphi}'(t)| \equiv \frac{c}{\|\dot{X}(\hat{\varphi}(t))\|_{X(\hat{\varphi}(t))}} \equiv \text{constant} \end{aligned} \quad (1.20)$$

for $t \in [a, b] \setminus \{a_0, \dots, a_{m+1}\}$. Then the restrictions $\hat{\varphi}(a) = a$, $\hat{\varphi}(b) = b$ and $\hat{\varphi}' > 0$ only allow for $\hat{\varphi} \equiv 1$. \blacksquare

For $X \in \mathcal{A}([a, b])$, we denote by

$$X_c := X \circ \varphi \in \mathcal{A}([a, b])$$

the (unique) reparametrization from Lemma 1.5.

An important property of the length functional is its invariance under reparametrizations:

Lemma 1.6. *Let $X \in \mathcal{A}([a, b])$, and let $\varphi \in C^0([c, d]; [a, b])$ be a piecewise continuously differentiable function with $\varphi(c) = a$, $\varphi(d) = b$ and $\varphi'(t) > 0$. Then*

$$L(X \circ \varphi) = L(X). \quad (1.21)$$

Proof. Choose $c = a_0 < a_1 < \dots < a_{m+1} = d$ such that φ is differentiable on $[c, d] \setminus \{a_0, \dots, a_{m+1}\}$ with $\varphi' > 0$ and X is differentiable on $[a, b] \setminus \{\varphi(a_0), \dots, \varphi(a_{m+1})\}$. Then

$$\begin{aligned} L(X \circ \varphi) &= \int_c^d \left\| \frac{d}{dt}(X \circ \varphi)(t) \right\|_{X \circ \varphi(t)} dt = \sum_{i=0}^m \int_{a_i}^{a_{i+1}} \|\varphi'(t) \dot{X}(\varphi(t))\|_{X \circ \varphi(t)} dt \\ &= \sum_{i=0}^m \int_{a_i}^{a_{i+1}} |\varphi'(t)| \|\dot{X}(\varphi(t))\|_{X(\varphi(t))} dt = \sum_{i=0}^m \int_{a_i}^{a_{i+1}} \varphi'(t) \|\dot{X}(\varphi(t))\|_{X(\varphi(t))} dt \\ &= \sum_{i=0}^m \int_{\varphi(a_i)}^{\varphi(a_{i+1})} \|\dot{X}(t)\|_{X(t)} dt = \int_a^b \|\dot{X}(t)\|_{X(t)} dt = L(X). \end{aligned} \quad \blacksquare$$

To explicitly compute the geodesic distance, we will primarily search for *length minimizers*, i.e. curves

$$X \in \mathcal{A}_A^B : L(X) = \inf_{Y \in \mathcal{A}_A^B} L(Y) = \text{dist}_{\text{geod}}(A, B). \quad (1.22)$$

Since every curve on $[a, b]$ can be reparametrized (by scaling and shifting) to a curve of the same length on an arbitrary interval $[c, d]$, a restriction of the admissible interval does not change the infimal length:

$$\inf_{Y \in \mathcal{A}_A^B} L(Y) = \inf_{Y \in \mathcal{A}_A^B([a, b])} L(Y); \quad (1.23)$$

recall from definition 1.1 that $\mathcal{A}_A^B([a, b])$ denotes the set of piecewise differentiable curves defined on $[a, b]$ connecting A and B . Furthermore, if there exists a length minimizer in \mathcal{A}_A^B , there exists one in $\mathcal{A}_A^B([a, b])$ as well. If we are interested only in the length of a curve X , we will therefore often assume without loss of generality that X is defined on $[0, 1]$ and that X has constant speed.

However, the energy functional is *not* invariant under reparameterization: For a given curve $X : [0, 1] \rightarrow \text{GL}(n)$ and $\lambda > 0$, the energy of the curve $Y : [0, \lambda] \rightarrow \text{GL}(n) : Y(t) = X(\frac{t}{\lambda})$ is

$$\begin{aligned} E(Y) &= \int_0^\lambda g_{Y(t)}(\dot{Y}(t), \dot{Y}(t)) dt = \int_0^\lambda g_{X(\frac{t}{\lambda})}(\frac{1}{\lambda}\dot{X}(\frac{t}{\lambda}), \frac{1}{\lambda}\dot{X}(\frac{t}{\lambda})) dt \\ &= \frac{1}{\lambda} \int_0^\lambda \frac{1}{\lambda} g_{X(\frac{t}{\lambda})}(\dot{X}(\frac{t}{\lambda}), \dot{X}(\frac{t}{\lambda})) dt \\ &= \frac{1}{\lambda} \int_0^1 g_{X(t)}(\dot{X}(t), \dot{X}(t)) dt = \frac{1}{\lambda} E(X). \end{aligned} \quad (1.24)$$

For $\lambda \rightarrow \infty$, we see that by admitting arbitrary parametrizations, the infimal energy of curves connecting A, B is zero whenever A and B can be connected:

$$\inf_{Y \in \mathcal{A}_A^B} E(Y) = 0 \quad \forall A, B \in \text{GL}(n) : \det(AB) > 0. \quad (1.25)$$

We will therefore call $X : [a, b] \rightarrow \text{GL}(n) : X \in \mathcal{A}_A^B([a, b])$ an *energy minimizer* if and only if it minimizes the energy over all curves *over the same parameter interval* connecting A and B :

$$X \in \mathcal{A}_A^B([a, b]) : E(X) = \inf_{Y \in \mathcal{A}_A^B([a, b])} E(Y). \quad (1.26)$$

The next two lemmas show important relations between length minimizers and energy minimizers.

Lemma 1.7. *Let $X \in \mathcal{A}([a, b])$ be a piecewise differentiable curve. Then*

$$L(X) \leq \sqrt{b-a} \sqrt{E(X)}, \quad (1.27)$$

and equality holds if and only if X has constant speed.

Proof.

$$\begin{aligned} [L(X)]^2 &= \left(\int_a^b \|\dot{X}(t)\|_{X(t)} dt \right)^2 = \left(\int_a^b 1 \cdot \|\dot{X}(t)\|_{X(t)} dt \right)^2 \\ &\leq \int_a^b 1 dt \cdot \int_a^b \|\dot{X}(t)\|_{X(t)}^2 dt = (b-a)E(X). \end{aligned} \quad (1.28)$$

The inequality in (1.28) is due to the Hölder inequality, and equality holds if and only if $\|\dot{X}\|_X$ and 1 are linearly dependent in $L^2([a, b])$, meaning $\|\dot{X}\|_X \equiv \text{constant}$ on $[a, b]$. \blacksquare

Lemma 1.8. For $X \in \mathcal{A}_A^B([a, b])$, the following are equivalent:

$$\begin{aligned} i) \quad E(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} E(Y), \\ ii) \quad L(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} L(Y) = \min_{Y \in \mathcal{A}_A^B} L(Y) \text{ and } X \text{ has constant speed.} \end{aligned}$$

Proof. Recall that X_c denotes the (unique) parameterization of X with constant speed on $[a, b]$. The invariance of the length under reparameterization implies $L(X_c) = L(X)$, and inequality (1.27) from the previous lemma yields

$$E(X) \geq \frac{1}{b-a} L(X)^2 = \frac{1}{b-a} L(X_c)^2 = E(X_c), \quad (1.29)$$

where, again, the equality holds if and only if X has constant speed, i.e. $X = X_c$. Therefore, every minimizer of E must have constant speed, and it remains to show that, if X has constant speed, the equivalence

$$E(X) = \min_{Y \in \mathcal{A}_A^B([a, b])} E(Y) \iff L(X) = \min_{Y \in \mathcal{A}_A^B([a, b])} L(Y) \quad (1.30)$$

holds. We find

$$\begin{aligned} E(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} E(Y) & (1.31) \\ \Rightarrow L(Y) &= L(Y_c) = \sqrt{b-a} \sqrt{E(Y_c)} \geq \sqrt{b-a} \sqrt{E(X)} = L(X) \quad \forall Y \in \mathcal{A}_A^B([a, b]) \\ \Rightarrow L(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} L(Y), \end{aligned}$$

as well as

$$\begin{aligned} L(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} L(Y) & (1.32) \\ \Rightarrow E(Y) &\geq \frac{1}{b-a} L(Y)^2 \geq \frac{1}{b-a} L(X)^2 = E(X) \quad \forall Y \in \mathcal{A}_A^B([a, b]) \\ \Rightarrow E(X) &= \min_{Y \in \mathcal{A}_A^B([a, b])} E(Y). \quad \blacksquare \end{aligned}$$

1.5. Left-invariant, right- $O(n)$ -invariant Riemannian metrics

In the following sections we will only consider Riemannian metrics that are left-invariant as well as right-invariant under $O(n)$.

Definition 1.9. A Riemannian metric g on $GL(n)$ is called *left $GL(n)$ -invariant* (or simply *left-invariant*) if

$$g_{CA}(CM, CN) = g_A(M, N) \quad (1.33)$$

for all $A, C \in GL(n)$ and $M, N \in \mathfrak{gl}(n)$.

The left-invariance of a Riemannian metric g can be applied directly to the geodesic distance: let $A, B, C \in GL(n)$. Then for every given curve $X \in \mathcal{A}_A^B$ connecting A and B we can define a curve $Y = CX \in \mathcal{A}_{(CA)}^{(CB)}$ connecting CA and CB by $Y(t) = C \cdot X(t)$. We assume without loss of generality that X is defined on the interval $[0, 1]$ and find

$$\begin{aligned} L(Y) &= \int_0^1 \sqrt{g_{Y(t)}(\dot{Y}(t), \dot{Y}(t))} dt & (1.34) \\ &= \int_0^1 \sqrt{g_{CX(t)}(C\dot{X}(t), C\dot{X}(t))} dt = \int_0^1 \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))} dt = L(X). \end{aligned}$$

Analogously, for every curve Y connecting CA and CB , the curve $X = C^{-1}Y$ connects A and B with $L(X) = L(Y)$. Thus for every curve connecting A and B , we can find a curve of equal length connecting CA and CB and vice versa. Therefore

$$\text{dist}_{\text{geod}}(CA, CB) = \inf_{Y \in \mathcal{A}_{(CA)}^{(CB)}} L(Y) = \inf_{X \in \mathcal{A}_A^B} L(X) = \text{dist}_{\text{geod}}(A, B). \quad (1.35)$$

In particular, this left invariance of the geodesic distance implies

$$\text{dist}_{\text{geod}}(A, B) = \text{dist}_{\text{geod}}(\mathbb{1}, A^{-1}B). \quad (1.36)$$

We will therefore often focus on the case $A = \mathbb{1}$. Note that $A^{-1}B \in \text{GL}^+(n)$ for $A, B \in \text{GL}^-(n)$, hence the geodesic distance on $\text{GL}^-(n)$ is completely determined by the geodesic distance on $\text{GL}^+(n)$.

Definition 1.10. A Riemannian metric g on $\text{GL}(n)$ is called *right-O(n)-invariant* if

$$g_{AQ}(MQ, NQ) = g_A(M, N) \quad (1.37)$$

for all $A \in \text{GL}(n)$, $M, N \in \mathfrak{gl}(n)$ and $Q \in \text{O}(n)$.

Such invariant metrics appear in the theory of elasticity, where right-O(n)-invariance follows from material isotropy, while objectivity implies the left invariance. We will show that a Riemannian metric fulfilling both invariances is uniquely determined up to three parameters $\mu, \mu_c, \kappa > 0$ and given by an *isotropic inner product* $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$ of the form

$$\begin{aligned} \langle M, N \rangle_{\mu, \mu_c, \kappa} &:= \mu \langle \text{dev sym } M, \text{dev sym } N \rangle + \mu_c \langle \text{skew } M, \text{skew } N \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } N \\ &= \mu \text{tr}[(\text{dev sym } M)^T \text{dev sym } N] + \mu_c \text{tr}[(\text{skew } M)^T \text{skew } N] + \frac{\kappa}{n} \text{tr } M \text{tr } N, \end{aligned} \quad (1.38)$$

where $\text{sym } M = \frac{1}{2}(M + M^T)$ denotes the symmetric part, $\text{skew } M = \frac{1}{2}(M - M^T)$ is the skew symmetric part and $\text{dev } M = M - \frac{\text{tr } M}{n} \mathbb{1}$ denotes the deviator of M . Note also that $\langle M, N \rangle = \text{tr}[M^T N]$ denotes the canonical inner product on $\mathfrak{gl}(n)$ here and throughout.

We will note some basic properties of $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$. The proof can be found in the appendix (Lemma A.7).

Lemma 1.11. *Let $M, N \in \mathfrak{gl}(n)$. Then*

$$i) \quad \langle Q^T M Q, Q^T N Q \rangle_{\mu, \mu_c, \kappa} = \langle M, N \rangle_{\mu, \mu_c, \kappa} \quad \text{for all } Q \in \text{O}(n), \quad (1.39)$$

$$\begin{aligned} ii) \quad \langle M, N \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev sym } M, N \rangle + \mu_c \langle \text{skew } M, N \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } N, \\ &= \langle \mu \text{dev sym } M + \mu_c \text{skew } M + \frac{\kappa}{n} \text{tr}[M] \mathbb{1}, N \rangle \end{aligned} \quad (1.40)$$

$$iii) \quad \langle M, N \rangle_{1,1,1} = \langle M, N \rangle, \quad (1.41)$$

$$iv) \quad \langle S, W \rangle_{\mu, \mu_c, \kappa} = 0 \quad \text{for all } S \in \text{Sym}(n), W \in \mathfrak{so}(n), \quad (1.42)$$

where $\text{Sym}(n)$ and $\mathfrak{so}(n)$ denote the sets of symmetric and skew symmetric matrices in $\mathbb{R}^{n \times n}$ respectively.

Note that (1.41) implies that the canonical inner product can be interpreted as a special case of $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$ with $\mu = \mu_c = \kappa = 1$. Furthermore we will denote by

$$\|M\|_{\mu, \mu_c, \kappa} = \sqrt{\langle M, M \rangle_{\mu, \mu_c, \kappa}} = \sqrt{\mu \|\text{dev sym } M\|^2 + \mu_c \|\text{skew } M\|^2 + \frac{\kappa}{n} \text{tr}[M]^2} \quad (1.43)$$

the norm induced by $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$. Note that $\|M\| = \sum_{i,j=1}^n M_{i,j}^2$ denotes the canonical matrix norm here and throughout.

We can now show the connection between the isotropic inner product and left-invariant, right-O(n)-invariant Riemannian metrics.

Proposition 1.12.

i) A Riemannian metric g on $\mathrm{GL}(n)$ is left-invariant if and only if g is of the form

$$g_A(M, N) = \langle A^{-1}M, A^{-1}N \rangle_g, \quad (1.44)$$

where $\langle \cdot, \cdot \rangle_g$ is an inner product on the tangent space $\mathfrak{gl}(n) = T_{\mathbb{I}} \mathrm{GL}(n)$ at the identity \mathbb{I} .

ii) A left-invariant metric g is additionally right- $\mathrm{O}(n)$ -invariant if and only if g is of the form

$$g_A(M, N) = \langle A^{-1}M, A^{-1}N \rangle_{\mu, \mu_c, \kappa} \quad (1.45)$$

with $\mu, \mu_c, \kappa > 0$.

Proof. i): Let $\langle \cdot, \cdot \rangle_g$ be an inner product on $\mathfrak{gl}(n)$ and g be defined by $g_A(M, N) = \langle A^{-1}M, A^{-1}N \rangle_g$. Then we find

$$\begin{aligned} g_{CA}(CM, CN) &= \langle (CA)^{-1}CM, (CA)^{-1}CN \rangle_g \\ &= \langle A^{-1}C^{-1}CM, A^{-1}C^{-1}CN \rangle_g = \langle A^{-1}M, A^{-1}N \rangle_g = g_A(M, N), \end{aligned} \quad (1.46)$$

hence g is left-invariant.

Now let g be an arbitrary left-invariant Riemannian metric on $\mathrm{GL}(n)$. Then

$$\langle M, N \rangle_g := g_{\mathbb{I}}(M, N) \quad (1.47)$$

defines an inner product on $\mathfrak{gl}(n)$, and we find

$$g_A(M, N) = g_{A \cdot \mathbb{I}}(AA^{-1}M, AA^{-1}N) = g_{\mathbb{I}}(A^{-1}M, A^{-1}N) = \langle A^{-1}M, A^{-1}N \rangle_g.$$

ii): If g is defined by $g_A(M, N) = \langle A^{-1}M, A^{-1}N \rangle_{\mu, \mu_c, \kappa}$, then g is left-invariant according to i). Let $Q \in \mathrm{O}(n)$. Then

$$g_{AQ}(MQ, NQ) = \langle (AQ)^{-1}MQ, (AQ)^{-1}NQ \rangle_{\mu, \mu_c, \kappa} = \langle Q^T A^{-1}MQ, Q^T A^{-1}NQ \rangle_{\mu, \mu_c, \kappa}.$$

We apply the isotropic property (1.39) to find

$$\langle Q^T A^{-1}MQ, Q^T A^{-1}NQ \rangle_{\mu, \mu_c, \kappa} = \langle A^{-1}M, A^{-1}N \rangle_{\mu, \mu_c, \kappa} = g_A(M, N), \quad (1.48)$$

which implies the right- $\mathrm{O}(n)$ -invariance of g .

Finally, let g be an arbitrary left-invariant, right- $\mathrm{O}(n)$ -invariant metric on $\mathrm{GL}(n)$. Again we define the inner product $\langle \cdot, \cdot \rangle_g$ by $\langle M, N \rangle_g := g_{\mathbb{I}}(M, N)$. Then for every $Q \in \mathrm{O}(n)$

$$\begin{aligned} \langle Q^T MQ, Q^T NQ \rangle_g &= g_{\mathbb{I}}(Q^T MQ, Q^T NQ) \\ &= g_{Q \cdot \mathbb{I}}(MQ, NQ) = g_{\mathbb{I}Q}(MQ, NQ) = g_{\mathbb{I}}(M, N) = \langle M, N \rangle_g. \end{aligned} \quad (1.49)$$

According to a well-known representation formula for isotropic linear mappings on $\mathbb{R}^{n \times n}$ [4], this invariance directly implies that $\langle \cdot, \cdot \rangle_g$ has the desired form (1.38). \blacksquare

Remark 1.13. Note that we required the Riemannian metric g to be defined on all of $\mathrm{GL}(n)$ and to be right-invariant under $\mathrm{O}(n)$ in order to obtain the representation formulae (1.45), (1.38) for g from equality (1.49). However, these requirements could be relaxed in all cases except $n = 4$: if g is defined only on $\mathrm{GL}^+(n)$ and is left-invariant under $\mathrm{GL}^+(n)$ as well as right invariant under $\mathrm{SO}(n)$, then (1.49) holds for all $Q \in \mathrm{SO}(n)$. If $n \neq 4$, then this already implies that $\langle \cdot, \cdot \rangle_g$ is of the form (1.38).

In the following sections g will denote a left-invariant, right- $\mathrm{O}(n)$ -invariant Riemannian metric as given in (1.45), unless otherwise indicated.

2. Energy minimizers and the calculus of variations

In the theory of Lie Groups and, more generally, Riemannian geometry, it can be shown that every length minimizing curve on a sufficiently smooth Riemannian manifold is a geodesic [12]. However, by focussing solely on the considered special case, we can circumvent the methods of abstract Riemannian geometry (such as the Levi-Civita-connection) and find a differential equation characterizing the minimizing curves by a straightforward application of the calculus of variations. A similar approach can be found in [15], where a geodesic equation similar to (2.8) is computed for the *right-invariant* Riemannian metric induced by the canonical inner product on $\mathfrak{gl}(n)$.

In preparation we need the following two lemmas which are taken from [20] and [19], respectively. Note that for a continuously differentiable function $f \in C^1(\mathrm{GL}(n), \mathrm{GL}(n))$, $Df[M] \in L(\mathfrak{gl}(n), \mathfrak{gl}(n))$ denotes the total derivative of f at $M \in \mathrm{GL}(n)$ and $Df[M].H \in \mathfrak{gl}(n)$ denotes its directional derivative at M in direction H .

Lemma 2.1. *Define $f \in C^1(\mathrm{GL}(n), \mathrm{GL}(n))$ by $f(X) = X^{-1}$. Then*

$$Df[A].H = -A^{-1}HA^{-1}. \quad (2.1)$$

Proof. A short proof can be found in [20]. ■

Lemma 2.2. *Let $M, N, P \in \mathfrak{gl}(n)$. Then*

$$\langle M, PN \rangle = \langle P^T M, N \rangle \quad \text{and} \quad \langle M, NP \rangle = \langle MP^T, N \rangle. \quad (2.2)$$

Proof. See [19]. ■

2.1. The geodesic equation for the canonical inner product

First, assume that $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$ is the canonical inner product on $\mathbb{R}^{n \times n}$, i.e. $\mu = \mu_c = \kappa = 1$. Let $X \in \mathcal{A}_A^B \cap C^2([0, 1]; \mathrm{GL}^+(n))$ be a two-times differentiable length minimizing curve from $[0, 1]$ to $\mathrm{GL}^+(n)$, where the length is measured by the left-invariant Riemannian metric g with

$$g_Z(X, Y) := \langle Z^{-1}X, Z^{-1}Y \rangle = \mathrm{tr}((Z^{-1}X)^T Z^{-1}Y),$$

and assume without loss of generality (Lemma 1.5) that X has constant speed. Then it follows from Lemma 1.8 that X also minimizes the energy functional E . We will characterize the minimizer X by the Euler–Lagrange equation corresponding to the energy, which we will call the *geodesic equation*, in consistence with a more general differential geometric definition of geodesics [12, Definition 1.4.2].

Let

$$C_0^1([0, 1]; \mathfrak{gl}(n)) := \{V \in C^1([0, 1]; \mathfrak{gl}(n)) \mid \mathrm{supp}(V) \subset (0, 1)\} \quad (2.3)$$

denote the set of *variations* in $\mathfrak{gl}(n)$ with compact support. For a fixed variation $V \in C_0^1([0, 1]; \mathfrak{gl}(n))$, define $Z_h := X + hV$ for $h \in (-\varepsilon, \varepsilon)$. For sufficiently small ε , we have $Z_h(t) \in \mathrm{GL}^+(n)$ for all $t \in [0, 1]$ and therefore $Z_h \in \mathcal{A}_A^B$ (note that $V(0) = V(1) = 0$ and thus $Z_h(0) = X(0) = A, Z_h(1) = B$). The minimizing property of X implies the stationarity condition

$$\begin{aligned} 0 &= \frac{1}{2} \frac{d}{dh} E(X + hV) \Big|_{h=0} = \frac{1}{2} \frac{d}{dh} \int_0^1 \langle Z_h^{-1} \dot{Z}_h, Z_h^{-1} \dot{Z}_h \rangle dt \Big|_{h=0} \\ &= \frac{1}{2} \int_0^1 \frac{d}{dh} \langle Z_h^{-1} \dot{Z}_h, Z_h^{-1} \dot{Z}_h \rangle \Big|_{h=0} dt \\ &= \frac{1}{2} \int_0^1 2 \langle Z_h^{-1} \dot{Z}_h \Big|_{h=0}, \frac{d}{dh} Z_h^{-1} \dot{Z}_h \Big|_{h=0} \rangle dt \\ &= \int_0^1 \langle Z_h^{-1} \dot{Z}_h \Big|_{h=0}, \frac{d}{dh} Z_h^{-1} \dot{Z}_h \Big|_{h=0} \rangle dt. \end{aligned} \quad (2.4)$$

We easily calculate $Z_h^{-1}|_{h=0} = X^{-1}$, $\dot{Z}_h|_{h=0} = \dot{X}$ and, using Lemma 2.1,

$$\frac{d}{dh} Z_h^{-1} \dot{Z}_h \Big|_{h=0} = \frac{d}{dh} (X + hV)^{-1} \Big|_{h=0} = -(X + hV)^{-1} V (X + hV)^{-1} \Big|_{h=0} = X^{-1} V X^{-1}.$$

Thus, with integration by parts, (2.4) computes to

$$\begin{aligned} \int_0^1 \langle X^{-1} \dot{X}, -X^{-1} V X^{-1} \dot{X} + X^{-1} \dot{V} \rangle dt &= \int_0^1 -\langle X^{-T} X^{-1} \dot{X}, V X^{-1} \dot{X} \rangle + \langle X^{-T} X^{-1} \dot{X}, \dot{V} \rangle dt \\ &= \int_0^1 -\langle X^{-T} X^{-1} \dot{X} (X^{-1} \dot{X})^T, V \rangle - \left\langle \frac{d}{dt} (X^{-T} X^{-1} \dot{X}), V \right\rangle dt \\ &= - \int_0^1 \langle X^{-T} X^{-1} \dot{X} (X^{-1} \dot{X})^T + \frac{d}{dt} (X^{-T} X^{-1} \dot{X}), V \rangle dt. \end{aligned}$$

Since this equation holds for all $V \in C_0^1([0, 1]; \mathfrak{gl}(n))$, we can apply the fundamental lemma of the calculus of variations to obtain the differential equation

$$\frac{d}{dt} (X^{-T} X^{-1} \dot{X}) = -X^{-T} X^{-1} \dot{X} (X^{-1} \dot{X})^T. \quad (2.5)$$

We compute the left hand term (using the product rule for matrix valued functions [20] as well as, again, Lemma 2.1):

$$\begin{aligned} \frac{d}{dt} (X^{-T} X^{-1} \dot{X}) &= (-X^{-T} \dot{X}^T X^{-T}) (X^{-1} \dot{X}) + X^{-T} \frac{d}{dt} (X^{-1} \dot{X}) \\ &= -X^{-T} \dot{X}^T X^{-T} X^{-1} \dot{X} + X^{-T} (-X^{-1} \dot{X} X^{-1} \dot{X} + X^{-1} \ddot{X}) \\ &= X^{-T} (-\dot{X}^T X^{-T} X^{-1} \dot{X} - X^{-1} \dot{X} X^{-1} \dot{X} + X^{-1} \ddot{X}). \end{aligned} \quad (2.6)$$

To simplify notation we define $U := X^{-1} \dot{X}$. Then

$$\dot{U} = -X^{-1} \dot{X} X^{-1} \dot{X} + X^{-1} \ddot{X} = -UU + X^{-1} \ddot{X}, \quad (2.7)$$

and combining (2.5) with (2.6) we obtain

$$\begin{aligned} -X^{-T} \overbrace{X^{-1} \dot{X}}^{=U} \overbrace{(X^{-1} \dot{X})^T}^{=U^T} &= X^{-T} (-\overbrace{(X^{-1} \dot{X})^T X^{-1} \dot{X}}^{=U^T U} - \overbrace{X^{-1} \dot{X} X^{-1} \dot{X}}^{=UU} + X^{-1} \ddot{X}) \\ \Rightarrow -UU^T &= -U^T U - UU + X^{-1} \ddot{X} \\ \Rightarrow U^T U - UU^T &= -UU + X^{-1} \ddot{X} \\ \stackrel{(2.7)}{\Rightarrow} \dot{U} &= U^T U - UU^T. \end{aligned} \quad (2.8)$$

Equation (2.8) can also be written as $\dot{U} = [U^T, U]$, where $[A, B] = AB - BA$ denotes the *commutator* on $\mathfrak{gl}(n)$. Vandereycken et al. [25] give an analogous equation for the case of the canonical right-invariant metric.

2.2. The geodesic equation for the general metric

Let us now consider the general case of arbitrary $\mu, \mu_c, \kappa > 0$. To devise a differential equation for this case, we need the following lemma.

Lemma 2.3. *Let $A, B \in \mathfrak{gl}(n)$. Then*

$$\langle A, [A, B] \rangle_{\mu, \mu_c, \kappa} = \frac{\mu + \mu_c}{2} \langle A, [A, B] \rangle. \quad (2.9)$$

Proof. Since $\text{tr}(AB) = \text{tr}(BA)$, we immediately see that

$$\text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0.$$

Therefore

$$\begin{aligned} \langle \text{sym } A, \text{sym}(AB - BA) \rangle &= \langle \text{dev sym } A, \text{dev sym}(AB - BA) \rangle + \text{tr } A \text{tr}(AB - BA) \\ &= \langle \text{dev sym } A, \text{dev sym}(AB - BA) \rangle \end{aligned} \quad (2.10)$$

and

$$\begin{aligned} &\langle \text{sym } A, \text{sym}(AB - BA) \rangle - \langle \text{skew } A, \text{skew}(AB - BA) \rangle \\ &= \langle \text{sym } A, AB - BA \rangle - \langle \text{skew } A, AB - BA \rangle \\ &= \langle \text{sym } A - \text{skew } A, AB - BA \rangle = \langle A^T, AB - BA \rangle = \text{tr}(AAB) - \text{tr}(ABA) = 0, \end{aligned} \quad (2.11)$$

thus

$$\begin{aligned} \langle A, [A, B] \rangle_{\mu, \mu_c, \kappa} &= \langle A, (AB - BA) \rangle_{\mu, \mu_c, \kappa} \\ &= \mu \langle \text{dev sym } A, \text{dev sym}(AB - BA) \rangle + \mu_c \langle \text{skew } A, \text{skew}(AB - BA) \rangle \\ &= \mu \langle \text{sym } A, \text{sym}(AB - BA) \rangle + \mu_c \langle \text{skew } A, \text{skew}(AB - BA) \rangle. \\ &\stackrel{(2.11)}{=} (\mu + \mu_c) \langle \text{sym } A, \text{sym}(AB - BA) \rangle \\ &\stackrel{(2.11)}{=} \frac{(\mu + \mu_c)}{2} (\langle \text{sym } A, \text{sym}(AB - BA) \rangle + \langle \text{skew } A, \text{skew}(AB - BA) \rangle) \\ &= \frac{(\mu + \mu_c)}{2} \langle A, (AB - BA) \rangle. \quad \blacksquare \end{aligned}$$

Let $X \in \mathcal{A}_A^B$ be a *piecewise two-times differentiable* energy minimizer with regard to the Riemannian metric

$$\begin{aligned} g_Z(X, Y) &:= \langle Z^{-1}X, Z^{-1}Y \rangle_{\mu, \mu_c, \kappa}, \\ \langle A, B \rangle_{\mu, \mu_c, \kappa} &:= \mu \langle \text{dev sym } A, \text{dev sym } B \rangle + \mu_c \langle \text{skew } A, \text{skew } B \rangle + \frac{\kappa}{n} \text{tr } A \text{tr } B. \end{aligned} \quad (2.12)$$

Note that, in contrast to section 2.1, we do not require X to be differentiable everywhere, but $X|_{[a_k, a_{k+1}]} \in C^2([a_k, a_{k+1}]; \text{GL}^+(n))$ for $k \in \{1, \dots, m\}$. Furthermore, following an approach by Lee et al. [15], we consider a different variation: for $Y_h := X(\mathbb{1} + hV)$, $V \in C_0^1([0, 1]; \mathfrak{gl}(n))$, the equation

$$0 = \frac{1}{2} \frac{d}{dh} E(Y_h) \Big|_{h=0} \quad (2.13)$$

holds for the energy minimizer X . Again, define $U := X^{-1}\dot{X}$, as well as $U_h := Y_h^{-1}\dot{Y}_h$. We compute

$$\begin{aligned} &Y_h(U + h \overbrace{(UV - VU + \dot{V})}^{=:C}) - h^2 X V \overbrace{(UV - VU + \dot{V})}^{=:C} \\ &= Y_h(U + hC) - h^2 XVC \\ &= (X + hXV)(U + hC) - h^2 XVC \\ &= XU + hXVU + hXC + h^2 XVC - h^2 XVC \\ &= \dot{X} + hXVU + hXUV - hXVU + hX\dot{V} \\ &= \dot{X} + hXUV + hX\dot{V} = \dot{X} + h\dot{X}V + hX\dot{V} = \dot{Y}_h. \end{aligned} \quad (2.14)$$

We therefore obtain

$$\begin{aligned}
U_h &= Y_h^{-1} \dot{Y}_h \stackrel{(2.14)}{=} Y_h^{-1} (Y_h(U + hC) - h^2 XVC) = Y_h^{-1} (Y_h(U + hC) - h^2 XVC) \\
&= U + hC - h^2 Y_h^{-1} XVC \\
&= U + h(UV - VU + \dot{V}) - h^2 Y_h^{-1} XVC, \quad (2.15)
\end{aligned}$$

implying

$$U_h|_{h=0} = U, \quad \frac{d}{dh} U_h|_{h=0} = UV - VU + \dot{V}. \quad (2.16)$$

Coming back to equation (2.13), we find that

$$\begin{aligned}
0 &= \frac{1}{2} \frac{d}{dh} E(Y_h) \Big|_{h=0} = \frac{1}{2} \int_0^1 \frac{d}{dh} \langle Y_h^{-1} \dot{Y}_h, Y_h^{-1} \dot{Y}_h \rangle_{\mu, \mu_c, \kappa} \Big|_{h=0} dt \\
&= \frac{1}{2} \int_0^1 \frac{d}{dh} \langle U_h, U_h \rangle_{\mu, \mu_c, \kappa} \Big|_{h=0} dt \\
&= \int_0^1 \langle U_h, \frac{d}{dh} U_h \rangle_{\mu, \mu_c, \kappa} \Big|_{h=0} dt \stackrel{(2.16)}{=} \int_0^1 \langle U, UV - VU + \dot{V} \rangle_{\mu, \mu_c, \kappa} dt. \quad (2.17)
\end{aligned}$$

Using Lemma 2.3, we can write (2.17) as

$$\begin{aligned}
&\int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \langle U, UV - VU \rangle_{\mu, \mu_c, \kappa} dt \\
&= \int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2} \langle U, UV - VU \rangle dt \\
&= \int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2} \langle V, \underbrace{U^T U - U U^T}_{\in \text{Sym}(n)} \rangle dt \\
&= \int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} (\mu \langle \text{dev sym } V, \text{dev sym}(U^T U - U U^T) \rangle \\
&\quad + \underbrace{\mu_c \langle \text{skew } V, \text{skew}(U^T U - U U^T) \rangle + \frac{\kappa}{n} \text{tr } V \text{tr}(U^T U - U U^T)}_{=0}) dt \\
&= \int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt \\
&= \sum_{k=0}^m \int_{a_k}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt. \quad (2.18)
\end{aligned}$$

For $V \in C_0^1([0, 1]; \mathfrak{gl}(n))$ with $V(a_k) = 0$ for $k \in \{1, \dots, m\}$, integration by parts yields

$$\begin{aligned}
0 &= \sum_{k=0}^m \int_{a_k}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt \\
&= \sum_{k=0}^m \int_{a_k}^{a_{k+1}} -\langle V, \dot{U} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt. \quad (2.19)
\end{aligned}$$

We apply the fundamental lemma of the calculus of variations to find

$$\dot{U}(t) = \frac{1 + \frac{\mu_c}{\mu}}{2} (U(t)^T U(t) - U(t) U(t)^T) \quad \forall t \in [0, 1] \setminus \{a_1, \dots, a_m\}. \quad (2.20)$$

To show that X is two-times differentiable on $[0, 1]$ and equation (2.20) is fulfilled everywhere, we first show that $U = X^{-1}\dot{X}$ is continuous at each a_k (and therefore on the whole interval $[0, 1]$). We choose $V \in C_0^1((a_{k-1}, a_{k+1}); \mathfrak{gl}(n))$ for $k \in \{1, \dots, m\}$ and compute

$$\begin{aligned}
0 &= \int_0^1 \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt \\
&= \int_{a_{k-1}}^{a_k} \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt \\
&\quad + \int_{a_{k-1}}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt \\
&= \int_{a_{k-1}}^{a_k} -\langle V, \dot{U} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt + [\langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa}]_{t=a_{k-1}}^{a_k} \\
&\quad + \int_{a_k}^{a_{k+1}} -\langle V, \dot{U} \rangle_{\mu, \mu_c, \kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} dt + [\langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa}]_{t=a_k}^{a_{k+1}}. \tag{2.21}
\end{aligned}$$

According to (2.20) the equality

$$\dot{U}(t) = \frac{\mu + \mu_c}{2\mu} (U(t)^T U(t) - U(t) U(t)^T)$$

holds on $[a_{k-1}, a_k]$ as well as $[a_k, a_{k+1}]$. Thus the integrals in (2.21) are zero and we find

$$\begin{aligned}
0 &= [\langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa}]_{t=a_{k-1}}^{a_k} + [\langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa}]_{t=a_k}^{a_{k+1}} \\
&= \lim_{t \nearrow a_k} \langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa} - \lim_{t \searrow a_k} \langle V(t), U(t) \rangle_{\mu, \mu_c, \kappa}. \tag{2.22}
\end{aligned}$$

Since V can be chosen with arbitrary values for $V(a_k)$, we find

$$\lim_{t \nearrow a_k} U(t) = \lim_{t \searrow a_k} U(t). \tag{2.23}$$

Therefore U is continuous on $[0, 1]$. But then (2.20) implies that \dot{U} is continuous as well. Thus $\dot{X} = XU$ is continuous on the whole interval, and therefore X is continuously differentiable. But then $\dot{X} = XU$ is continuously differentiable as well, and thus $X \in C^2([0, 1]; \text{GL}^+(n))$ and $U = X^{-1}\dot{X}$ fulfils the *geodesic equation*

$$\dot{U} = \frac{1 + \frac{\mu_c}{\mu}}{2} (U^T U - U U^T) \tag{2.24}$$

everywhere on $[0, 1]$. The results of section 2.2 can be summarized as follows:

Proposition 2.4. *Let $A, B \in \text{GL}^+(n)$ and let $X \in \mathcal{A}_A^B$ be a piecewise two-times differentiable energy minimizer with regard to the left-invariant Riemannian metric induced by the inner product $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$. Then $X \in C^2([a, b]; \text{GL}^+(n))$, and $U = X^{-1}\dot{X}$ is a solution to (2.24).*

Remark 2.5. In particular, every length minimizing curve can be reparameterized to an energy minimizer, according to Lemmas 1.5 and 1.8. Thus every *length minimizer* has a reparameterization that fulfils the geodesic equation.

2.3. Properties of solutions U

We will now state some properties for every solution $U \in C^1([0, 1]; \mathfrak{gl}(n))$ to equation (2.24).

Lemma 2.6. *Let U be a solution to (2.24). Then:*

$$i) \quad \|U\|_{\mu, \mu_c, \kappa} \equiv \text{constant}, \quad (2.25)$$

$$ii) \quad \text{tr}(U) \equiv \text{constant}, \quad (2.26)$$

$$iii) \quad \det(U) \equiv \text{constant}, \quad (2.27)$$

$$iv) \quad \text{tr}(\text{Cof } U) \equiv \text{constant}, \quad (2.28)$$

where $\text{Cof } U$ denotes the cofactor of U .

Proof.

$$\begin{aligned} i) \quad \frac{d}{dt} \|U\|_{\mu, \mu_c, \kappa}^2 &= \frac{d}{dt} \langle U, U \rangle_{\mu, \mu_c, \kappa} = 2 \langle U, \dot{U} \rangle_{\mu, \mu_c, \kappa} = (1 + \frac{\mu_c}{\mu}) \langle U, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} \\ &= (1 + \frac{\mu_c}{\mu}) \left(\mu \langle \text{sym } U, U^T U - U U^T \rangle_{\mu, \mu_c, \kappa} + \mu_c \langle \text{skew } U, \underbrace{U^T U - U U^T}_{\in \text{Sym}(n)} \rangle_{\mu, \mu_c, \kappa} \right. \\ &\quad \left. + \frac{\kappa}{n} \text{tr } U \underbrace{\text{tr}(U^T U - U U^T)}_{=0} \right) \\ &= (\mu + \mu_c) \langle \frac{1}{2}(U + U^T), U^T U - U U^T \rangle \\ &= \frac{\mu + \mu_c}{2} \text{tr}(U^T U U - U U^T U + U^T U^T U - U^T U U^T) \\ &= \frac{\mu + \mu_c}{2} \text{tr}(U^T U U) - \text{tr}(U U^T U) + \text{tr}(U^T U^T U) - \text{tr}(U^T U U^T) = 0, \\ ii) \quad \frac{d}{dt} \text{tr } U &= \text{tr } \dot{U} = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^T U - U U^T) = 0, \\ iii) \quad \frac{d}{dt} \det U &= \text{tr}(\text{Adj}(U) \dot{U}) = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\text{Adj}(U)(U^T U - U U^T)) \\ &= \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\text{Adj}(U) U^T U - \text{Adj}(U) U U^T) \\ &= \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^T U \text{Adj}(U) - \det(U) U^T) = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\det(U) U^T - \det(U) U^T) = 0, \\ iv) \quad \frac{d}{dt} \text{tr } U^{-1} &= -\text{tr}(U^{-1} \dot{U} U^{-1}) \\ &= -\frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^{-1}(U^T U - U U^T) U^{-1}) = -\frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^{-1} U^T - U^T U^{-1}) = 0 \\ \implies \text{tr } \text{Cof}(U) &= \det(U) \cdot \text{tr } U^{-T} = \det(U) \cdot \text{tr } U^{-1} \equiv \text{constant}, \end{aligned}$$

where $\text{Adj}(A) = \text{Cof } A^T$ is the adjugate matrix of A with $\text{Adj}(A) \cdot A = A \cdot \text{Adj}(A) = (\det A) \mathbb{1}$. For the derivative of the determinant function used in equality (2.29), see e.g. [20]. \blacksquare

2.4. Existence and uniqueness of solutions

Since the function $U \mapsto \frac{1 + \frac{\mu_c}{\mu}}{2}(U^T U - U U^T)$ is locally Lipschitz continuous, the Picard–Lindelöf theorem implies that equation (2.24) has a unique local solution for given $U(0)$ and that if a global solution exists, it is uniquely defined. Then for given $U \in C^1(I, \mathfrak{gl}(n))$, the initial value problem

$$\begin{cases} \dot{X} = XU, \\ X(0) = X_0 \end{cases} \quad (2.29)$$

has a unique solution on the interval I as well. Using estimates (2.25) to (2.27), we can even show that this solution is globally defined:

Proposition 2.7. For $X_0 \in \text{GL}^+(n)$, $M_0 \in \mathfrak{gl}(n)$, the initial value problem

$$\begin{cases} U = X^{-1}\dot{X}, \\ \dot{U} = \frac{1 + \frac{\mu_c}{\mu}}{2}(U^T U - U U^T), \\ X(0) = X_0, \quad \dot{X}(0) = M_0 \end{cases} \quad (2.30)$$

for the geodesic equation has a unique global solution $X : \mathbb{R} \rightarrow \text{GL}^+(n)$.

Proof. Since $U(0) = X(0)^{-1}\dot{X}(0)$ is determined by $X(0)$ and $\dot{X}(0)$, (2.30) has a unique local solution, as demonstrated above. Then, according to (2.25), $\|U\|_{\mu, \mu_c, \kappa}$ is constant, hence U is bounded in $\mathbb{R}^{n \times n}$ and therefore defined on all of \mathbb{R} . The same estimate shows that the function $X \mapsto UX$ is globally Lipschitz continuous. Therefore there exists a unique global solution $X : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ to 2.30. It remains to show that $X(t) \in \text{GL}^+(n)$ for all $t \in \mathbb{R}$. First we observe that, for $\det X > 0$,

$$\frac{d}{dt}(\det X) = \text{tr}(\text{Adj}(X)\dot{X}) = \text{tr}(\det(X)X^{-1}\dot{X}) = \det(X) \text{tr} U \stackrel{(2.26)}{=} \det(X) \text{tr}(U(0)). \quad (2.31)$$

Then, on every interval $I \subseteq \mathbb{R}$ with $0 \in I$ and $X(I) \subseteq \text{GL}^+(n)$, the function $t \mapsto \det X(t)$ solves the initial value problem

$$\begin{cases} \frac{d}{dt}(\det X(t)) = \text{tr}[U(0)] \det X(t), \\ \det X(0) = \det X_0. \end{cases} \quad (2.32)$$

Assume now that $X(\mathbb{R}) \not\subseteq \text{GL}^+(n)$. Since $I^+ := \{t \in \mathbb{R} : \det X(t) > 0\}$ is open and $0 \in I^+$ by definition of X_0 , we can find the minimal $t_0 > 0$ with $-t_0 \notin I^+$ or $t_0 \notin I^+$. But then (2.32) yields

$$\det X(t) = (\det X_0) e^{t \cdot \text{tr}(U(0))} \quad (2.33)$$

on $(-t_0, t_0)$, and thus

$$0 < \lim_{t \searrow -t_0} \det X(t) = \det X(-t_0), \quad 0 < \lim_{t \nearrow t_0} \det X(t) = \det X(t_0),$$

in contradiction to the definition of t_0 . ■

3. The parametrization of geodesics

3.1. Solving the geodesic equation

We will now give an explicit solution to the geodesic initial value problem (2.30). The result is inspired by the development of Mielke [17], who obtained formula (3.3) as a parametrization of the geodesics on $\text{SL}(n)$. For the canonical inner product, i.e. for the special case $\mu = \mu_c = \kappa = 1$, the result can also be found in [1]. Note that while geodesics on a Lie group equipped with a bi-invariant Riemannian metric are translated one-parameter groups which can easily be characterized through the matrix exponential [24] (which, for example, allows for an easy explicit computation of the geodesic distance on the special orthogonal group $\text{SO}(n)$ with respect to the canonical metric [18]), the metric discussed here is not bi- $\text{GL}(n)$ -invariant and thus the solution to the geodesic equation takes on a more complicated form.

Let $A \in \text{GL}^+(n)$, $M \in \mathfrak{gl}(n)$ and $\omega = \frac{\mu_c}{\mu}$. We define

$$\Phi(M) := \exp(\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M), \quad (3.1)$$

$$\Phi_A(M) := A \cdot \Phi(M) = A \exp(\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M), \quad (3.2)$$

$$X(t) := \Phi_A(tM) = A \exp(t(\text{sym } M - \omega \text{ skew } M)) \exp(t(1 + \omega) \text{ skew } M), \quad (3.3)$$

where $\exp : \mathfrak{gl}(n) \rightarrow \text{GL}^+(n)$ denotes the matrix exponential (see Appendix A.1 for further information). We use equation (A.12) to compute

$$\begin{aligned}\dot{X} &= A \exp(t(\text{sym } M - \omega \text{ skew } M)) (\text{sym } M - \omega \text{ skew } M) \exp(t(1 + \omega) \text{ skew } M) \\ &\quad + A \exp(t(\text{sym } M - \omega \text{ skew } M)) (1 + \omega) \text{ skew } M \exp(t(1 + \omega) \text{ skew } M) \\ &= A \exp(t(\text{sym } M - \omega \text{ skew } M)) (\text{sym } M + \text{skew } M) \exp(t(1 + \omega) \text{ skew } M) \\ &= A \exp(t(\text{sym } M - \omega \text{ skew } M)) M \exp(t(1 + \omega) \text{ skew } M)\end{aligned}\tag{3.4}$$

and

$$\begin{aligned}U &= X^{-1} \dot{X} = \exp(t(1 + \omega) \text{ skew } M)^{-1} \exp(t(\text{sym } M - \omega \text{ skew } M))^{-1} A^{-1} \\ &\quad \cdot A \exp(t(\text{sym } M - \omega \text{ skew } M)) M \exp(t(1 + \omega) \text{ skew } M) \\ &= \exp(t(1 + \omega) \text{ skew } M)^{-1} M \underbrace{\exp(t(1 + \omega) \text{ skew } M)}_{\substack{\in \mathfrak{so}(n) \\ =: Q \in \text{SO}(n)}} = Q^T M Q,\end{aligned}\tag{3.5}$$

where we used (A.8) stated in the appendix to infer $Q \in \text{SO}(n)$. Therefore

$$U^T U - U U^T = Q^T M^T M Q - Q^T M M^T Q = Q^T (M^T M - M M^T) Q,\tag{3.6}$$

as well as

$$\begin{aligned}\dot{U} &= \dot{Q}^T M Q + Q^T M \dot{Q} \stackrel{(A.12)}{=} Q^T (1 + \omega) (\text{skew } M)^T M Q + Q^T (1 + \omega) M (\text{skew } M) Q \\ &= (1 + \omega) Q^T ((\text{skew } M)^T M + M (\text{skew } M)) Q \\ &= (1 + \omega) Q^T (M (\text{skew } M) - (\text{skew } M) M) Q \\ &= (1 + \omega) Q^T \left(\frac{1}{2} (M M - M M^T) - \frac{1}{2} (M M - M^T M) \right) Q \\ &= \frac{(1 + \omega)}{2} Q^T (M^T M - M M^T) Q = \frac{(1 + \omega)}{2} (U^T U - U U^T),\end{aligned}\tag{3.7}$$

and thus X solves the geodesic equation. To solve the initial value problem (2.30) we use the equality $\exp(0) = \mathbb{I}$ as well as equation (3.4) to obtain

$$X(0) = A, \quad \dot{X}(0) = A M,\tag{3.8}$$

hence we can solve (2.30) by choosing $A = X_0$, $M = X_0^{-1} M_0$. We conclude:

Theorem 3.1. *Let $X_0 \in \text{GL}^+(n)$, $M_0 \in \mathfrak{gl}(n)$. Then the curve $X : \mathbb{R} \rightarrow \text{GL}^+(n)$ with*

$$X(t) = \Phi_{X_0}(t X_0^{-1} M_0) = X_0 \exp(t(\text{sym}(X_0^{-1} M_0) - \frac{\mu_c}{\mu} (X_0^{-1} M_0))) \exp(t(1 + \frac{\mu_c}{\mu}) \text{skew}(X_0^{-1} M_0))$$

is the unique global solution to the geodesic initial value problem

$$\begin{cases} U = X^{-1} \dot{X}, \\ \dot{U} = \frac{1 + \frac{\mu_c}{\mu}}{2} (U^T U - U U^T), \\ X(0) = X_0, \quad \dot{X}(0) = M_0. \end{cases} \quad \blacksquare$$

Since the left invariance of g implies

$$\text{dist}_{\text{geod}}(A, B) = \text{dist}_{\text{geod}}(\mathbb{I}, A^{-1} B),$$

the distance between two matrices can be computed as the length of a geodesic starting in \mathbb{I} ; we can therefore mostly focus on the case $A = \mathbb{I}$, $\Phi_A = \Phi_{\mathbb{I}} = \Phi$ and

$$X(t) = \exp(t(\text{sym } M - \omega \text{ skew } M)) \exp(t(1 + \omega) \text{ skew } M).$$

3.2. Properties of geodesic curves

Since the geodesic curves

$$X(t) = A \exp(t(\text{sym } M - \omega \text{ skew } M)) \exp(t(1 + \omega) \text{ skew } M) \quad (3.9)$$

solve the geodesic initial value problem (2.30), the properties given in Lemma 2.6 can be directly applied to $U = X^{-1}\dot{X}$. Furthermore, we can compute the length of X on the interval $[0, t_0]$ for given $M \in \mathfrak{gl}(n)$ (recall from (3.5) that $X^{-1}\dot{X} = U = Q^T M Q$ with $Q(t) \in \text{SO}(n)$):

$$\begin{aligned} L(X) &= \int_0^{t_0} g_X(\dot{X}, \dot{X}) \, dt = \int_0^{t_0} \sqrt{\langle X^{-1}\dot{X}, X^{-1}\dot{X} \rangle_{\mu, \mu_c, \kappa}} \, dt \\ &= \int_0^{t_0} \sqrt{\langle U, U \rangle_{\mu, \mu_c, \kappa}} \, dt = \int_0^{t_0} \sqrt{\langle Q^T M Q, Q^T M Q \rangle_{\mu, \mu_c, \kappa}} \, dt \\ &= \int_0^{t_0} \sqrt{\langle M, M \rangle_{\mu, \mu_c, \kappa}} \, dt = t_0 \|M\|_{\mu, \mu_c, \kappa}, \end{aligned} \quad (3.10)$$

where the second to last equality follows from the isotropic property (1.39) of the inner product $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$.

4. The existence of length minimizers

As we have seen in section 2 (Remark 2.5), every sufficiently smooth length minimizing curve can be reparameterized to solve the geodesic equation. Thus Theorem 3.1 shows that every length minimizer, after possible reparameterization, is of the form (3.3).

However, it is not clear that, for given $A, B \in \text{GL}^+(n)$, such a minimizer connecting A and B actually exists. In the broader setting of differential geometry it can be shown that the *local* existence of minimizers is guaranteed on any (differentiable) Riemannian manifold [12, Corollary 1.4.2]. In our effort to keep this paper self contained we will mimic the proof of local existence given in [8] with regard to our special case. While the *global* existence is not obvious on the non-compact set $\text{GL}^+(n)$, it will also be shown later on.

4.1. Preparations

In the context of Riemannian geometry the function $\Phi_A : \mathfrak{gl}(n) \rightarrow \text{GL}^+(n)$ with

$$\Phi_A(M) = A \exp(\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M) \quad (4.1)$$

can be considered the *geodesic exponential* at A [12, Definition 1.4.3]. From this point of view the following lemma is an application of *Gauss's lemma* [7, 2.93], although here we will prove it through direct computation.

Lemma 4.1. *Let $A \in \text{GL}^+(n)$, $M, T \in \mathfrak{gl}(n)$. Then*

$$g_{\Phi_A(M)}(D\Phi_A[M].M, D\Phi_A[M].T) = \langle M, T \rangle_{\mu, \mu_c, \kappa}. \quad (4.2)$$

Remark 4.2. Note carefully that M occurs multiple times in (4.2), including the direction of one of the derivatives. This is a necessary restriction; in general, the equality $g_{\Phi_A(M)}(D\Phi_A[M].S, D\Phi_A[M].T) = \langle S, T \rangle_{\mu, \mu_c, \kappa}$ does *not* hold for arbitrary $S, T \in \mathfrak{gl}(n)$.

Proof. First note that the left invariance of the Riemannian metric g implies

$$\begin{aligned} g_{\Phi_A(M)}(D\Phi_A[M].M, D\Phi_A[M].T) &= g_{A\Phi(M)}(AD\Phi[M].M, AD\Phi[M].T) \\ &= g_{\Phi(M)}(D\Phi[M].M, D\Phi[M].T). \end{aligned} \quad (4.3)$$

Using (A.12), we compute

$$\begin{aligned}
D\Phi[M].M &= (D \exp[\text{sym } M - \omega \text{ skew } M].(\text{sym } M - \omega \text{ skew } M)) \exp((1 + \omega) \text{ skew } M) \\
&\quad + \exp(\text{sym } M - \omega \text{ skew } M) (D \exp[(1 + \omega) \text{ skew } M].((1 + \omega) \text{ skew } M)) \\
&= \exp(\text{sym } M - \omega \text{ skew } M) (\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M) \\
&\quad + \exp(\text{sym } M - \omega \text{ skew } M) ((1 + \omega) \text{ skew } M) \exp((1 + \omega) \text{ skew } M) \\
&= \exp(\text{sym } M - \omega \text{ skew } M) (\text{sym } M + \text{skew } M) \exp((1 + \omega) \text{ skew } M) \\
&= \exp(\text{sym } M - \omega \text{ skew } M) M \underbrace{\exp((1 + \omega) \text{ skew } M)}_{=:Q \in \text{SO}(n)} \\
&= \exp(\text{sym } M - \omega \text{ skew } M) M Q, \tag{4.4}
\end{aligned}$$

while (A.11) yields

$$\begin{aligned}
D\Phi[M].T &= (D \exp[\text{sym } M - \omega \text{ skew } M].(\text{sym } T - \omega \text{ skew } T)) \exp((1 + \omega) \text{ skew } M) \\
&\quad + \exp(\text{sym } M - \omega \text{ skew } M) (D \exp[(1 + \omega) \text{ skew } M].(1 + \omega) \text{ skew } T) \\
&= \int_0^1 \exp(s(\text{sym } M - \omega \text{ skew } M)) (\text{sym } T - \omega \text{ skew } T) \cdot \exp((1 - s)(\text{sym } M - \omega \text{ skew } M)) \exp((1 + \omega) \text{ skew } M) ds \\
&\quad + \int_0^1 \exp(\text{sym } M - \omega \text{ skew } M) \exp(s(1 + \omega) \text{ skew } M) ((1 + \omega) \text{ skew } T) \cdot \exp((1 - s)(1 + \omega) \text{ skew } M) ds \\
&= \exp(\text{sym } M - \omega \text{ skew } M) \int_0^1 \underbrace{[\exp((s - 1)(\text{sym } M - \omega \text{ skew } M)) \cdot (\text{sym } T - \omega \text{ skew } T) \exp((1 - s)(\text{sym } M - \omega \text{ skew } M)) \\
&\quad + \exp(s(1 + \omega) \text{ skew } M) ((1 + \omega) \text{ skew } T) \exp(-s(1 + \omega) \text{ skew } M)]}_{=:R_s \in \text{SO}(n)} ds \cdot \exp((1 + \omega) \text{ skew } M) \\
&= \exp(\text{sym } M - \omega \text{ skew } M) \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s] ds \cdot \exp((1 + \omega) \text{ skew } M). \tag{4.5}
\end{aligned}$$

We also find

$$\Phi(M)^{-1} = \underbrace{\exp((1 + \omega) \text{ skew } M)^{-1}}_{=:Q^T} \exp(\text{sym } M - \omega \text{ skew } M)^{-1}. \tag{4.6}$$

Hence we can compute

$$\begin{aligned}
g_{\Phi(M)}(D\Phi[M].M, D\Phi[M].T) &= \langle \Phi(M)^{-1} D\Phi[M].M, \Phi(M)^{-1} D\Phi[M].T \rangle_{\mu, \mu_c, \kappa} \\
&= \langle Q^T M Q, Q^T \cdot \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s] ds \cdot Q \rangle_{\mu, \mu_c, \kappa} \\
&= \langle M, \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s] ds \rangle_{\mu, \mu_c, \kappa} \\
&= \int_0^1 \langle M, P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s \rangle_{\mu, \mu_c, \kappa} ds \\
&= \int_0^1 \langle \mu \text{ dev sym } M + \mu_c \text{ skew } M, P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} \rangle \\
&\quad + \frac{\kappa}{n} \text{tr } M \text{tr}(P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1}) + (1 + \omega) \langle M, \underbrace{R_s^T(\text{skew } T)R_s}_{\in \text{so}(n)} \rangle_{\mu, \mu_c, \kappa} ds \\
&= \int_0^1 \langle P_s^T(\mu \text{ dev sym } M + \mu_c \text{ skew } M)P_s^{-T}, \text{sym } T - \omega \text{ skew } T \rangle \\
&\quad + \frac{\kappa}{n} \text{tr } M \text{tr}(\text{sym } T - \omega \text{ skew } T) + (1 + \omega) \mu_c \langle \text{skew } M, R_s^T \text{skew } T R_s \rangle ds \\
&= \int_0^1 \mu \langle P_s^T(\text{dev sym } M + \omega \text{ skew } M)P_s^{-T}, \text{sym } T - \omega \text{ skew } T \rangle \\
&\quad + \frac{\kappa}{n} \text{tr } M \text{tr } T + (1 + \omega) \mu_c \langle R_s \text{skew } M R_s^T, \text{skew } T \rangle ds. \tag{4.7}
\end{aligned}$$

It is not difficult to see that the matrices

$$\text{dev sym } M + \omega \text{ skew } M, \quad \text{sym } M + \omega \text{ skew } M \tag{4.8}$$

commute. Therefore, because $P_s^T = \exp((s-1)(\text{sym } M + \omega \text{ skew } M))$, we can use equation (A.6) to infer that $\text{dev } \text{sym } M + \omega \text{ skew } M$ and P_s^T commute. Analogously, we see that $\text{skew } M$ and $R_s = \exp(\exp((1+\omega) \text{ skew } M))$ commute as well. Thus expression (4.7) can be simplified to

$$\begin{aligned}
& \int_0^1 \mu \langle (\text{dev } \text{sym } M + \omega \text{ skew } M) P_s^T P_s^{-T}, \text{sym } T - \omega \text{ skew } T \rangle \\
& + \frac{\kappa}{n} \text{tr } M \text{tr } T + (1+\omega) \mu_c \langle \text{skew } M R_s R_s^T, \text{skew } T \rangle \text{ ds} \\
& = \int_0^1 \mu \langle \text{dev } \text{sym } M + \omega \text{ skew } M, \text{sym } T - \omega \text{ skew } T \rangle \\
& \quad + \frac{\kappa}{n} \text{tr } M \text{tr } T + (1+\omega) \mu_c \langle \text{skew } M, \text{skew } T \rangle \text{ ds} \\
& = \int_0^1 \mu \langle \text{dev } \text{sym } M, \text{dev } \text{sym } T \rangle + \mu_c \langle \text{skew } M, -\omega \text{ skew } T \rangle \\
& \quad + \frac{\kappa}{n} \text{tr } M \text{tr } T + (\mu_c + \omega \mu_c) \langle \text{skew } M, \text{skew } T \rangle \text{ ds} \\
& = \int_0^1 \mu \langle \text{dev } \text{sym } M, \text{dev } \text{sym } T \rangle + \mu_c \langle \text{skew } M, \text{skew } T \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } T \text{ ds} = \langle M, T \rangle_{\mu, \mu_c, \kappa}. \blacksquare
\end{aligned} \tag{4.9}$$

This lemma can now be used to establish a lower bound for the length of curves “close to A ”. To do so, we choose $\varepsilon > 0$ such that Φ_A is a diffeomorphism from $B_\varepsilon(0) \subseteq \mathfrak{gl}(n)$ to an open neighbourhood of A in $\text{GL}^+(n)$; note that, according to (3.8), $D\Phi_A[0]$ is surjective and therefore nonsingular. Then any “short enough” curve Y with $Y(0) = A$ can be represented as $Y = \Phi_A \circ \gamma$ with a curve γ in $B_\varepsilon(0) \subseteq \mathfrak{gl}(n)$. To compute the length of Y , we need the following lemma, which is a modified version of Proposition 5.3.2 in [5]:

Lemma 4.3. *Let $\gamma \in C^0([0, 1]; B_\varepsilon(0))$ be a piecewise differentiable curve in $B_\varepsilon(0) \subseteq \mathfrak{gl}(n)$ with $\gamma(0) = 0$, $\gamma(1) = M$. Then:*

$$L(\Phi_A \circ \gamma) \geq \|M\|_{\mu, \mu_c, \kappa}, \tag{4.10}$$

$$E(\Phi_A \circ \gamma) \geq \|M\|_{\mu, \mu_c, \kappa}^2. \tag{4.11}$$

Proof. Let such a curve γ be given. Then the length of $\Phi_A \circ \gamma$ is

$$\begin{aligned}
L(\Phi_A \circ \gamma) &= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}\left(\frac{d}{dt}(\Phi_A \circ \gamma)(t), \frac{d}{dt}(\Phi_A \circ \gamma)(t)\right)} dt \\
&= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \dot{\gamma}, D\Phi_A[\gamma] \cdot \dot{\gamma})} dt.
\end{aligned} \tag{4.12}$$

Note that since γ and $\dot{\gamma}$ are generally not equal, the previous lemma can not be applied directly. Therefore we decompose $\dot{\gamma}$ into the sum of γ_1 and γ_2 , where γ_1 is a multiple of γ and γ_2 is orthogonal to γ with regard to the inner product on $\mathfrak{gl}(n)$: define r, γ_1, γ_2 by

$$r(t) := \begin{cases} \frac{\langle \gamma, \dot{\gamma} \rangle_{\mu, \mu_c, \kappa}}{\|\gamma(t)\|_{\mu, \mu_c, \kappa}^2} & : \gamma(t) \neq 0 \\ 0 & : \gamma(t) = 0 \end{cases}, \tag{4.13}$$

$$\gamma_1(t) := r(t) \gamma(t),$$

$$\gamma_2(t) := \dot{\gamma}(t) - \gamma_1(t).$$

Then

$$\dot{\gamma} = \gamma_1 + \gamma_2, \quad \langle \gamma, \gamma_2 \rangle_{\mu, \mu_c, \kappa} = 0 \tag{4.14}$$

and thus

$$\begin{aligned}
L(\Phi_A \circ \gamma) &= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \dot{\gamma}, D\Phi_A[\gamma] \cdot \dot{\gamma})} dt \\
&= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot (\gamma_1 + \gamma_2), D\Phi_A[\gamma] \cdot (\gamma_1 + \gamma_2))} dt \\
&= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma_1 + D\Phi_A[\gamma] \cdot \gamma_2, D\Phi_A[\gamma] \cdot \gamma_1 + D\Phi_A[\gamma] \cdot \gamma_2)} dt \\
&\geq \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma_1, D\Phi_A[\gamma] \cdot \gamma_1) + 2g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma_1, D\Phi_A[\gamma] \cdot \gamma_2)} dt \quad (4.15) \\
&= \int_0^1 \sqrt{r(t)^2 g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma, D\Phi_A[\gamma] \cdot \gamma) + 2r(t)g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma, D\Phi_A[\gamma] \cdot \gamma_2)} dt \\
&\stackrel{(4.2)}{=} \int_0^1 \sqrt{r(t)^2 \langle \gamma, \gamma \rangle_{\mu, \mu_c, \kappa} + 2r(t) \underbrace{\langle \gamma, \gamma_2 \rangle_{\mu, \mu_c, \kappa}}_{=0}} dt.
\end{aligned}$$

Without loss of generality we assume $M \neq 0$. Since γ is continuous and $\gamma(1) = M$, we can choose $a := \max\{s \in [0, 1] : \gamma(s) = 0\}$ such that $\gamma(t) \neq 0$ for all $t > a$. We obtain

$$\begin{aligned}
L(\Phi \circ \gamma) &= \int_0^1 \sqrt{r(t)^2 \langle \gamma, \gamma \rangle_{\mu, \mu_c, \kappa}} dt = \int_0^1 r(t) \|\gamma\|_{\mu, \mu_c, \kappa} dt \quad (4.16) \\
&\geq \int_a^1 r(t) \|\gamma\|_{\mu, \mu_c, \kappa} dt = \int_a^1 \frac{\langle \gamma, \dot{\gamma} \rangle_{\mu, \mu_c, \kappa}}{\|\gamma\|_{\mu, \mu_c, \kappa}} dt = \int_a^1 \frac{d}{dt} \|\gamma(t)\|_{\mu, \mu_c, \kappa} dt.
\end{aligned}$$

The curve γ is piecewise differentiable by assumption, so choose $a = a_1 < \dots < a_n < a_{n+1} = 1$ such that γ is continuously differentiable on $(a, 1) \setminus \{a_1, \dots, a_n\}$.

We compute

$$\begin{aligned}
L(\Phi \circ \gamma) &\geq \int_a^1 \frac{d}{dt} \|\gamma(t)\|_{\mu, \mu_c, \kappa} dt = \sum_{i=1}^m \int_{a_i}^{a_{i+1}} \frac{d}{dt} \|\gamma(t)\|_{\mu, \mu_c, \kappa} dt = \sum_{i=1}^m (\|\gamma(a_{i+1})\|_{\mu, \mu_c, \kappa} - \|\gamma(a_i)\|_{\mu, \mu_c, \kappa}) \\
&= \|\gamma(a_{m+1})\|_{\mu, \mu_c, \kappa} - \|\gamma(a)\|_{\mu, \mu_c, \kappa} = \|\gamma(1)\|_{\mu, \mu_c, \kappa} - 0 = \|M\|_{\mu, \mu_c, \kappa}. \quad (4.17)
\end{aligned}$$

Finally, Lemma 1.7 yields

$$E(\Phi \circ \gamma) \geq L(\Phi \circ \gamma)^2 \geq \|M\|_{\mu, \mu_c, \kappa}^2. \quad \blacksquare$$

We can now give a lower bound for the length of a curve $Z \in \mathcal{A}_A^B : [0, 1] \rightarrow \text{GL}^+(n)$ with $B \notin \Phi_A(B_\varepsilon(0))$: with $t_0 = \min\{t \in [0, 1] : Z(t) \notin \Phi_A(B_\varepsilon(0))\}$, we can write $Z(t) = \Phi_A(\gamma(t))$ for $t \in [0, t_0)$, and $Z(t_0) \notin \Phi_A(B_\varepsilon(0))$ implies

$$\lim_{t \nearrow t_0} \|\gamma(t)\|_{\mu, \mu_c, \kappa} = \varepsilon. \quad (4.18)$$

Thus, using (4.10), we find $L(Z) \geq \varepsilon$ and therefore

$$\text{dist}_{\text{geod}}(A, B) > \varepsilon \quad \forall B \notin \Phi_A(B_\varepsilon(0)). \quad (4.19)$$

Furthermore, for $\|M\|_{\mu, \mu_c, \kappa} < \varepsilon$, we can directly compute the distance between A and $B = \Phi_A(M) \in B_\varepsilon(A)$: For the geodesic curve $X \in \mathcal{A}_A^B : [0, 1] \rightarrow \text{GL}^+(n)$ with

$$X(t) = \Phi_A(tM) \quad (4.20)$$

we find $X(0) = A, X(1) = \Phi_A(M) = B$ and equation (3.10) yields $L(X) = \|M\|_{\mu, \mu_c, \kappa}$; thus $\text{dist}_{\text{geod}}(A, B) \leq \|M\|_{\mu, \mu_c, \kappa}$. For another curve $Y \in \mathcal{A}_A^B : [0, 1] \rightarrow \text{GL}^+(n)$, we find that Y is either contained in

$B_\varepsilon(A)$, in which case (4.10) implies $L(Y) \geq \|M\|_{\mu, \mu_c, \kappa}$; or Y leaves $B_\varepsilon(A)$. But then, as shown above, $L(Y) \geq \varepsilon > \|M\|_{\mu, \mu_c, \kappa}$, and therefore

$$\text{dist}_{\text{geod}}(A, B) = \|M\|_{\mu, \mu_c, \kappa}, \quad (4.21)$$

and X is a length minimizer. Finally these positive lower bounds imply

$$\text{dist}_{\text{geod}}(A, N) > 0 \quad (4.22)$$

for all $B \neq A$. Therefore $\text{dist}_{\text{geod}}$ defines a *metric* on $\text{GL}^+(n)$. Note also that equality in (4.15) holds if and only if

$$g_{\Phi_A(\gamma)}(D\Phi_A[\gamma] \cdot \gamma_2, D\Phi_A[\gamma] \cdot \gamma_2) = 0 \quad (4.23)$$

on $[0, 1]$. Since $g_{\Phi_A(\gamma)}$ is positive definite, this is equivalent to $0 = D\Phi_A[\gamma] \cdot \gamma_2 = AD\Phi[\gamma] \cdot \gamma_2$, and since $D\Phi$ is non-singular in a neighbourhood of 0, this equality holds (for small enough ε) if and only if

$$0 = \gamma_2(t) = \dot{\gamma}(t) - \gamma_1(t) = \dot{\gamma}(t) - r(t) \gamma(t)$$

for all $t \in [0, 1]$. But then γ and $\dot{\gamma}$ are linearly dependent, and thus γ is a parameterization of a straight line in $\mathfrak{gl}(n)$ connecting 0 and M . Hence $\Phi_A \circ \gamma$ is a reparametrization of the curve $t \mapsto \Phi_A(tM) = X(t)$, which shows that this length minimizing curve is (locally) uniquely determined. The above considerations are summarized in the following lemmas:

Lemma 4.4. *Let $A \in \text{GL}^+(n)$. Then there exists $\varepsilon(A) > 0$ such that for all $0 < \varepsilon < \varepsilon(A)$:*

- i) $\Phi_A : \mathfrak{gl}(n) \supseteq B_\varepsilon(0) \rightarrow B_\varepsilon(A) \subseteq \text{GL}^+(n)$ is a diffeomorphism,
- ii) For every $B \in B_\varepsilon(A)$ there exists a unique length minimizer connecting A and B ,
- iii) $\text{dist}_{\text{geod}}(A, \Phi_A(M)) = \|M\|_{\mu, \mu_c, \kappa}$ for $\|M\|_{\mu, \mu_c, \kappa} \leq \varepsilon$,
- iv) $\Phi_A(S_\varepsilon(0)) = S_\varepsilon(A)$,

where

$$S_\varepsilon(0) = \{M \in \mathfrak{gl}(n) : \|M\|_{\mu, \mu_c, \kappa} = \varepsilon\}$$

and

$$S_\varepsilon(A) = \{B \in \text{GL}^+(n) : \text{dist}_{\text{geod}}(A, B) = \varepsilon\}. \quad \blacksquare$$

Lemma 4.5. *The geodesic distance $\text{dist}_{\text{geod}}$ is a metric on $\text{GL}^+(n)$. \blacksquare*

4.2. Global existence of minimizers

After these preparations, we can now prove the global existence of length minimizing curves. We will closely follow the proof of the Hopf-Rinow theorem as given in [12, Theorem 1.4.8].

Proposition 4.6. *Let $A, B \in \text{GL}^+(n)$. Then there exists a length minimizing curve connecting A and B , i.e. $X \in \mathcal{A}_A^B$ with $L(X) = \text{dist}_{\text{geod}}(A, B)$.*

Proof. Let $B \in \text{GL}^+(n)$ be fixed. For $A \in \text{GL}^+(n)$, we define $r_A := \text{dist}_{\text{geod}}(A, B)$ and construct a length minimizer $X_A : [0, r_A] \rightarrow \text{GL}^+(n)$ as follows: Choose $\varepsilon(A)$ as in Lemma 4.4 and ε with $0 < \varepsilon < \varepsilon(A)$, $\varepsilon < \text{dist}_{\text{geod}}(A, B)$. Then the continuous function

$$\mathfrak{gl}(n) \supset S_\varepsilon(0) \rightarrow \mathbb{R}^+, N \mapsto \text{dist}_{\text{geod}}(\Phi_A(N), B) \quad (4.24)$$

attains a minimum at some $M_A \in S_\varepsilon(0)$. Note that by this definition, $\Phi_A(M_A)$ is closest to B among the geodesic sphere $S_\varepsilon(A)$. Next we define $X_A \in C^\infty([0, r_A]; \text{GL}^+(n))$ by

$$X_A(t) := \Phi_A(t\varepsilon^{-1}M_A). \quad (4.25)$$

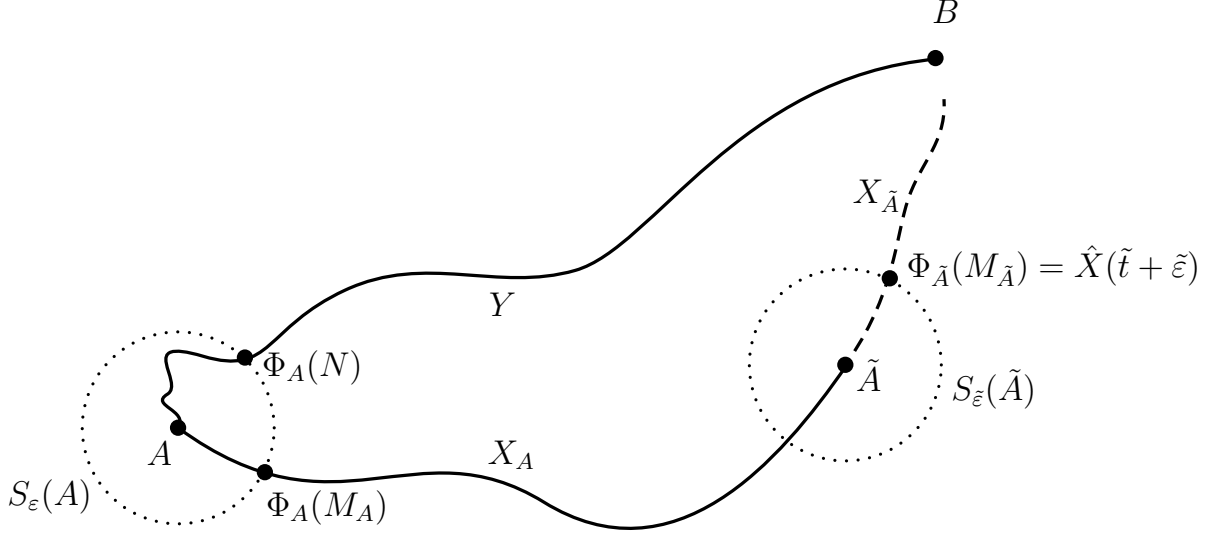


Figure 1: Global existence of length minimizers

To simplify notation, denote by $L_t(X_A) := L(X_A|_{[0,t]})$ the length of X_A up to $t \in [0, r_A]$. Then equation (3.10) yields $L_t(X_A) = t \varepsilon^{-1} \|M_A\|_{\mu, \mu_c, \kappa} = t$. Thus we find $\text{dist}_{\text{geod}}(X_A(t), A) \leq t$, and therefore

$$\text{dist}_{\text{geod}}(X_A(t), B) \geq \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(X_A(t), A) \geq \text{dist}_{\text{geod}}(A, B) - t. \quad (4.26)$$

We will denote by

$$I_A := \{t \in [0, r_A] : \text{dist}_{\text{geod}}(X_A(t), B) = r_A - t\} \quad (4.27)$$

the set of all t where equality holds in (4.26). Geometrically, I_A measures for how long X_A runs “in an optimal direction towards B ”. Since $L(X_A) = L_{r_A}(X_A) = r_A = \text{dist}_{\text{geod}}(A, B)$, X_A is a length minimizer between A and B if $X_A(r_A) = B$, thus it remains to show that $r_A \in I_A$. We will first show that $\varepsilon \in I_A$:

Assume $\text{dist}_{\text{geod}}(X_A(\varepsilon), B) > r_A - \varepsilon$. Then $\text{dist}_{\text{geod}}(A, B) = r_A < \varepsilon + \text{dist}_{\text{geod}}(X_A(\varepsilon), B)$, hence we can find $Y : [a, b] \rightarrow \text{GL}^+(n)$, $Y \in \mathcal{A}_A^B$ with

$$L(Y) < \varepsilon + \text{dist}_{\text{geod}}(X_A(\varepsilon), B) = \varepsilon + \text{dist}_{\text{geod}}(\Phi_A(M_A), B). \quad (4.28)$$

Since $\text{dist}_{\text{geod}}(A, B) > \varepsilon$ by definition of ε , the curve Y intersects $S_\varepsilon(A)$, so there exist $s \in (a, b)$ and, according to Lemma 4.4, $N \in S_\varepsilon(0)$ with $Y(s) = \Phi_A(N)$. Since $L(Y) = L(Y|_{[a,s]}) + L(Y|_{[s,b]})$, we obtain

$$\begin{aligned} \text{dist}_{\text{geod}}(\Phi_A(N), B) &\leq L(Y|_{[s,b]}) = L(Y) - L(Y|_{[a,s]}) \leq L(Y) - \text{dist}_{\text{geod}}(Y(a), Y(s)) \\ &= L(Y) - \text{dist}_{\text{geod}}(A, \Phi_A(N)) = L(Y) - \varepsilon < \text{dist}_{\text{geod}}(\Phi_A(M_A), B), \end{aligned} \quad (4.29)$$

in contradiction to the definition of M_A . Therefore $\varepsilon \in I_A$, and in particular I_A is nonempty.

Now assume $r_A \notin I_A$. It is not difficult to see that I_A is closed, so let $\tilde{t} := \max I_A \neq r_A$ and, for $\tilde{A} := X_A(\tilde{t})$, choose $\tilde{\varepsilon} < \varepsilon(\tilde{A})$, $M_{\tilde{A}}$, $X_{\tilde{A}}$, $I_{\tilde{A}}$ accordingly. We find $\tilde{\varepsilon} \in I_{\tilde{A}}$, and thus $\text{dist}_{\text{geod}}(X_{\tilde{A}}(\tilde{\varepsilon}), B) = \text{dist}_{\text{geod}}(\tilde{A}, B) - \tilde{\varepsilon}$.

Let \hat{X} denote the piecewise smooth curve

$$\hat{X} : [0, \tilde{t} + \tilde{\varepsilon}] \rightarrow \text{GL}^+(n), \quad \hat{X}(t) = \begin{cases} X_A(t) & : t \leq \tilde{t} \\ X_{\tilde{A}}(t - \tilde{t}) & : t > \tilde{t} \end{cases} \quad (4.30)$$

obtained by “attaching” $X_{\tilde{A}}$ to X_A . Then from

$$L(\hat{X}) = L(\hat{X}|_{[0, \tilde{t}]}) + L(\hat{X}|_{[\tilde{t}, \tilde{t} + \tilde{\varepsilon}]}) = \tilde{t} + \tilde{\varepsilon} \quad (4.31)$$

and

$$\begin{aligned} \text{dist}_{\text{geod}}(A, \hat{X}(\tilde{t} + \tilde{\varepsilon})) &= \text{dist}_{\text{geod}}(A, X_{\tilde{A}}(\tilde{\varepsilon})) \geq \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(X_{\tilde{A}}(\tilde{\varepsilon}), B) \\ &\stackrel{\tilde{\varepsilon} \in I_{\tilde{A}}}{=} \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(\tilde{A}, B) + \tilde{\varepsilon} \\ &= \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(X_A(\tilde{t}), B) + \tilde{\varepsilon} \stackrel{\tilde{t} \in I_A}{=} \tilde{t} + \tilde{\varepsilon}, \end{aligned}$$

we infer that \hat{X} is a length minimizer between A and $\hat{X}(\tilde{t} + \tilde{\varepsilon})$. Since $X_A, X_{\tilde{A}}$ have (the same) constant speed, \hat{X} is a piecewise smooth energy minimizer in $\mathcal{A}_A^{\hat{X}(\tilde{t} + \tilde{\varepsilon})}$. Then according to Proposition 2.4, \hat{X} is smooth and fulfils (2.24) everywhere. Finally, because X_A fulfils the differential equation with the same initial values as \hat{X} , the uniqueness from Proposition 2.7 yields

$$\hat{X} = X_A \text{ on } [0, \tilde{t} + \tilde{\varepsilon}], \quad (4.32)$$

and therefore

$$\begin{aligned} \text{dist}_{\text{geod}}(X_A(\tilde{t} + \tilde{\varepsilon}), B) &= \text{dist}_{\text{geod}}(X_{\tilde{A}}(\tilde{\varepsilon}), B) = \text{dist}_{\text{geod}}(\tilde{A}, B) - \tilde{\varepsilon} \\ &= \text{dist}_{\text{geod}}(X_A(\tilde{t}), B) - \tilde{\varepsilon} \\ &= \text{dist}_{\text{geod}}(A, B) - (\tilde{t} + \tilde{\varepsilon}) \Rightarrow \tilde{t} + \tilde{\varepsilon} \in I_A, \end{aligned} \quad (4.33)$$

in contradiction to the choice of \tilde{t} . Therefore $r_A \in I_A$, and thus $\text{dist}_{\text{geod}}(X_A(r_A), B) = r_A - r_A = 0$. Since $\text{dist}_{\text{geod}}$ is a metric on $\text{GL}^+(n)$ according to 4.5, we find $X(r_A) = B$. Since $L(X_A) = r_A$, this concludes the proof. \blacksquare

4.3. Conclusion

As above, let g be a left-invariant, right- $O(n)$ -invariant Riemannian metric on $\text{GL}(n)$ given by

$$\begin{aligned} g_A(M, N) &= \langle A^{-1}M, A^{-1}N \rangle_{\mu, \mu_c, \kappa}, \\ \langle M, N \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev sym } M, \text{dev sym } N \rangle + \mu_c \langle \text{skew } M, \text{skew } N \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } N, \\ \langle M, N \rangle &= \text{tr}(M^T N) \end{aligned}$$

for $A \in \text{GL}(n)$, $M, N \in \mathfrak{gl}(n)$. Combining the global existence of length minimizing geodesics (Proposition 4.6) with the fact that a length minimizer can be reparameterized into an energy minimizer $X : [0, 1] \rightarrow \text{GL}(n)$ (Lemmas 1.5 and 1.8) and the representation formula of energy minimizers from Theorem 3.1, we obtain the central result of this paper.

Theorem 4.7. *Let $A, B \in \text{GL}^+(n)$. Then there exists an initial direction $M \in \mathfrak{gl}(n)$ such that the curve $X : [0, 1] \rightarrow \text{GL}^+(n)$ with*

$$X(t) = A \exp(t(\text{sym } M - \omega \text{skew } M)) \exp(t(1 + \omega) \text{skew } M) \quad (4.34)$$

is a shortest curve connecting A and B , i.e. $X(0) = A$, $X(1) = B$ and

$$\|M\|_{\mu, \mu_c, \kappa} = L(X) = \text{dist}_{\text{geod}}(A, B) = \inf_{Y \in \mathcal{A}_A^B} L(Y). \quad (4.35)$$

Corollary 4.8. *The geodesic distance between A and B is given by*

$$\begin{aligned} \text{dist}_{\text{geod}}(A, B) &= \min\{\|M\|_{\mu, \mu_c, \kappa} \mid M \in \mathfrak{gl}(n), \\ &\quad A \exp(\text{sym } M - \omega \text{skew } M) \exp((1 + \omega) \text{skew } M) = B\}. \end{aligned} \quad (4.36)$$

In particular, the set is non-empty for all $A, B \in \text{GL}^+(n)$.

5. Special cases and applications

For given $A, B \in \text{GL}^+(n)$ it is still quite difficult to compute the geodesic distance $\text{dist}_{\text{geod}}(A, B)$ using formula (4.36): There is no known closed form solution for finding an $M \in \mathfrak{gl}(n)$ with

$$\exp(\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M) = A^{-1}B, \quad (5.1)$$

let alone one that is minimal with regard to the norm $\|\cdot\|_{\mu, \mu_c, \kappa}$. We will therefore consider some easier special cases.

5.1. The geodesic distance on $\text{GL}^+(1)$

In the one dimensional case, we can identify $\text{GL}^+(1)$ with \mathbb{R}^+ and the matrix multiplication on $\mathfrak{gl}(1)$ with the usual multiplication on \mathbb{R} . The most general inner product on \mathbb{R} is given by

$$\langle x, y \rangle_\kappa = \kappa xy, \quad \kappa > 0, \quad (5.2)$$

with the corresponding left-invariant Riemannian metric

$$g_p(x, y) = \langle p^{-1}x, p^{-1}y \rangle_\kappa = \left\langle \frac{x}{p}, \frac{y}{p} \right\rangle_\kappa = \kappa \frac{xy}{p^2}. \quad (5.3)$$

The length and energy of a piecewise differentiable curve $X \in C^0([0, 1]; \mathbb{R}^+)$ are therefore

$$\begin{aligned} L(X) &= \int_0^1 \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))} dt = \int_0^1 \sqrt{\frac{\dot{X}(t) \cdot \dot{X}(t)}{X(t)^2}} dt = \kappa \int_0^1 \left| \frac{\dot{X}(t)}{X(t)} \right| dt, \\ E(X) &= \int_0^1 g_{X(t)}(\dot{X}(t), \dot{X}(t)) dt = \kappa^2 \int_0^1 \left(\frac{\dot{X}(t)}{X(t)} \right)^2 dt. \end{aligned} \quad (5.4)$$

It is easy to see that, in order to minimize the length over all curves connecting $p, q \in \mathbb{R}^+$, we can assume that X is strictly monotone. In this case, X is uniquely determined by p and q up to a reparameterization.

We recall from Lemma 1.8 that a curve $X \in \mathcal{A}_p^q$ is an energy minimizer if and only if it is a length minimizer of constant speed. Since the length is invariant under reparameterization (Lemma 1.6), L is constant on the set of strictly monotone curves $X \in \mathcal{A}_p^q$, and therefore any X with

$$g_X(\dot{X}, \dot{X}) \equiv \text{constant}, \quad X(0) = p, \quad X(1) = q \quad (5.5)$$

is an energy minimizer. To solve (5.5), we define

$$X : [0, 1] \mapsto \text{GL}^+(1), \quad X(t) = p \cdot \exp(t \ln(\frac{q}{p})), \quad (5.6)$$

and check

$$g_{X(t)}(\dot{X}(t), \dot{X}(t)) = \kappa \frac{\dot{X}(t)^2}{X(t)^2} = \kappa \frac{p^2 \exp(t \ln(\frac{q}{p}))^2 \ln(\frac{q}{p})^2}{p^2 \exp(t \ln(\frac{q}{p}))^2} \equiv \kappa (\ln(\frac{q}{p}))^2, \quad (5.7)$$

as well as

$$X(0) = p, \quad X(1) = p \frac{q}{p} = q. \quad (5.8)$$

Thus X is an energy minimizer, and its energy and length are given by

$$\min_{Y \in \mathcal{A}_p^q} E(Y) = E(X) = \kappa^2 \int_0^1 \left(\frac{\dot{X}(t)}{X(t)} \right)^2 dt = \kappa^2 \int_0^1 \left(\ln(\frac{q}{p}) \right)^2 dt = \kappa^2 |\ln(\frac{q}{p})|^2 = \kappa^2 (\ln(q) - \ln(p))^2$$

and

$$\min_{Y \in \mathcal{A}_p^q} L(Y) = L(X) = \kappa \int_0^1 \left| \frac{\dot{X}(t)}{X(t)} \right| dt = \kappa \int_0^1 |\ln(\frac{q}{p})| dt = \kappa |\ln(\frac{q}{p})| = \kappa |\ln(q) - \ln(p)|. \quad (5.9)$$

We conclude:

Proposition 5.1. *The geodesic distance between $p, q \in \text{GL}^+(1) \cong \mathbb{R}^+$ is*

$$\text{dist}_{\text{geod}}(p, q) = \kappa |\ln(\frac{q}{p})|, \quad (5.10)$$

and a shortest geodesic connecting p and q is given by

$$X : [0, 1] \rightarrow \text{GL}^+(1), \quad X(t) = p \cdot \exp(t \ln(\frac{q}{p})). \quad (5.11)$$

5.2. Normal matrices

The following Lemma states some properties of normal matrices and their relation to the matrix exponential. A matrix $M \in \mathfrak{gl}(n)$ is called normal if $MM^T = M^T M$.

Lemma 5.2. *Let $A \in \text{GL}^+(n)$, $M \in \mathfrak{gl}(n)$. Then:*

- i) *M is normal if and only if $\text{sym } M$ and $\text{skew } M$ commute.*
- ii) *If M is normal, then $\exp(M)$ is normal.*
- iii) *If A is normal, then there exists a normal matrix $N \in \mathfrak{gl}(n)$ with $\exp(N) = A$.*

Proof. i) can be shown by direct computation:

$$\begin{aligned} & (\text{sym } M)(\text{skew } M) = (\text{skew } M)(\text{sym } M) \\ \Leftrightarrow & \quad \frac{1}{2}(M + M^T) \cdot \frac{1}{2}(M - M^T) = \frac{1}{2}(M - M^T) \cdot \frac{1}{2}(M + M^T) \\ \Leftrightarrow & \quad \frac{1}{4}(M^2 + M^T M - M M^T - (M^T)^2) = \frac{1}{4}(M^2 - M^T M + M M^T - (M^T)^2) \\ \Leftrightarrow & \quad M^T M - M M^T = -M^T M + M M^T \\ \Leftrightarrow & \quad M^T M = M M^T. \end{aligned} \quad (5.12)$$

Furthermore, if M and M^T commute, then according to (A.5), $\exp(M)$ and $\exp(M^T) = \exp(M)^T$ commute as well, proving ii). For a proof of iii) we refer to [2, Proposition 11.2.8]. ■

5.2.1. Geodesics with normal initial tangents

Let $N \in \mathfrak{gl}(n)$ be a normal matrix. Then, according to Theorem 3.1, the geodesic curve $X : [0, 1] \rightarrow \text{GL}(n)$ with

$$X(t) = \Phi(tN) = \exp(t(\text{sym } N - \omega \text{ skew } N)) \exp(t(1 + \omega) \text{ skew } N) \quad (5.13)$$

is the unique solution to the geodesic equation

$$\begin{cases} U = X^{-1} \dot{X}, \\ \dot{U} = \frac{1 + \mu \epsilon}{2} \mu (U^T U - U U^T) \end{cases} \quad (5.14)$$

with the initial values

$$X(0) = \mathbb{1}, \quad \dot{X}(t) = N. \quad (5.15)$$

According to Lemma 5.2, $\text{sym } N$ and $\text{skew } N$ commute if N is normal. But then $\lambda_1 \text{sym } N$ and $\lambda_2 \text{skew } N$ commute as well for all $\lambda_1, \lambda_2 \in \mathbb{R}$, and (A.6) yields $\exp(\lambda_1 \text{sym } N + \lambda_2 \text{skew } N) = \exp(\lambda_1 \text{sym } N) \exp(\lambda_2 \text{skew } N)$. This allows us to simplify X : we find

$$\begin{aligned} X(t) &= \exp(t \text{sym } N - t \omega \text{skew } N) \exp(t(1 + \omega) \text{skew } N) \\ &= \exp(t \text{sym } N) \exp(-t \omega \text{skew } N) \exp(t(1 + \omega) \text{skew } N) \\ &= \exp(t \text{sym } N) \exp(-t \omega \text{skew } N + t(1 + \omega) \text{skew } N) \\ &= \exp(t \text{sym } N) \exp(t \text{skew } N) = \exp(t \text{sym } N + t \text{skew } N) = \exp(tN). \end{aligned} \quad (5.16)$$

This representation of geodesics with normal initial tangents can also be found in [24, Section 8.5.1]. The length of X is

$$\begin{aligned} L(X) &= \int_0^1 \sqrt{\langle X^{-1} \dot{X}, X^{-1} \dot{X} \rangle_{\mu, \mu_c, \kappa}} dt \\ &= \int_0^1 \sqrt{\langle \exp(tN)^{-1} \exp(tN) N, \exp(tN)^{-1} \exp(tN) N \rangle_{\mu, \mu_c, \kappa}} dt \\ &= \int_0^1 \sqrt{\langle N, N \rangle_{\mu, \mu_c, \kappa}} dt = \|N\|_{\mu, \mu_c, \kappa} = \sqrt{\mu \|\text{dev sym } N\|^2 + \mu_c \|\text{skew } N\|^2 + \frac{\kappa}{n} \text{tr}[N]^2} \end{aligned} \quad (5.17)$$

and $X(0) = \mathbb{1}$, $X(1) = \exp(N)$. In particular, $X(1)$ is normal according to Lemma 5.2 as it is the exponential of a normal matrix.

5.2.2. Geodesics connecting $\mathbb{1}$ with a normal matrix

Now let $A \in \text{GL}^+(n)$ be normal. To find a geodesic X connecting $\mathbb{1}$ and A , we need to find $M \in \mathfrak{gl}(n)$ solving

$$A = \exp(\text{sym } M - \omega \text{skew } M) \exp((1 + \omega) \text{skew } M) = X(1). \quad (5.18)$$

But, again according to Lemma 5.2, a normal matrix A has a (generally not uniquely determined) *normal logarithm*, i.e. a normal matrix $\text{Log } A \in \mathfrak{gl}(n)$ with $\exp(\text{Log } A) = A$. Then we have shown in (5.16) that, with $N = \text{Log } A$, the geodesic X has the form

$$X(t) = \exp(t(\text{sym } \text{Log } A - \omega \text{skew } \text{Log } A)) \exp(t(1 + \omega) \text{skew } \text{Log } A) = \exp(t \text{Log } A),$$

and hence

$$X(1) = \exp(\text{Log } A) = A.$$

Thus for normal matrices A , the curve $X : [0, 1] \rightarrow \text{GL}^+(n) : X(t) = \exp(t \text{Log } A)$ is a geodesic connecting $\mathbb{1}$ and A , and (5.17) yields $L(X) = \|\text{Log } A\|_{\mu, \mu_c, \kappa}$. We therefore obtain the upper bound

$$\text{dist}_{\text{geod}}(\mathbb{1}, A) \leq L(X) = \|\text{Log } A\|_{\mu, \mu_c, \kappa} \quad (5.19)$$

for the distance of a normal matrix $A \in \text{GL}^+(n)$ to the identity $\mathbb{1}$.

Note carefully that this does not immediately imply $\text{dist}_{\text{geod}}(\mathbb{1}, A) = \|\text{Log } A\|_{\mu, \mu_c, \kappa}$: While $\text{dist}_{\text{geod}}(\mathbb{1}, A)$ is indeed the length of a geodesic curve connecting $\mathbb{1}$ and A , such a geodesic is generally not uniquely determined, and it is therefore possible that X is not the *shortest* such geodesic. However, as shown in Lemma 4.4, geodesic curves are *locally* unique, which immediately implies the following proposition.

Proposition 5.3. *Let $\mu, \mu_c, \kappa > 0$. Then there exists $\varepsilon > 0$ such that for every normal $A \in \text{GL}^+(n)$ with $\|A - \mathbb{1}\| < \varepsilon$, the geodesic distance between $\mathbb{1}$ and A is given by*

$$\text{dist}_{\text{geod}}(\mathbb{1}, A) = \min \|\text{Log } A\|_{\mu, \mu_c, \kappa} := \min \{\|M\|_{\mu, \mu_c, \kappa} \mid \exp(M) = A\}.$$

In particular, for $F \in \text{GL}^+(n)$ with sufficiently small $\|F - \mathbb{1}\|$, Proposition 5.3 can be applied to the positive definite symmetric (and therefore normal) matrix $F^T F$ to obtain the distance

$$\begin{aligned} \text{dist}^2(F, F^{-T}) &= \text{dist}^2(F^T F, \mathbb{1}) = \min\{\|M\|_{\mu, \mu_c, \kappa}^2 \mid \exp(M) = F^T F\} \\ &= \mu \|\text{dev} \log(F^T F)\|^2 + \frac{\kappa}{n} [\text{tr} \log(F^T F)]^2, \end{aligned} \quad (5.20)$$

where $\log(F^T F)$ is the symmetric *principal matrix logarithm* of $F^T F$ on $\text{PSym}(n)$ [2, 10]. Note, again, that it is not obvious at this point that equality (5.20) holds for all $F \in \text{GL}^+(n)$ since it is not immediately clear that there is no shorter geodesic curve connecting $\mathbb{1}$ and $F^T F$.

5.3. Application to nonlinear elasticity

An example for the application of geodesic distance measures is the theory of elasticity and, more specific, hyperelasticity, where the deformation of a solid body is modelled through an energy functional W which depends on the gradient $F \in \text{GL}^+(n)$ of the deformation (an introduction to elasticity theory can be found in [3]). Similarly, Mielke's work on the geodesic distance in $\text{SL}(n)$ was primarily motivated by an application in the theory of plasticity [17].

Among the energy functions considered in nonlinear elasticity is the *isotropic Hencky strain energy*

$$W : \text{GL}^+(n) \rightarrow \mathbb{R}, \quad W(F) = \mu \|\text{dev} \log \sqrt{F^T F}\|^2 + \frac{\kappa}{2} [\text{tr}(\log \sqrt{F^T F})]^2, \quad (5.21)$$

which was introduced in 1928 by Heinrich Hencky [9]. Note that, because $\log \sqrt{F^T F}$ is symmetric, W can be written as

$$W(F) = \|\log \sqrt{F^T F}\|_{\mu, \mu_c, \kappa}^2 \quad (5.22)$$

for arbitrary $\mu_c > 0$. As was shown by Neff et al. [22, 21], the Hencky energy can be characterized as the geodesic distance (with respect to a left- $\text{GL}(n)$ -invariant, right $\text{O}(n)$ -invariant Riemannian metric) of the deformation gradient F to the group $\text{SO}(n)$ of rigid rotations. The proof of this result employs the parametrization of geodesic curves given in Theorem 3.1 and the characterization of the geodesic distance stated in Corollary 4.8 as well as a recently discovered optimality result regarding the matrix logarithm [14, 23].

Proposition 5.4. *Let g be the left-invariant Riemannian metric on $\text{GL}(n)$ induced by the isotropic inner product $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$ on $\mathfrak{gl}(n)$ with $\mu_c \geq 0$, and let $F \in \text{GL}^+(n)$. Then*

$$\text{dist}_{\text{geod}}(F, \text{SO}(n)) = \text{dist}_{\text{geod}}(F, R) = \|\log \sqrt{F^T F}\|_{\mu, \mu_c, \kappa} = \|\log U\|_{\mu, \mu_c, \kappa}, \quad (5.23)$$

where $F = RU$ with $R \in \text{SO}(n)$, $U = \sqrt{F^T F} \in \text{PSym}(n)$ is the polar decomposition of F and

$$\text{dist}_{\text{geod}}(F, \text{SO}(n)) = \inf_{Q \in \text{SO}(n)} \text{dist}_{\text{geod}}(F, Q) \quad (5.24)$$

denotes the geodesic distance of F to $\text{SO}(n)$.

Proof. See [21]. ■

Proposition 5.4 also shows that equality (5.20) holds globally, i.e. for all $F \in \text{GL}^+(n)$ and arbitrarily large $\|F - \mathbb{1}\|$.

5.4. Open Problems

Although the explicit parametrization of geodesic curves makes it possible to establish some lower bounds for the geodesic distance in certain special cases, there is no known general formula or algorithm to compute the distance between two elements A, B of $\text{GL}^+(n)$. However, recent results indicate that it might be possible to compute the geodesic distance for a number of additional special cases, including

the (non-local) distance between arbitrary $A, B \in \text{SO}(n)$ [16]. Note that although the geodesic distance on $\text{SO}(n)$ with respect to the canonical bi-invariant metric is already well known [18], the distance discussed here takes into account the length of connecting curves which do not lie completely in $\text{SO}(n)$. Furthermore, it might be useful to obtain some basic geometric properties of $\text{GL}(n)$ or $\text{SO}(n)$ with the considered metrics (e.g. to explicitly compute the curvature tensors).

Acknowledgements

We are grateful to Prof. Robert Bryant (Duke University), for his helpful remarks regarding geodesics on Lie groups and invariances of inner products on $\mathfrak{gl}(n)$ as well as to Prof. Alexander Mielke (Weierstraß-Institut, Berlin) for pertinent discussions on geodesics in $\text{GL}(n)$.

References

- [1] E. Andruchow, G. Larotonda, L. Recht, and A. Varela. The left invariant metric in the general linear group. *preprint arXiv:1109.0520*, 2011.
- [2] D. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton reference. Princeton University Press, 2009.
- [3] P. Ciarlet. *Three-Dimensional Elasticity*. Number 1 in Studies in mathematics and its applications. Elsevier Science, 1988.
- [4] C. De Boor. A naive proof of the representation theorem for isotropic, linear asymmetric stress-strain relations. *Journal of Elasticity*, 15(2):225–227, 1985.
- [5] J.-H. Eschenburg and J. Jost. *Differentialgeometrie und Minimalflächen*. Springer, 2007.
- [6] G. Fischer. *Lineare Algebra*. Vieweg-Studium: Grundkurs Mathematik. Vieweg, 2005.
- [7] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. U.S. Government Printing Office, 2004.
- [8] A. Gastel. Riemannsche Geometrie, 2004. Lecture notes.
- [9] H. Hencky. Über die Form des Elastizitätsgesetzes bei ideal elastischen Stoffen. *Zeitschrift für technische Physik*, 9:215–220, 1928. www.uni-due.de/imperia/md/content/mathematik/ag_neff/hencky1928.pdf.
- [10] N. J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
- [11] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.
- [12] J. Jost. *Riemannian Geometry and Geometric Analysis (2nd ed.)*. Springer, 1998.
- [13] K. Königsberger. *Analysis 1*. Analysis. Springer, 2004.
- [14] J. Lankeit, P. Neff, and Y. Nakatsukasa. The minimization of matrix logarithms: On a fundamental property of the unitary polar factor. *Linear Algebra and its Applications*, 449(0):28–42, 2014.
- [15] S. Lee, M. Choi, H. Kim, and F. C. Park. Geometric direct search algorithms for image registration. *IEEE Transactions on Image Processing*, 16(9):2215–2224, 2007.
- [16] R. Martin and P. Neff. On the geodesic convexity of $\text{SO}(n)$ in $\text{GL}(n)$. *in preparation*, 2014.
- [17] A. Mielke. Finite elastoplasticity, Lie groups and geodesics on $\text{SL}(d)$. In P. Newton, P. Holmes, and A. Weinstein, editors, *Geometry, Mechanics, and Dynamics*, pages 61–90. Springer New York, 2002.
- [18] M. Moakher. Means and averaging in the group of rotations. *SIAM J. Matrix Anal. Appl.*, 24(1):1–16, Jan. 2002.
- [19] P. Neff. *Mathematische Analyse multiplikativer Viskoplastizität*. Ph.D. Thesis, Technische Universität Darmstadt. Shaker Verlag, Aachen, 2000. available at http://www.uni-due.de/%7Ehm0014/Download_files/neffdiss.ps.
- [20] P. Neff. Convexity and coercivity in nonlinear, anisotropic elasticity and some useful relations. Technical report, Technische Universität Darmstadt, 2008. available at http://www.uni-due.de/%7Ehm0014/Download_files/cism_convexity08.pdf.
- [21] P. Neff, B. Eidel, F. Osterbrink, and R. Martin. Appropriate strain measures: The Hencky shear strain energy $\|\text{dev} \log \sqrt{F^T F}\|^2$ measures the geodesic distance of the isochoric part of the deformation gradient $\bar{F} \in \text{SL}(3)$ to $\text{SO}(3)$ in the canonical left invariant Riemannian metric on $\text{SL}(3)$. *in preparation*, 2014.

- [22] P. Neff, B. Eidel, F. Osterbrink, and R. Martin. A Riemannian approach to strain measures in nonlinear elasticity. *Comptes Rendus Mécanique*, 342(4):254–257, 2014.
- [23] P. Neff, Y. Nakatsukasa, and A. Fischle. A logarithmic minimization property of the unitary polar factor in the spectral norm and the Frobenius matrix norm. *to appear in SIAM J. Matrix Analysis*, 2014. available at arXiv:1302.3235.
- [24] C. H. Taubes. *Differential geometry: bundles, connections, metrics and curvature*. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2011.
- [25] B. Vandereycken, P. Absil, and S. Vandewalle. A Riemannian geometry with complete geodesics for the set of positive semidefinite matrices of fixed rank. *IMA Journal of Numerical Analysis*, 33:481–514, 2013.

A. Appendix

A.1. The matrix exponential

A.1.1. Basic properties of \exp on $\mathfrak{gl}(n)$

We will list, without proof, some basic properties of the exponential function.

Lemma A.1. *Let $B_i \in \mathfrak{gl}(d_i)$, $i \in \{1, \dots, k\}$ be square block matrices of size $d_i \times d_i$, $\lambda \in \mathbb{R}$ and $M, N \in \mathfrak{gl}(n)$. Then:*

$$i) \quad \det(\exp(M)) = e^{\text{tr } M}, \quad (\text{A.1})$$

$$ii) \quad \exp\left(\frac{\lambda}{n} \mathbb{1}\right) = e^\lambda \mathbb{1}, \quad (\text{A.2})$$

$$iii) \quad \exp(M^T) = \exp(M)^T, \quad (\text{A.3})$$

$$iv) \quad \exp(-M) = \exp(M)^{-1}, \quad (\text{A.4})$$

$$v) \quad MN = NM \Rightarrow \exp(M + N) = \exp(M) \exp(N) = \exp(N) \exp(M), \quad (\text{A.5})$$

$$vi) \quad MN = NM \Rightarrow M \exp(N) = \exp(N) M, \quad (\text{A.6})$$

$$vii) \quad T \in \text{GL}(n) \Rightarrow \exp(T^{-1} M T) = T^{-1} \exp(M) T, \quad (\text{A.7})$$

$$viii) \quad M \in \mathfrak{so}(n) \Rightarrow \exp(M) \in \text{SO}(n), \quad (\text{A.8})$$

$$ix) \quad M \in \text{Sym}(n) \Rightarrow \exp(M) \in \text{PSym}(n), \quad (\text{A.9})$$

$$x) \quad \exp \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_k \end{pmatrix} = \begin{pmatrix} \exp B_1 & & 0 \\ & \ddots & \\ 0 & & \exp B_k \end{pmatrix}, \quad (\text{A.10})$$

$$xi) \quad D \exp[M].T = \int_0^1 \exp(sM) T \exp((1-s)M) ds. \quad (\text{A.11})$$

If M and T commute, formula $xi)$ simplifies to

$$D \exp[M].T = \exp(M) T = T \exp(M). \quad (\text{A.12})$$

Proof for these elementary properties as well as more thorough introduction to the matrix exponential can be found in [10].

A.1.2. \exp on $\mathfrak{so}(2)$

In the two-dimensional case, the group $\text{SO}(2)$ can be described as the set of planar rotations: every $Q \in \text{SO}(2)$ is of the form

$$Q = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (\text{A.13})$$

To show the surjectivity of $\exp : \mathfrak{so}(2) \rightarrow \text{SO}(2)$, it therefore suffices to show that

$$\exp \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (\text{A.14})$$

Since $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^4 = \mathbb{1}$, this can be done by a straightforward computation:

$$\begin{aligned} & \exp \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^k \\ &= \sum_{k=0}^{\infty} \frac{1}{(4k)!} \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^{(4k)} + \sum_{k=0}^{\infty} \frac{1}{(4k+1)!} \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^{(4k+1)} + \sum_{k=0}^{\infty} \frac{1}{(4k+2)!} \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^{(4k+2)} + \sum_{k=0}^{\infty} \frac{1}{(4k+3)!} \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^{(4k+3)} \\ &= \sum_{k=0}^{\infty} \frac{\theta^{4k}}{(4k)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^0 + \sum_{k=0}^{\infty} \frac{\theta^{4k+1}}{(4k+1)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^1 + \sum_{k=0}^{\infty} \frac{\theta^{4k+2}}{(4k+2)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 + \sum_{k=0}^{\infty} \frac{\theta^{4k+3}}{(4k+3)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 \\ &= \sum_{k=0}^{\infty} \frac{\theta^{4k}}{(4k)!} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=0}^{\infty} \frac{\theta^{4k+1}}{(4k+1)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \sum_{k=0}^{\infty} \frac{\theta^{4k+2}}{(4k+2)!} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \sum_{k=0}^{\infty} \frac{\theta^{4k+3}}{(4k+3)!} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \end{aligned}$$

A.1.4. exp on $\text{Sym}(n)$

Proposition A.4. *The function*

$$\exp : \text{Sym}(n) \rightarrow \text{PSym}(n) \quad (\text{A.24})$$

is surjective, in other words: For every symmetric, positive definite matrix P (with $\det P = 1$), there is a symmetric (trace free) matrix S with $\exp(S) = P$.

Proof. Let $S \in \text{SL}(n)$ be given. Since every symmetric matrix is orthogonally diagonalizable, we can write

$$S = R^T D R \quad (\text{A.25})$$

with $R \in \text{SO}(n)$, $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Since S is positive definite and λ_i are the eigenvalues of S , we know that $\lambda_i > 0$ for all $i \in \{1, \dots, n\}$ and thus

$$\tilde{D} := \text{diag}(\ln(\lambda_1), \dots, \ln(\lambda_n)) \quad (\text{A.26})$$

is well-defined. Again, (A.7) and (A.10) yield

$$\exp(R^T \tilde{D} R) = R^T \exp(\tilde{D}) R = R^T \text{diag}(\exp(\ln \lambda_1), \dots, \exp(\ln \lambda_n)) R = R^T D R = S, \quad (\text{A.27})$$

■

Proposition A.5. *The function*

$$\exp : \text{Sym}(n) \rightarrow \text{PSym}(n)$$

is a diffeomorphism from $\text{Sym}(n)$ to its open subset $\text{PSym}(n)$. Its inverse

$$\exp|_{\text{PSym}(n)}^{-1} : \text{PSym}(n) \rightarrow \text{Sym}(n),$$

is called the principal logarithm on $\text{PSym}(n)$ and is denoted by \log .

Proof. We refer to [2, Proposition 11.4.5] for a proof that $\exp : \text{Sym}(n) \rightarrow \text{PSym}(n)$ is indeed injective and infinitely differentiable. To see that its inverse is differentiable, we refer to [11, Theorem 6.6.14], where it is shown that a primary matrix function F defined through a real valued function f acting on its eigenvalues is differentiable if f is smooth on the set of eigenvalues attained on the domain of F . Since the principal logarithm on $\text{PSym}(n)$ is such a function defined through $f = \ln$, and because \ln is smooth on \mathbb{R}^+ , i.e. on the set of eigenvalues attained on $\text{PSym}(n)$, it is differentiable. For further information on the matrix exponential, the matrix logarithm and matrix functions in general refer to [10]. ■

A.2. Additional proofs

Lemma A.6. *Let $M, N \in \mathfrak{gl}(n)$ and $\langle \cdot, \cdot \rangle$ denote the canonical inner product $\langle M, N \rangle = \text{tr}(M^T N)$ on $\mathfrak{gl}(n)$. Then:*

$$i) \quad \langle Q^T M Q, Q^T N Q \rangle = \langle M, N \rangle \quad \text{for all } Q \in \text{O}(n), \quad (\text{A.28})$$

$$ii) \quad \langle M, \mathbb{1} \rangle = \text{tr}(M), \quad (\text{A.29})$$

$$iii) \quad \langle S, W \rangle = 0 \quad \text{for all } S \in \text{Sym}(n), W \in \mathfrak{so}(n), \quad (\text{A.30})$$

Proof.

$$\begin{aligned} i): \quad \langle Q^T M Q, Q^T N Q \rangle &= \text{tr}((Q^T M Q)^T (Q^T N Q)) = \text{tr}(Q^T M^T Q Q^T N Q) \\ &= \text{tr}(Q^T M^T N Q) = \text{tr}(M^T N Q Q^T) = \text{tr}(M^T N) = \langle M, N \rangle. \end{aligned} \quad (\text{A.31})$$

$$ii): \quad \langle M, \mathbb{1} \rangle = \text{tr}(M^T \mathbb{1}) = \text{tr}(M^T) = \text{tr}(M). \quad (\text{A.32})$$

$$\begin{aligned} iii): \quad \langle S, W \rangle &= \text{tr}(S^T W) = \text{tr}(S W) = \text{tr}(W S) \\ &= -\text{tr}(W^T S) = -\langle W, S \rangle = -\langle S, W \rangle \implies \langle S, W \rangle = 0. \end{aligned} \quad (\text{A.33})$$

■

Lemma A.7. *Let $M, N \in \mathfrak{gl}(n)$ and $\mu, \mu_c, \kappa \geq 0$. Then*

$$i) \quad \langle Q^T M Q, Q^T N Q \rangle_{\mu, \mu_c, \kappa} = \langle M, N \rangle_{\mu, \mu_c, \kappa} \quad \text{for all } Q \in \text{O}(n), \quad (\text{A.34})$$

$$\begin{aligned} ii) \quad \langle M, N \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev sym } M, N \rangle + \mu_c \langle \text{skew } M, N \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } N \\ &= \langle \mu \text{dev sym } M + \mu_c \text{skew } M + \frac{\kappa}{n} \text{tr}[M] \mathbb{1}, N \rangle, \end{aligned} \quad (\text{A.35})$$

$$iii) \quad \langle M, N \rangle_{1,1,1} = \langle M, N \rangle, \quad (\text{A.36})$$

$$iv) \quad \langle S, W \rangle_{\mu, \mu_c, \kappa} = 0 \quad \text{for all } S \in \text{Sym}(n), W \in \mathfrak{so}(n), \quad (\text{A.37})$$

where $\text{Sym}(n)$ and $\mathfrak{so}(n)$ denote the set of symmetric and skew symmetric matrices in $\mathbb{R}^{n \times n}$ respectively.

Proof. We first show that for $M \in \mathfrak{gl}(n)$ the matrices $\text{dev sym } M$, $\text{skew } M$ and \mathbb{I} are pairwise perpendicular with respect to the canonical inner product $\langle \cdot, \cdot \rangle$. Lemma A.6 yields $\langle \text{dev sym } M, \text{skew } M \rangle = 0$ (note that $\text{dev sym } M$ is symmetric) as well as

$$\begin{aligned} \langle W, \mathbb{I} \rangle &= \text{tr}(W^T \mathbb{I}) = \text{tr}(W^T) = -\text{tr}(W) = -\text{tr}(\mathbb{I}^T W) = \langle W, \mathbb{I} \rangle, \\ \langle \text{dev sym } M, \mathbb{I} \rangle &= \text{tr}(\text{dev sym } M) = \text{tr}(M - \frac{\text{tr } M}{n} \mathbb{I}) = \text{tr } M - \frac{\text{tr } M}{n} \text{tr } \mathbb{I} = 0; \end{aligned} \quad (\text{A.38})$$

observe that $\text{tr } \mathbb{I} = n$ for $\mathbb{I} \in \mathbb{R}^{n \times n}$. We compute

$$\begin{aligned} &\langle \mu \text{dev sym } M + \mu_c \text{skew } M + \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, N \rangle \\ &= \langle \mu \text{dev sym } M + \mu_c \text{skew } M + \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \text{sym } N + \text{skew } N \rangle \\ &= \langle \mu \text{dev sym } M + \mu_c \text{skew } M + \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \text{dev sym } N + \text{skew } N + \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &= \langle \mu \text{dev sym } M, \text{dev sym } N + \text{skew } N + \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &\quad + \langle \mu_c \text{skew } M, \text{dev sym } N + \text{skew } N + \frac{\text{tr } N}{n} \mathbb{I} \rangle + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \text{dev sym } N + \text{skew } N + \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &= \langle \mu \text{dev sym } M, \text{dev sym } N \rangle + \langle \mu \text{dev sym } M, \text{skew } N \rangle + \langle \mu \text{dev sym } M, \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &\quad + \langle \mu_c \text{skew } M, \text{dev sym } N \rangle + \langle \mu_c \text{skew } M, \text{skew } N \rangle + \langle \mu_c \text{skew } M, \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &\quad + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \text{dev sym } N \rangle + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \text{skew } N \rangle + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &= \langle \mu \text{dev sym } M, \text{dev sym } N \rangle + \langle \mu_c \text{skew } M, \text{skew } N \rangle + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &= \mu \langle \text{dev sym } M, \text{dev sym } N \rangle + \mu_c \langle \text{skew } M, \text{skew } N \rangle + \langle \frac{\kappa}{n} \text{tr}[M] \mathbb{I}, \frac{\text{tr } N}{n} \mathbb{I} \rangle \\ &= \mu \langle \text{dev sym } M, \text{dev sym } N \rangle + \mu_c \langle \text{skew } M, \text{skew } N \rangle + \frac{\kappa}{n} \text{tr}[M] \text{tr}[N] = \langle M, N \rangle_{\mu, \mu_c, \kappa}, \end{aligned} \quad (\text{A.39})$$

proving ii). Then we can obtain i) by

$$\langle M, N \rangle_{1,1,1} = \langle \text{dev sym } M + \text{skew } M + \frac{\text{tr } M}{n} \mathbb{I}, N \rangle = \langle M, N \rangle \quad (\text{A.40})$$

as well as iii) by

$$\langle S, W \rangle = \langle W, S \rangle = \langle \mu \text{dev sym } W + \mu_c \text{skew } W + \frac{\kappa}{n} \text{tr}[W] \mathbb{I}, S \rangle = \mu_c \langle W, S \rangle = 0. \quad (\text{A.41})$$

Finally, we observe that

$$\begin{aligned} &\text{tr}(Q^T M Q) = \text{tr}(M Q Q^T) = \text{tr}(M), \\ &\text{sym}(Q^T M Q) = \frac{1}{2}(Q^T M Q + (Q^T M Q)^T) = \frac{1}{2}(Q^T (M + M^T) Q) = Q^T (\text{sym } M) Q, \\ &\text{skew}(Q^T M Q) = \frac{1}{2}(Q^T M Q - (Q^T M Q)^T) = \frac{1}{2}(Q^T (M - M^T) Q) = Q^T (\text{skew } M) Q, \\ &\text{dev sym}(Q^T M Q) = \text{dev}(Q^T (\text{sym } M) Q) = Q^T (\text{sym } M) Q - \frac{\text{tr}(Q^T (\text{sym } M) Q)}{n} \mathbb{I} \\ &= Q^T (\text{sym } M) Q - \frac{\text{tr}(\text{sym } M)}{n} \mathbb{I} = Q^T (\text{sym } M - \frac{\text{tr}(\text{sym } M)}{n} \mathbb{I}) Q = Q^T (\text{dev sym } M) Q \end{aligned} \quad (\text{A.42})$$

and show i) by computing

$$\begin{aligned} \langle Q^T M Q, Q^T N Q \rangle_{\mu, \mu_c, \kappa} &= \mu \langle \text{dev sym}(Q^T M Q), \text{dev sym}(Q^T N Q) \rangle + \mu_c \langle \text{skew}(Q^T M Q), \text{skew}(Q^T N Q) \rangle \\ &\quad + \frac{\kappa}{n} \text{tr}(Q^T M Q) \text{tr}(Q^T N Q) \\ &= \mu \langle Q^T (\text{dev sym } M) Q, Q^T (\text{dev sym } N) Q \rangle + \mu_c \langle Q^T (\text{skew } M) Q, Q^T (\text{skew } N) Q \rangle \\ &\quad + \frac{\kappa}{n} \text{tr } M \text{tr } N \\ &= \mu \langle \text{dev sym } M, \text{dev sym } N \rangle + \mu_c \langle \text{skew } M, \text{skew } N \rangle + \frac{\kappa}{n} \text{tr } M \text{tr } N = \langle M, N \rangle_{\mu, \mu_c, \kappa}. \quad \blacksquare \end{aligned}$$

Lemma A.8. *Let $Q \in C^1([a, b]; \text{O}(n))$. Then $Q^T(t) \dot{Q}(t) \in \mathfrak{so}(n)$ for all $t \in [a, b]$.*

Proof. $Q(t) \in \text{O}(n)$ implies $Q^T Q \equiv \mathbb{I}$ and thus

$$0 = \frac{d}{dt}(Q^T(t) Q(t)) = \dot{Q}^T(t) Q(t) + Q^T(t) \dot{Q}(t) = Q^T(t) \dot{Q}(t) + (Q^T(t) \dot{Q}(t))^T = 2 \text{sym}(Q^T(t) \dot{Q}(t)). \quad \blacksquare$$