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

Robert Martin\textsuperscript{1} and Patrizio Neff\textsuperscript{2}

June 4, 2021

Abstract

We provide an easy approach to the geodesic distance on the general linear group $\text{GL}(n)$ for left-invariant Riemannian metrics which are also right-$\text{O}(n)$-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.

Contents

1 Introduction and preliminaries \hfill 2
  1.1 Basic Definitions \hfill 2
  1.2 Distance functions on $\mathbb{R}^{n\times n}$ \hfill 2
  1.3 The geodesic distance \hfill 3
  1.4 Length and energy of curves \hfill 5
  1.5 Left-invariant, right-$\text{O}(n)$-invariant Riemannian metrics \hfill 7

2 Energy minimizers and the calculus of variations \hfill 9
  2.1 The geodesic equation for the canonical inner product \hfill 10
  2.2 The geodesic equation for the general metric \hfill 14
  2.3 Properties of solutions $U$ \hfill 16
  2.4 Existence and uniqueness of solutions \hfill 16

3 The parameterization of geodesics \hfill 17
  3.1 Solving the geodesic equation \hfill 17
  3.2 Properties of geodesic curves \hfill 18

4 The existence of length minimizers \hfill 18
  4.1 Preparations \hfill 19
  4.2 Global existence of minimizers \hfill 23
  4.3 Conclusion \hfill 25

5 Special cases and applications \hfill 25
  5.1 The geodesic distance on $\text{GL}^+(1)$ \hfill 26
  5.2 Normal matrices \hfill 27
  5.3 Application to nonlinear elasticity \hfill 29
  5.4 Open Problems \hfill 29

6 References \hfill 30

A Appendix \hfill 31
  A.1 Basic properties of the matrix exponential \hfill 31
  A.2 Additional proofs \hfill 31

\textsuperscript{1}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

\textsuperscript{2}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
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 tensor $\log F^T F$ of a deformation gradient $F \in \text{GL}^+(n)$ and the geodesic distance of $F$ to the special orthogonal group $\text{SO}(n)$. Since the requirements of objectivity and isotropy strongly suggest a distance measure on $\text{GL}^+(n)$ which is right-invariant under rotations and left-invariant with respect to action of $\text{GL}^-(n)$, we restrict our considerations to Riemannian metrics on $\text{GL}(n)$ which are left-$\text{GL}(n)$-invariant as well as right $\text{O}(n)$-invariant.

Although the theory of Lie groups is obviously applicable to $\text{GL}(n)$ 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 $\text{GL}^+(n)$ [1] as well as for left-invariant, right-$\text{SO}(n)$-invariant metrics on $\text{SL}(n)$ [18], analogous results are not found in the literature for the more general case on $\text{GL}(n)$.

The aim of this paper is therefore to provide a more accessible approach to this type of Riemannian metrics on $\text{GL}(n)$ 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 $\text{GL}(n)$.

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

$\text{GL}(n) = \{ X \in \mathbb{R}^{n \times n} \mid \det(X) \neq 0 \}$ \hspace{1cm} (general linear group),

$\text{GL}^+(n) = \{ X \in \mathbb{R}^{n \times n} \mid \det(X) > 0 \}$,

$\text{GL}^-(n) = \{ X \in \mathbb{R}^{n \times n} \mid \det(X) < 0 \}$,

$\text{SL}(n) = \{ X \in \mathbb{R}^{n \times n} \mid \det(X) = 1 \}$ \hspace{1cm} (special linear group),

$\text{O}(n) = \{ X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1} \}$ \hspace{1cm} (orthogonal group),

$\text{SO}(n) = \{ X \in \mathbb{R}^{n \times n} \mid X^T X = \mathbb{1} \text{ and } \det(X) = 1 \}$ \hspace{1cm} (special orthogonal group).

Furthermore, we define the set of symmetric matrices $\text{Sym}(n) = \{ X \in \mathbb{R}^{n \times n} \mid X^T = X \}$, the set of positive definite symmetric matrices $\text{Sym}^+(n) = \{ X \in \mathbb{R}^{n \times n} \mid X^T = X \text{ and } v^T X v > 0 \forall v \in \mathbb{R}^n \setminus \{0\} \}$ 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 $\text{dist}: M \times M \to [0, \infty]$ with

$\text{dist}(A, A) = 0$, \hspace{1cm} $\text{dist}(A, B) = \text{dist}(B, A)$

and \hspace{1cm} $\text{dist}(A, C) \leq \text{dist}(A, B) + \text{dist}(B, C)$ \hspace{1cm} (triangle inequality) \hspace{1cm} (1.1)

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

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

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

$\langle M, N \rangle = \text{tr}(M^T N) = \sum_{i,j=1}^{n} M_{ij} N_{ij}$, \hspace{1cm} $M, N \in \mathbb{R}^{n \times n}$, \hspace{1cm} (1.2)
where \( \text{tr} M = \sum_{i=1}^{n} M_{i,i} \) denotes the trace of \( M \), induces the Euclidean norm (or Frobenius matrix norm)

\[
\|M\| = \sqrt{\langle M, M \rangle} = \sqrt{\sum_{i,j=1}^{n} M_{ij}^2},
\]

and the Euclidean distance is the metric given by

\[
\text{dist}_{\text{Euclid}}(A, B) = \|A - B\|.
\]

While \( \text{dist}_{\text{Euclid}} \) induces a distance function on \( \text{GL}(n) \) as well, it does not appear as a “natural” inner property of the general linear group: since \( \text{GL}(n) \) 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 \( \text{GL}(n) \) is not a closed subset of \( \mathbb{R}^{n \times n} \), it is not complete with respect to the Euclidean distance.

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

\[
T_A \text{GL}(n) = A \cdot T_1 \text{GL}(n) \cong A \cdot \mathbb{R}^{n \times n} = \mathbb{R}^{n \times n} =: \mathfrak{gl}(n).
\]

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

### 1.3 The geodesic distance

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

\[
g: \text{GL}(n) \times \mathfrak{gl}(n) \times \mathfrak{gl}(n) \to \mathbb{R}, \ (A, M, N) \mapsto g_A(M, N)
\]

such that for every fixed \( A \in \text{GL}(n) \) the function \( g_A(\cdot, \cdot): \mathfrak{gl}(n) \times \mathfrak{gl}(n) \to \mathbb{R} \) 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 \( \text{GL}(n) \): the length \( L \) and energy \( E \) of a curve \( X \in C^1([a, b]; \text{GL}(n)) \) are given by

\[
L(X) := \int_0^1 \sqrt{g_X(t)(\dot{X}(t), \dot{X}(t))} \, dt \quad \text{and} \quad E(X) := \int_0^1 g_X(t)(\dot{X}(t), \dot{X}(t)) \, dt,
\]

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 \( \text{dist}_{\text{geod}}(A, B) \) between \( A, B \in \text{GL}(n) \) 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

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

the set of regular \( k \)-times differentiable curves in \( \text{GL}(n) \) over the interval \([a, b]\). Note that, by the usual definition of differentiability on closed intervals as the restriction of differentiable functions on \( \mathbb{R} \), the \( i \)-th

\footnote{Note that in the theory of Lie Groups \( \mathfrak{gl}(n) \) usually denotes (or, more precisely, is identified with) the tangent space \( T_1 \text{GL}(n) \) at the identity only. From the perspective of classical analysis, all tangent spaces \( T_A \text{GL}(n) \) can simply be identified with \( \mathbb{R}^{n \times n} \), allowing us to employ a much simpler notion of “smoothness”.}
derivative \( X^{(i)}(a) \) and \( X^{(i)}(b) \) at the boundaries is well-defined via the one-sided limits \( \lim_{t \downarrow a} X^{(i)}(t) \) and \( \lim_{t \uparrow b} X^{(i)}(t) \). We now define the set of \textit{piecewise k-times differentiable curves} in \( \text{GL}(n) \) over the interval \([a, b] \) by

\[
\mathcal{A}^k([a, b]) := \{ X \in C^0([a, b]; \text{GL}(n)) \mid \exists a = a_0 < a_1 < \cdots < a_{m+1} = b, \forall j = 0, \ldots, m : X|_{[a_j, a_{j+1}]} \in C^k([a_j, a_{j+1}]; \text{GL}(n)) \}.
\]  

(1.9)

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

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

(1.10)

the admissible set over the fixed interval \([a, b] \) is denoted by

\[
\mathcal{A}^B_A := \{ X : [a, b] \to \text{GL}(n), X \in \mathcal{A}^B_A \}
\]  

(1.11)

and the general \textit{admissible set of curves} is

\[
\mathcal{A} := \bigcup_{A, B \in \text{GL}(n)} \mathcal{A}^B_A.
\]  

(1.12)

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

We can now properly define the geodesic distance function:

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

\[
\text{dist}_{\text{geod}}(A, B) := \inf_{X \in \mathcal{A}^B_A} L(X)
\]  

(1.13)

is called the \textit{geodesic distance} between two matrices \( A \) and \( B \).

\textbf{Remark 1.4.} It is easy to verify that \( \text{dist}_{\text{geod}} \) is indeed a distance function; in order to see that it satisfies the triangle inequality \( (1.1) \), choose a curve \( X \in \mathcal{A}^B_A \) connecting \( A \) and \( B \) with \( L(X) \leq \text{dist}_{\text{geod}}(A, B) + \varepsilon \) as well as a curve \( Y \in \mathcal{A}^C_B \) 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.

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

(1.14)

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

\[
L(Z) = \int_0^2 \sqrt{g_Z(\dot{Z}, \dot{Z})} \, dt = \int_0^1 \sqrt{g_Z(\dot{X}, \dot{X})} \, dt + \int_1^2 \sqrt{g_Y(\dot{Y}, \dot{Y})} \, dt
\]  

(1.15)

\[
= \int_0^1 \sqrt{g_X(\dot{X}, \dot{X})} \, dt + \int_0^1 \sqrt{g_Y(\dot{Y}, \dot{Y})} \, dt
\]

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.

\footnote{Here and throughout we will often omit the integration variable and write e.g. \( \dot{Z} \) instead of \( \dot{Z}(t) \).}
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, $\text{GL}(n)$ is not connected, but can be decomposed into two connected components $\text{GL}^+(n) = \{ A \in \text{GL}(n) \mid \det A > 0 \}$ and $\text{GL}^-(n) = \{ A \in \text{GL}(n) \mid \det A < 0 \}$. Thus $\text{dist}_{\text{geod}}(A, B) < \infty$ if and only if $A$ and $B$ are in the same connected component, i.e. if $\det(AB) > 0$. 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

In order to further investigate the geodesic distance, some basic properties of curves in $\text{GL}(n)$, the length and the energy functional are required. Some of these properties can be found in any textbook on differential geometry; however, in order to keep this article self contained and accessible to readers unfamiliar with the methods of general differential geometry, we explicitly state and prove them here. Many properties of curves in $\text{GL}(n)$ also correspond directly to the case of curves in the Euclidean space (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 $T_A \text{GL}(n) = \mathbb{R}^{n \times n} = \mathfrak{gl}(n)$, 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 differentiable curve $X \in \mathcal{A}([a, b])$ has constant speed with regard to the Riemannian metric $g$ if the mapping $t \mapsto \|\dot{X}(t)\|_{X(t)} = \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))}$ is constant on $[a, b]$. Furthermore, if $X$ is only piecewise differentiable, we say that $X$ has constant speed if there exists $c > 0$ such that $\|\dot{X}(t)\|_{X(t)} = c$ for all $t \in (a, b)$ at which $X$ is differentiable.

**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.** See Lemma A.3 in the appendix.

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

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

**Proof.** See Lemma A.4 in the appendix.

To explicitly compute the geodesic distance, we will primarily search for length minimizers, i.e. curves $X \in \mathcal{A}_A^B$ which satisfy

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

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:

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

5
recall from Definition 1.1 that $\mathcal{A}_A^R([a,b])$ denotes the set of piecewise differentiable curves defined on $[a,b]$ connecting $A$ and $B$. Furthermore, if there exists a length minimizer in $\mathcal{A}_A^R$, then there exists one in $\mathcal{A}_A^R([a,b])$ 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 $X : [0,1] \to \text{GL}(n)$ and $\lambda > 0$, the energy of the curve $Y : [0,\lambda] \to \text{GL}(n), Y(t) = X(\frac{t}{\lambda})$ is

$$E(Y) = \int_0^\lambda g_{Y(t)}(\dot{Y}(t), \dot{Y}(t)) \, dt = \int_0^\lambda g_{X(\frac{t}{\lambda})}(\frac{1}{\lambda} \dot{X}(\frac{t}{\lambda}), \frac{1}{\lambda} \dot{X}(\frac{t}{\lambda})) \, dt$$

$$= \frac{1}{\lambda^2} \int_0^\lambda \dot{X}(\frac{t}{\lambda})(\dot{X}(\frac{t}{\lambda}), \dot{X}(\frac{t}{\lambda})) \, dt$$

$$= \frac{1}{\lambda} \int_0^1 g_{X(t)}(\dot{X}(t), \dot{X}(t)) \, dt = \frac{1}{\lambda} E(X). \quad (1.19)$$

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

$$\inf_{Y \in \mathcal{A}_A^R(A,B)} E(Y) = 0 \quad \forall A, B \in \text{GL}(n) : \det(AB) > 0. \quad (1.20)$$

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

$$X \in \mathcal{A}_A^R([a,b]) \quad \text{and} \quad E(X) = \inf_{Y \in \mathcal{A}_A^R([a,b])} E(Y). \quad (1.21)$$

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

$$L(X) \leq \sqrt{b-a} \sqrt{E(X)}, \quad (1.22)$$

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

**Proof.**

$$[L(X)]^2 = \left( \int_a^b \|\dot{X}(t)\|_{X(t)} \, dt \right)^2 = \left( \int_a^b 1 \cdot \|\dot{X}(t)\|_{X(t)} \, dt \right)^2 \leq \int_a^b 1 \, dt \cdot \int_a^b \|\dot{X}(t)\|_{X(t)}^2 \, dt = (b-a)E(\gamma). \quad (1.23)$$

The inequality in (1.23) 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 $\|\dot{X}\|_{X} \equiv \text{constant on } [a,b].$ 

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

i) $E(X) = \min_{Y \in \mathcal{A}_A^R([a,b])} E(Y)$,

ii) $L(X) = \min_{Y \in \mathcal{A}_A^R([a,b])} L(Y) = \min_{Y \in \mathcal{A}_A^R([a,b])} L(Y)$ and $X$ has constant speed.

**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 (1.22) from the previous lemma yields

$$E(X) \geq \frac{1}{b-a} [L(X)]^2 = \frac{1}{b-a} [L(X_c)]^2 = E(X_c), \quad (1.24)$$
where, again, 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, then the equivalence

$$E(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} E(Y) \iff L(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} L(Y)$$

(1.25)

holds. If we assume that $X = X_c$, then

$$E(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} E(Y)$$

(1.26)

$$\Rightarrow L(Y) = L(Y_c) = \sqrt{b-a} \sqrt{E(Y_c)} \geq \sqrt{b-a} \sqrt{E(X)} = L(X) \quad \forall Y \in A^{R}_{\mathcal{A}}([a,b])$$

$$\Rightarrow L(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} L(Y),$$

as well as

$$L(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} L(Y)$$

(1.27)

$$\Rightarrow E(Y) \geq \frac{1}{b-a} [L(Y)]^2 \geq \frac{1}{b-a} [L(X)]^2 = E(X) \quad \forall Y \in A^{R}_{\mathcal{A}}([a,b])$$

$$\Rightarrow E(X) = \min_{Y \in A^{R}_{\mathcal{A}}([a,b])} E(Y),$$

which concludes the proof. 

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

In the following, 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

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

(1.28)

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 $X \in A^{R}_{\mathcal{A}}$ connecting $A$ and $B$ we can define a curve $Y = CX \in A^{R(CB)}_{(C,A)}$ connecting $CA$ and $CB$ by $Y(t) = CX(t)$. We assume without loss of generality that $X$ is defined on the interval $[0,1]$ and find

$$L(Y) = \int_0^1 \sqrt{g_{r(C)}}(\dot{Y}(t),\dot{Y}(t)) \, dt$$

(1.29)

$$= \int_0^1 \sqrt{g_{C}}(CX(t),CX(t)) \, dt = \int_0^1 \sqrt{g_{X(t)}}(\dot{X}(t),\dot{X}(t)) \, dt = L(X).$$

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

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

(1.30)

In particular, this left-invariance of the geodesic distance implies

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

(1.31)
Lemma 1.11. Let $A Riemannian metric $A$ on $\mathbb{R}^3$ is completely determined by the geodesic distance on $GL^+(n)$ for $A, B \in GL^-(n)$, the geodesic distance on $GL^-(n)$ is completely determined by the geodesic distance on $GL^+(n)$ for left-$GL(n)$-invariant metrics.

**Definition 1.10.** A Riemannian metric $g$ on $GL(n)$ is called right-$O(n)$-invariant if 
\begin{equation}
g_{AQ}(MQ, NQ) = g_A(M, N)
\end{equation}
for all $A \in GL(n)$, $M, N \in \mathfrak{gl}(n)$ and $Q \in O(n)$.

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 satisfying 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{equation}
\langle M, N \rangle_{\mu, \mu_c, \kappa} = \mu \langle \text{dev} \text{ sym} M, \text{dev} \text{ sym} N \rangle + \mu_c \langle \text{skew} M, \text{skew} N \rangle + \frac{\kappa}{n} \text{tr}(M)(\text{tr} N)
\end{equation}
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 $\text{dev} M = M - \frac{1}{3n}(\text{tr} M)\mathbf{1}$ denotes the deviator of $M$. Here and throughout, $\langle M, N \rangle = \text{tr}(M^T N)$ denotes the canonical inner product on $\mathfrak{gl}(n)$.

We will state some basic properties of $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$. The proof can be found in the appendix (Lemma A.6).

**Lemma 1.11.** Let $M, N \in \mathfrak{gl}(n)$. Then
\begin{enumerate}
  \item $\langle Q^T M Q, Q^T N Q \rangle_{\mu, \mu_c, \kappa} = \langle M, N \rangle_{\mu, \mu_c, \kappa}$ for all $Q \in O(n)$,
  \item $\langle M, N \rangle_{\mu, \mu_c, \kappa} = \mu \langle \text{dev} \text{ sym} M, \text{dev} \text{ sym} N \rangle + \mu_c \langle \text{skew} M, \text{skew} N \rangle + \frac{\kappa}{n} \text{tr}(M) \text{tr} N$, \\
  \item $\langle M, N \rangle_{1,1,1} = \langle M, N \rangle$,
  \item $\langle S, W \rangle_{\mu, \mu_c, \kappa} = 0$ for all $S \in \text{Sym}(n), W \in \mathfrak{so}(n)$,
\end{enumerate}
where $\text{Sym}(n)$ and $\mathfrak{so}(n)$ denote the sets of symmetric and skew symmetric matrices in $\mathbb{R}^{n \times n}$ respectively.

In particular, (1.36) 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
\begin{equation}
\|M\|_{\mu, \mu_c, \kappa} = \sqrt{\langle M, M \rangle_{\mu, \mu_c, \kappa}} = \sqrt{\mu \|\text{dev} \text{ sym} M\|^2 + \mu_c \|\text{skew} M\|^2 + \frac{\kappa}{n} \text{tr}(M)^2}
\end{equation}
the norm induced by $\langle \cdot, \cdot \rangle_{\mu, \mu_c, \kappa}$. Here and throughout, $\|M\| = \sum_{i,j=1}^n M_{ij}^2$ denotes the canonical matrix norm.

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

**Proposition 1.12.**
\begin{enumerate}
  \item A Riemannian metric $g$ on $GL(n)$ is left-invariant if and only if $g$ is of the form 
  \begin{equation}
g_A(M, N) = \langle A^{-1} M, A^{-1} N \rangle_*,
\end{equation}
  where $\langle \cdot, \cdot \rangle_*$ is an inner product on the tangent space $\mathfrak{gl}(n) = T_\mathbf{1} GL(n)$ at the identity $\mathbf{1}$.
  \item A left-invariant metric $g$ is additionally right-$O(n)$-invariant if and only if $g$ is of the form 
  \begin{equation}
g_A(M, N) = \langle A^{-1} M, A^{-1} N \rangle_{\mu, \mu_c, \kappa}
\end{equation}
with $\mu, \mu_c, \kappa > 0$.
\end{enumerate}
Proof. i) Let \( \langle \cdot, \cdot \rangle_* \) be an inner product on \( \mathfrak{gl}(n) \) and \( g \) be defined by \( g_A(M, N) = \langle A^{-1}M, A^{-1}N \rangle_* \). Then we find

\[
g_{CA}(CM, CN) = \langle (CA)^{-1}CM, (CA)^{-1}CN \rangle_* \]
\[
= \langle A^{-1}C^{-1}CM, A^{-1}C^{-1}CN \rangle_* = \langle A^{-1}M, A^{-1}N \rangle_* = g_A(M, N),
\]

hence \( g \) is left-invariant.

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

\[
\langle M, N \rangle_* := g_A(M, N)
\]

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

\[
g_A(M, N) = g_{A1}(AA^{-1}M, AA^{-1}N) = g_A(A^{-1}M, A^{-1}N) = \langle A^{-1}M, A^{-1}N \rangle_*. \]

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

\[
g_{AQ}(MQ, NQ) = \langle (AQ)^{-1}MQ, (AQ)^{-1}NQ \rangle_{\mu, \mu, \kappa} = \langle Q^TA^{-1}MQ, Q^TA^{-1}NQ \rangle_{\mu, \mu, \kappa}.
\]

We apply the isotropy property \([1.34]\) to find

\[
\langle Q^TA^{-1}MQ, Q^TA^{-1}NQ \rangle_{\mu, \mu, \kappa} = \langle A^{-1}M, A^{-1}N \rangle_{\mu, \mu, \kappa} = g_A(M, N),
\]

which implies the right-\(O(n)\)-invariance of \( g \).

Finally, let \( g \) be an arbitrary left-invariant, right-\(O(n)\)-invariant metric on \( GL(n) \). Again we define the inner product \( \langle \cdot, \cdot \rangle_* \) by \( \langle M, N \rangle_* := g_A(M, N) \). Then for every \( Q \in O(n) \)

\[
\langle Q^TMQ, Q^TNQ \rangle_* = g_A(Q^TMQ, Q^TNQ)
\]
\[
= g_{AQ}(MQ, NQ) = g_{AQ}(MQ, NQ) = g_A(M, N) = \langle M, N \rangle_*.
\]

According to a well-known representation formula for isotropic linear mappings on \( \mathbb{R}^{n \times n} \) [6], this invariance directly implies that \( \langle \cdot, \cdot \rangle_* \) has the desired form \([1.33]\). \( \blacksquare \)

In the following sections \( g \) will denote a left-invariant, right-\(O(n)\)-invariant Riemannian metric as given in \([1.40]\), 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 [4, Corollary 3.9]. 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 classical 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. For any continuously differentiable function \( f \in C^1(\text{GL}(n); \text{GL}(n)) \), we denote by \( Df[A] \) the total derivative of \( f \) at \( A \in \text{GL}(n) \), and \( Df[A]H \in \mathfrak{gl}(n) \) denotes its directional derivative at \( A \) in direction \( H \).

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

\[
Df[A].H = -A^{-1}HA^{-1}.
\]
Proof. A short proof can be found in [20]. ■

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

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

Proof. Direct computation yields

$$\langle M, PN \rangle = \text{tr}((PN)^T M) = \text{tr}(N^T(P^T M)) = \langle P^T M, N \rangle,$$

$$\langle M, NP \rangle = \text{tr}((NP)^T M) = \text{tr}(P^T N^T M) = \text{tr}(N^T(MP^T)) = \langle MP^T, N \rangle.$$ \hfill ■

### 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 $\mathfrak{gl}(n) = \mathbb{R}^{n \times n}$, i.e. that $\mu = \mu_c = \kappa = 1$. Let $X \in A^2_1 \cap C^2([0, 1]; GL^+(n))$ be a two-times differentiable length minimizing curve from $[0, 1]$ to $GL^+(n)$, where the length is measured by the left-invariant Riemannian metric $g$.

$$g_Z(M, N) := \langle Z^{-1} M, Z^{-1} N \rangle = \text{tr}((Z^{-1} M)^T Z^{-1} N),$$

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

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

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 $Z_h := X + h V$ for $h \in (-\varepsilon, \varepsilon)$. For sufficiently small $\varepsilon$, we have $Z_h(t) \in GL^+(n)$ for all $t \in [0, 1]$ and therefore $Z_h \in A^2_1$ (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

$$0 = \frac{1}{2} \frac{d}{dh} E(X + h V)_{|h=0} = \frac{1}{2} \frac{d}{dh} \int_0^1 \langle Z^{-1}_h \dot{Z}_h, Z^{-1}_h \dot{Z}_h \rangle_{|h=0} dt$$

$$= \frac{1}{2} \int_0^1 \frac{d}{dh} \langle Z^{-1}_h \dot{Z}_h, Z^{-1}_h \dot{Z}_h \rangle_{|h=0} dt$$

$$= \int_0^1 \langle Z^{-1}_h \dot{Z}_h_{|h=0}, \frac{d}{dh} Z^{-1}_h \dot{Z}_h_{|h=0} \rangle dt.$$ \hfill (2.4)

Since $Z^{-1}_h_{|h=0} = X^{-1}$ and $\dot{Z}_h_{|h=0} = \dot{X}$, we can use Lemma 2.1 to find

$$\frac{d}{dh} Z^{-1}_{|h=0} = \frac{d}{dh} (X + h V)^{-1}_{|h=0} = -(X + h V)^{-1} V (X + h V)^{-1}_{|h=0} = X^{-1} V X^{-1}. \quad \text{(2.4)}$$

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

$$\int_0^1 \langle X^{-1} \dot{X}, -X^{-1} V X^{-1} \dot{X} + X^{-1} V \rangle dt = \int_0^1 -\langle X^{-T} X^{-1} \dot{X}, V X^{-1} \dot{X} \rangle + \langle X^{-T} X^{-1} \dot{X}, V \rangle dt$$

$$= \int_0^1 -\langle X^{-T} X^{-1} \dot{X}, X^{-1} \dot{X} V \rangle + \langle \frac{d}{dt} (X^{-T} X^{-1} \dot{X}), V \rangle dt$$

$$= -\int_0^1 \langle X^{-T} X^{-1} \dot{X} (X^{-1} \dot{X})^T + \frac{d}{dt} (X^{-T} X^{-1} \dot{X}), V \rangle dt.$$
Since this equation holds for all $V \in C^1_0([0, 1]; \mathfrak{gl}(n))$, 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 as well as, again, Lemma 2.1):

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

$$= -X^{-T}\dot{X}^T X^{-T}X^{-1}\dot{X} + X^{-T}(-X^{-1}\dot{X}X^{-1}\dot{X} + X^{-1}\ddot{X})$$

$$= X^{-T}(-\ddot{X}X^{-T}X^{-1}\dot{X} - X^{-1}\dot{X}X^{-1}\dot{X} + X^{-1}\ddot{X}). \quad (2.6)$$

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

$$-X^{-T}X^{-1}\dot{X}X^{-1}\dot{X} + X^{-T}(\dot{X}^T X^{-1}\dot{X} - X^{-1}\dot{X}X^{-1}\dot{X} + X^{-1}\ddot{X}) \quad \Leftrightarrow \quad -UU = -UU^T - UU + X^{-1}\ddot{X}$$

$$\Leftrightarrow \quad U^T U - UU^T = -UU + X^{-1}\ddot{X} \quad \Leftrightarrow \quad \dot{U} = U^T U - UU^T. \quad (2.8)$$

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. 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 $\mu, \mu_v, \kappa > 0$. In order to find the geodesic equation for this case, we need the following lemma.

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

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

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

$$\text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0. \quad (2.10)$$

Therefore

$$\langle \text{sym} A, \text{sym}(AB - BA) \rangle = \langle \text{dev sym} A, \text{dev sym}(AB - BA) \rangle + \frac{1}{n} \text{tr} A \text{ tr}(AB - BA)$$

$$= \langle \text{dev sym} A, \text{dev sym}(AB - BA) \rangle \quad (2.11)$$

and

$$\langle \text{sym} A, \text{sym}(AB - BA) \rangle - \langle \text{skew} A, \text{skew}(AB - BA) \rangle$$

$$= \langle \text{sym} A, AB - BA \rangle - \langle \text{skew} A, AB - BA \rangle$$

$$= \langle \text{sym} A - \text{skew} A, AB - BA \rangle = \langle A^T, AB - BA \rangle = \text{tr}(AAB) - \text{tr}(ABA) = 0, \quad (2.12)$$
thus

\[ \langle A, [A, B] \rangle_{\mu, \mu, \kappa} = \langle A, (AB - BA) \rangle_{\mu, \mu, \kappa} \]
\[ \mu(\text{dev sym } A, \text{dev sym } (AB - BA)) + \mu_c(\text{skew } A, \text{skew } (AB - BA)), \]
\[ \mu(\text{sym } A, \text{sym } (AB - BA)) + \mu_c(\text{skew } A, \text{skew } (AB - BA)). \]
\[ = \langle A, (AB - BA) \rangle, \]

which completes the proof.

Let \( X \in A^B_A \) be a piecewise two-times differentiable energy minimizer with regard to the Riemannian metric

\[ g_Z(X, Y) := \langle Z^{-1}X, Z^{-1}Y \rangle_{\mu, \mu, \kappa}, \]
\[ \langle A, B \rangle_{\mu, \mu, \kappa} := \mu(\text{dev sym } A, \text{dev sym } B) + \mu_c(\text{skew } A, \text{skew } B) + \frac{\kappa}{n} \text{tr } A \text{ tr } B. \]  

(2.13)

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

\[ 0 = \frac{1}{2} \frac{d^2}{dh^2} E(Y_h) \big|_{h=0} \]

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

\[ Y_h(U + h(\dot{X}hV - hXV')) - h^2XV(\dot{X}hV - hXV')) = Y_h(U + hC) - h^2XVC \]
\[ = (X + hXV)(U + hC) - h^2XVC \]
\[ = XU + hXUV + hXC + h^2XVC - h^2XVC \]
\[ = X + hXUV + hXUV - hXVU + hXV \]
\[ = \dot{X} + hXUV + hXV = X + hXV = \dot{Y}_h. \]

(2.15)

We therefore obtain

\[ U_h = Y_h^{-1} \dot{Y}_h \]
\[ = Y_h^{-1}(Y_h(U + hC) - h^2XVC) \]
\[ = U + hC - h^2Y_h^{-1}XVC \]
\[ = U + h(UV - XU + V) - h^2Y_h^{-1}XVC, \]

(2.16)

(2.17)

which implies

\[ U_h \big|_{h=0} = U, \quad \frac{d}{dh} U_h \big|_{h=0} = UV - XU + V. \]

(2.18)
Coming back to equation (2.14), we find that
\[
0 = \frac{1}{2} \frac{d}{dh} E(Y_h)\Big|_{h=0} = \frac{1}{2} \int_0^1 \frac{d}{dh} \langle Y_h^{-1} \dot{Y}_h, Y_h^{-1} \dot{Y}_h \rangle_{\mu,\mu,\kappa} \big|_{h=0} dt
\]
\[
= \frac{1}{2} \int_0^1 \frac{d}{dh} \langle U_h, U_h \rangle_{\mu,\mu,\kappa} \big|_{h=0} dt
\]
\[
= \int_0^1 \langle U_h, \frac{d}{dh} U_h \rangle_{\mu,\mu,\kappa} \big|_{h=0} dt \tag{2.18}
\]
\[
= \int_0^1 \langle U, UV - VU + \dot{V} \rangle_{\mu,\mu,\kappa} dt. \tag{2.19}
\]
Using Lemma 2.3 we can write (2.19) as
\[
\int_0^1 \langle U, \dot{V} \rangle_{\mu,\mu,\kappa} + \langle U, UV - VU \rangle_{\mu,\mu,\kappa} dt
\]
\[
= \int_0^1 \langle U, \dot{V} \rangle_{\mu,\mu,\kappa} + \frac{\mu + \mu c}{2} \langle U, UV - VU \rangle dt
\]
\[
= \int_0^1 \langle U, \dot{V} \rangle_{\mu,\mu,\kappa} + \frac{\mu + \mu c}{2} \langle V, U^T U - UU^T \rangle dt \tag{2.20}
\]
\[
= \sum_{k=0}^{m} \int_{a_k}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu,\mu,\kappa} + \frac{\mu + \mu c}{2} \langle V, U^T U - UU^T \rangle_{\mu,\mu,\kappa} dt,
\]
where \( a = a_0 < a_1 < \cdots < a_{m+1} = b \) are chosen such that \( X|_{[a_j, a_{j+1}]} \in C^2([a_j, a_{j+1}]; \text{GL}(n)) \) for all \( j \in \{0, \ldots, m\} \). For \( V \in C^1([0,1]; \mathfrak{gl}(n)) \) with \( V(a_k) = 0 \) for \( k \in \{1, \ldots, m\} \), integration by parts yields
\[
0 = \sum_{k=0}^{m} \int_{a_k}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu,\mu,\kappa} + \frac{\mu + \mu c}{2} \langle V, U^T U - UU^T \rangle_{\mu,\mu,\kappa} dt
\]
\[
= \sum_{k=0}^{m} \int_{a_k}^{a_{k+1}} -\langle V, \dot{U} \rangle_{\mu,\mu,\kappa} + \frac{\mu + \mu c}{2} \langle V, U^T U - UU^T \rangle_{\mu,\mu,\kappa} dt. \tag{2.21}
\]
We can now apply the fundamental lemma of the calculus of variations to find
\[
\dot{U}(t) = \frac{1}{2} \frac{\mu}{\mu} (U(t)^T U(t) - U(t) U(t)^T) \quad \forall t \in [0,1] \setminus \{a_1, \ldots, a_m\}. \tag{2.22}
\]
To show that \( X \) is two-times differentiable on \([0,1]\) and that equation (2.22) is satisfied everywhere, we first show that \( U = X^{-1} X \) is continuous at each \( a_k \) (and therefore on the whole interval \([0,1]\)). We choose
\[ V \in C^0_t((a_{k-1}, a_{k+1}); \mathfrak{gl}(n)) \text{ for } k \in \{1, \ldots, m\} \text{ and compute} \]
\[
\begin{align*}
0 &= \int_0^{a_k} \langle U, \dot{V} \rangle_{\mu,\nu,\kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu,\nu,\kappa} \, dt \\
&= \int_{a_{k-1}}^{a_k} \langle U, \dot{V} \rangle_{\mu,\nu,\kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu,\nu,\kappa} \, dt \\
&\quad + \int_{a_k}^{a_{k+1}} \langle U, \dot{V} \rangle_{\mu,\nu,\kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu,\nu,\kappa} \, dt \\
&= \int_{a_{k-1}}^{a_k} -\langle V, \dot{U} \rangle_{\mu,\nu,\kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu,\nu,\kappa} \, dt + \langle [V(t), U(t)]_{\mu,\nu,\kappa} \rangle_{t=a_{k-1}}^{a_k} \\
&\quad + \int_{a_k}^{a_{k+1}} -\langle V, \dot{U} \rangle_{\mu,\nu,\kappa} + \frac{\mu + \mu_c}{2\mu} \langle V, U^T U - U U^T \rangle_{\mu,\nu,\kappa} \, dt + \langle [V(t), U(t)]_{\mu,\nu,\kappa} \rangle_{t=a_k}^{a_{k+1}}. \tag{2.23}
\end{align*}
\]

According to (2.22) 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 (2.23) are zero and we find
\[
\begin{align*}
0 &= \langle [V(t), U(t)]_{\mu,\nu,\kappa} \rangle_{t=a_{k-1}}^{a_k} + \langle [V(t), U(t)]_{\mu,\nu,\kappa} \rangle_{t=a_k}^{a_{k+1}} \\
&= \lim_{t \searrow a_k} \langle V(t), U(t) \rangle_{\mu,\nu,\kappa} - \lim_{t \nearrow a_k} \langle V(t), U(t) \rangle_{\mu,\nu,\kappa}. \tag{2.24}
\end{align*}
\]
Since \(V\) can be chosen with arbitrary values for \(V(a_k)\), we find
\[
\lim_{t \searrow a_k} U(t) = \lim_{t \nearrow a_k} U(t). \tag{2.25}
\]
Therefore \(U\) is continuous on \([0, 1]\). But then (2.22) 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 \(X \in C^2([0, 1]; \mathrm{GL}^+(n))\), and \(U = X^{-1} \dot{X}\) satisfies the \textit{geodesic equation}
\[
\dot{U} = \frac{1 + \mu_c}{2 \mu} (U^T U - U U^T) \tag{2.26}
\]
everywhere on \([0, 1]\). The results of Section 2.2 can be summarized as follows:

**Proposition 2.4.** Let \(A, B \in \mathrm{GL}^+(n)\) and let \(X \in A^B_\kappa\) be a piecewise two-times differentiable energy minimizer with regard to the left-invariant Riemannian metric induced by the inner product \(\langle \cdot, \cdot \rangle_{\mu,\nu,\kappa}\). Then \(X \in C^2([a, b]; \mathrm{GL}^+(n))\), and \(U = X^{-1} \dot{X}\) is a solution to (2.26).

**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 \textit{length minimizer} has a reparameterization that satisfies the geodesic equation.

**Remark 2.6.** Equation (2.26) could also be deduced from a more general formula given by Bloch et al. [3], which is derived via the Euler-Poincaré equations corresponding to the length minimization problem [10, p. 430ff].
2.3 Properties of solutions

Some properties of the solutions \(U \in C^{1}([0,1]; \mathfrak{gl}(n))\) to (2.26) can be inferred directly from the equation.

Lemma 2.7. Let \(U\) be a solution to (2.26). Then:

i) \[\|U\|_{\mu,\mu,\kappa} \equiv \text{constant},\] (2.27)

ii) \[\text{tr}(U) \equiv \text{constant},\] (2.28)

iii) \[\text{det}(U) \equiv \text{constant},\] (2.29)

iv) \[\text{tr}(\text{Cof} U) \equiv \text{constant},\] (2.30)

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

Proof. i)

\[
\frac{d}{dt} \|U\|_{\mu,\mu,\kappa}^2 = \frac{d}{dt} \langle U, U \rangle_{\mu,\mu,\kappa} = 2 \langle U, U \rangle_{\mu,\mu,\kappa} = (1 + \frac{\mu_c}{\mu})(U, U^T U - U U^T)_{\mu,\mu,\kappa}
\]

\[= (1 + \frac{\mu_c}{\mu})\left(\mu(\text{sym} U, U^T U - U U^T)_{\mu,\mu,\kappa} + \mu_c(\text{skew} U, U^T U - U U^T)_{\mu,\mu,\kappa}\right)
\]

\[+ \frac{\mu}{n} \text{tr}(U^T U - U U^T)_{\mu,\mu,\kappa} = 0.
\]

(2.31)

ii)

\[
\frac{d}{dt} \text{tr} U = \text{tr} \dot{U} = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^T U - U U^T) = 0.
\]

(2.32)

iii)

\[
\frac{d}{dt} \text{det} U = \text{tr}(\text{Adj}(U) \dot{U}) = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\text{Adj}(U)(U^T U - U U^T))
\]

\[= \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\text{Adj}(U) U^T U - \text{Adj}(U) U U^T)
\]

\[= \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^T U \text{Adj}(U) - \text{det}(U) U^T) = \frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(\text{det}(U) U^T - \text{det}(U) U T) = 0,
\]

(2.33)

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

iv) We first assume that \(U\) is invertible. Then

\[
\frac{d}{dt} \text{tr} U^{-1} = -\text{tr}(U^{-1} \dot{U} U^{-1})
\]

\[= -\frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^{-1}(U^T U - U U^T) U^{-1}) = -\frac{1 + \frac{\mu_c}{\mu}}{2} \text{tr}(U^{-1} U^T U^T - U U^T U^{-1}) = 0,
\]

(2.34)

and since \(\text{det}(U)\) is constant as well, we find

\[\text{tr} \text{Cof}(U) = \text{det}(U) \cdot \text{tr} U^{-T} = \text{det}(U) \cdot \text{tr}(U^{-1}) \equiv \text{constant}.
\]

The general case follows directly from the density of \(\text{GL}(n)\) in \(\mathfrak{gl}(n) = \mathbb{R}^{n \times n}\). ■
2.4 Existence and uniqueness of solutions

Since the mapping $U \mapsto \frac{1 + \frac{\mu c}{2}}{2} (U^T U - U U^T)$ is locally Lipschitz continuous, the Picard–Lindelöf theorem implies that equation (2.26) has a unique local solution for given $U(0)$ and that if a global solution exists, it is uniquely defined. Similarly, for given $U \in C^1(I, \mathfrak{gl}(n))$ on an arbitrary interval $I$, the initial value problem

$$\begin{cases}
\dot{X} = XU, \\
X(0) = X_0
\end{cases} \tag{2.35}$$

has a unique local solution as well. Using estimates (2.27) to (2.29), we can even show that this solution is globally defined:

**Proposition 2.8.** For $X_0 \in \text{GL}^+(n)$, $M_0 \in \mathfrak{gl}(n)$, the initial value problem

$$\begin{cases}
U = X^{-1} \dot{X}, \\
\dot{U} = \frac{1 + \frac{\mu c}{2}}{2} (U^T U - U U^T), \\
X(0) = X_0, \quad \dot{X}(0) = M_0
\end{cases} \tag{2.36}$$

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

**Proof.** Since $U(0) = X(0)^{-1} \dot{X}(0)$ is determined by $X(0)$ and $\dot{X}(0)$, (2.36) has a unique local solution, as demonstrated above. Then, according to (2.27), $\|U\|_{\mu', \mu, \kappa}$ 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 mapping $X \mapsto UX$ is globally Lipschitz continuous. Therefore (after rewriting (2.36) as $\dot{X} = UX$) there exists a unique global solution $X : \mathbb{R} \to \mathbb{R}^{n \times n}$ to (2.36).

It remains to show that $X(t) \in \text{GL}^+(n)$ for all $t \in \mathbb{R}$. First we observe that, for $\det X > 0$,

$$\frac{d}{dt} (\det X) = \text{tr}(\text{Adj}(X) \dot{X}) = \text{tr}(\det(X) X^{-1} \dot{X}) = \det(X) \text{tr} U \cdot \det(X) \text{tr}(U(0)). \tag{2.37}$$

Thus 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

$$\begin{cases}
\frac{d}{dt} (\det X(t)) = \text{tr}(U(0)) \det X(t), \\
\det X(0) = \det X_0.
\end{cases} \tag{2.38}$$

Assume now that $X(\mathbb{R}) \notin \text{GL}^+(n)$. Since $I^+ := \{ t \in \mathbb{R} \mid \det X(t) > 0 \}$ is open and $0 \in I^+$ by definition of $X_0$, there is a minimal $t_0 > 0$ with $-t_0 \notin I^+$ or $t_0 \notin I^+$. But then $\det X$ is the (unique) solution to (2.38), i.e.

$$\det X(t) = (\det X_0) e^{t \cdot \text{tr}(U(0))}. \tag{2.39}$$

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

$$0 < \lim_{t \to t_0^-} \det X(t) = \det X(-t_0), \quad 0 < \lim_{t \to t_0^+} \det X(t) = \det X(t_0),$$

in contradiction to the definition of $t_0$. \[\square\]
3 The parameterization of geodesics

3.1 Solving the geodesic equation

We will now give an explicit solution to the geodesic initial value problem \((2.36)\). The result is inspired by the work of Mielke [18], who obtained formula \((3.3)\) as a parametrization of the geodesics on \(SL(n)\). For the canonical inner product, i.e. for the special case \(\mu = \mu_\kappa = \kappa = 1\), 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 \(SO(n)\) with respect to the canonical metric [19]), 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 \(A \in \text{GL}^+(n)\), \(M \in \text{gl}(n)\) and \(\omega = \frac{\mu}{\rho}\). We define

\[
\Phi(M) := \exp(\text{sym} \ M - \omega \text{ skew} \ M) \exp((1 + \omega) \text{ skew} \ M),
\]

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

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

where \(\exp: \text{gl}(n) \rightarrow \text{GL}^+(n)\) denotes the matrix exponential (cf. Appendix A). We use equation \((A.12)\) to compute

\[
\dot{X} = A \exp(t(\text{sym} \ M - \omega \text{ skew} \ M)) (\text{sym} \ M - \omega \text{ skew} \ M) \exp(t(1 + \omega) \text{ skew} \ M) + A \exp(t(\text{sym} \ M - \omega \text{ skew} \ M)) (1 + \omega) \text{ skew} \ M \exp(t(1 + \omega) \text{ skew} \ M)
\]

\[
= A \exp(t(\text{sym} \ M - \omega \text{ skew} \ M)) (\text{sym} \ M + \text{ skew} \ M) \exp(t(1 + \omega) \text{ skew} \ M)
\]

\[
= A \exp(t(\text{sym} \ M - \omega \text{ skew} \ M)) M \exp(t(1 + \omega) \text{ skew} \ M)
\]

\[
(3.4)
\]

and

\[
U := X^{-1} \dot{X} = \exp(t(1 + \omega) \text{ skew} \ M)^{-1} \exp(t(\text{sym} \ M - \omega \text{ skew} \ M))^{-1} A^{-1} \cdot A \exp(t(\text{sym} \ M - \omega \text{ skew} \ M)) M \exp(t(1 + \omega) \text{ skew} \ M)
\]

\[
= \exp(t(1 + \omega) \text{ skew} \ M)^{-1} M \underbrace{\exp(t(1 + \omega) \text{ skew} \ M)}_{\in \text{so}(n)}^{\in \text{so}(n)} = q^T MQ,
\]

\[
(3.5)
\]

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

\[
U^T U - UU^T = Q^T M^T MQ - Q^T M M^T Q = Q^T (M^T M - MM^T) Q
\]

\[
(3.6)
\]

and

\[
\dot{U} = \dot{Q}^T MQ + Q^T M \dot{Q}^{\text{A.12}} = Q^T (1 + \omega)(\text{sym} \ M)^T MQ + Q^T (1 + \omega) M(\text{skew} \ M) Q
\]

\[
= (1 + \omega)Q^T ((\text{skew} \ M)^T M + M(\text{skew} \ M)) Q
\]

\[
= (1 + \omega)Q^T (M(\text{skew} \ M) - (\text{skew} \ M) M) Q
\]

\[
= (1 + \omega)Q^T (\frac{1}{2}(MM - MM^T) - \frac{1}{2}(MM - MTM)) Q
\]

\[
= \frac{(1 + \omega)}{2} Q^T (M^T M - MM^T) Q = \frac{(1 + \omega)}{2} (U^T U - UU^T),
\]

\[
(3.7)
\]

and thus \(X\) solves the geodesic equation. To solve the initial value problem \((2.36)\) we use the equality \(\exp(0) = 1\) as well as equation \((3.4)\) to obtain

\[
X(0) = A, \quad \dot{X}(0) = AM,
\]

\[
(3.8)
\]

hence we can solve \((2.36)\) by choosing \(A = X_0\), \(M = X_0^{-1} M_0\). We conclude:
Theorem 3.1. Let $X_0 \in \operatorname{GL}^+(n)$, $M_0 \in \mathfrak{gl}(n)$. Then the curve $X : \mathbb{R} \to \operatorname{GL}^+(n)$ with

$$X(t) = \Phi_{X_0}(t X_0^{-1} M_0) = X_0 \exp(t \langle \text{sym}(X_0^{-1} M_0) - \mu \mu^T (X_0^{-1} M_0) \rangle \mu, \mu, \kappa) \exp(t(1 + \mu \mu^T) \langle \text{skew}(X_0^{-1} M_0) \rangle \mu, \mu, \kappa)$$

is the unique global solution to the geodesic initial value problem

$$\begin{align*}
U &= X^{-1} \dot{X}, \\
\dot{U} &= \frac{1 + \mu_0}{2} \langle U^T U - U U^T \rangle, \\
X(0) &= X_0, \quad \dot{X}(0) = M_0.
\end{align*}$$

Because of the left-invariance of $g$, we can mostly focus on the case $A = 1$, $\Phi_A = \Phi = \Phi_1$ and

$$X(t) = \exp(t \langle \text{sym} M - \omega \text{skew} M \rangle \mu, \mu, \kappa) \exp(t(1 + \omega) \langle \text{skew} M \rangle \mu, \mu, \kappa).$$

3.2 Properties of geodesic curves

Since the geodesic curves

$$X(t) = A \exp(t \langle \text{sym} M - \omega \text{skew} M \rangle \mu, \mu, \kappa) \exp(t(1 + \omega) \langle \text{skew} M \rangle \mu, \mu, \kappa)$$

solve the geodesic initial value problem (2.36), the properties given in Lemma 2.7 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 $Q(t) \in \operatorname{SO}(n)$):

$$L(X) = \int_0^{t_0} g_X(X, \dot{X}) \, dt = \int_0^{t_0} \sqrt{\langle X^{-1} \dot{X}, X^{-1} \dot{X} \rangle \mu, \mu, \kappa} \, dt = \int_0^{t_0} \sqrt{\langle U, U \rangle \mu, \mu, \kappa} \, dt = \int_0^{t_0} \sqrt{\langle Q^T M Q, Q^T M Q \rangle \mu, \mu, \kappa} \, dt = \int_0^{t_0} \sqrt{\langle M, M \rangle \mu, \mu, \kappa} \, dt = t_0 \| M \|_{\mu, \mu, \kappa},$$

where the second to last equality follows from the isotropy property (1.34) of the inner product $\langle \cdot, \cdot \rangle_{\mu, \mu, \kappa}$.

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).

At this point, it is not clear that, for given $A, B \in \operatorname{GL}^+(n)$, 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]. Furthermore, the global existence of the geodesic curves demonstrated in Proposition 2.8 implies that $\operatorname{GL}(n)$ is geodesically complete with respect to $g$, and thus the Hopf-Rinow theorem also guarantees the global existence of length minimizers. Nonetheless, in our effort to keep this paper self contained, we will give a full proof for the existence of minimizers using only results from basic real analysis. Towards this aim, we will first show that a specific variant of Gauss’s lemma [8, p. 2.93] holds for $g$ on $\operatorname{GL}(n)$ and continue by following the proof of the Hopf-Rinow theorem as given in [12, Theorem 1.7.1] (a similar proof can be found in [4, Theorem 2.8]) as applied to our special case. Readers not interested in these rather basic proofs should skip the main part of Section 4 and continue with Theorem 4.7.
4.1 Preparations

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

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

(4.1)

can be considered the geodesic exponential at $A$ [12] Definition 1.4.3. Note again that from this point of view the following lemma is a direct application of Gauss’s lemma. We will prove it via direct computation.

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

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

(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 $g_{\Phi_A(M)}(D\Phi_A[M].S, D\Phi_A[M].T) = \langle S, T \rangle_{\mu, \mu, \kappa}$ does not hold for arbitrary $S, T \in \mathfrak{gl}(n)$.

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

$$g_{\Phi_A(M)}(D\Phi_A[M].M, D\Phi_A[M].T) = g_{\Phi(M)}(AD\Phi[M].M, AD\Phi[M].T)$$

$$= g_{\Phi(M)}(D\Phi[M].M, D\Phi[M].T).$$

(4.3)

Using [A.12] from the appendix, we compute

$$D\Phi[M].M = (D \exp[\text{sym } M - \omega \text{ skew } M].(\text{sym } M - \omega \text{ skew } M)) \exp((1 + \omega) \text{ skew } M)$$

$$+ \exp(\text{sym } M - \omega \text{ skew } M) (D \exp[(1 + \omega) \text{ skew } M].((1 + \omega) \text{ skew } M))$$

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

$$+ \exp(\text{sym } M - \omega \text{ skew } M) ((1 + \omega) \text{ skew } M) \exp((1 + \omega) \text{ skew } M)$$

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

$$= \exp(\text{sym } M - \omega \text{ skew } M) M Q,$$

(4.4)

while [A.11] yields

$$D\Phi[M].T$$

$$= (D \exp[\text{sym } M - \omega \text{ skew } M].(\text{sym } T - \omega \text{ skew } T)) \exp((1 + \omega) \text{ skew } M)$$

$$+ \exp(\text{sym } M - \omega \text{ skew } M) (D \exp[(1 + \omega) \text{ skew } M].(1 + \omega) \text{ skew } T)$$

$$= \int_0^1 \exp(s(\text{sym } M - \omega \text{ skew } M)) (\text{sym } T - \omega \text{ skew } T) \exp((1 - s)(\text{sym } M - \omega \text{ skew } M)) \exp((1 + \omega) \text{ skew } M) ds$$

$$+ \int_0^1 \exp(\text{sym } M - \omega \text{ skew } M) \exp(s(1 + \omega) \text{ skew } M) ((1 + \omega) \text{ skew } T) \exp((1 - s)(1 + \omega) \text{ skew } M) ds$$

$$= \exp(\text{sym } M - \omega \text{ skew } M) \left( \int_0^1 \exp((s - 1)(\text{sym } M - \omega \text{ skew } M)) (\text{sym } T - \omega \text{ skew } T) \exp((1 - s)(\text{sym } M - \omega \text{ skew } M)) \right.$$

$$+ \exp(s(1 + \omega) \text{ skew } M) ((1 + \omega) \text{ skew } T) \exp(-s(1 + \omega) \text{ skew } M) \exp((1 + \omega) \text{ skew } M)$$

$$= \exp(\text{sym } M - \omega \text{ skew } M) \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T) P_s^{-1} + (1 + \omega) R_s^T (\text{skew } T) R_s] ds \cdot \exp((1 + \omega) \text{ skew } M).$$

(4.5)

We also find

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

(4.6)
Hence we can compute
\[ g_{\Phi(M)}(D\Phi[M], M, D\Phi[M], T) = (\Phi(M)^{-1} D\Phi[M], M, \Phi(M)^{-1} D\Phi[M], T)_{\mu, \mu, n} \]
\[ = (Q^T MQ, Q^T \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s] \, ds \cdot Q)_{\mu, \mu, n} \]
\[ = (M, \int_0^1 [P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s] \, ds)_{\mu, \mu, n} \]
\[ = \int_0^1 (M, P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1} + (1 + \omega)R_s^T(\text{skew } T)R_s)_{\mu, \mu, n} \, ds \]
\[ = \int_0^1 (\mu \text{ dev } \text{sym } M + \mu_c \text{ skew } M, P_s(\text{sym } T - \omega \text{ skew } T)P_s^{-1}) \]
\[ + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) + (1 + \omega) \mu_c \langle \text{skew } M, R_s \rangle \, ds \]
\[ = \int_0^1 \mu (P_s^T (\text{dev } \text{sym } M + \omega \text{ skew } M)P_s^{-T}, \text{sym } T - \omega \text{ skew } T) \]
\[ + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) + (1 + \omega) \mu_c \langle \text{skew } M \text{ skew } M^T R_s, \text{skew } T \rangle \, ds . \]  

(4.7)

It is not difficult to see that the matrices
\[ \text{dev } \text{sym } M + \omega \text{ skew } M, \quad \text{sym } M + \omega \text{ skew } M \]
commute. Therefore, since \( P_s^T = \exp((s - 1)(\text{sym } M + \omega \text{ skew } M)) \), 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
\[ \int_0^1 \mu ((\text{dev } \text{sym } M + \omega \text{ skew } M)P_s^T P_s^{-T}, \text{sym } T - \omega \text{ skew } T) \]
\[ + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) + (1 + \omega) \mu_c \langle \text{skew } M \text{ skew } M^T R_s, \text{skew } T \rangle \, ds \]
\[ = \int_0^1 \mu (\text{dev } \text{sym } M + \omega \text{ skew } M, \text{sym } T - \omega \text{ skew } T) \]
\[ + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) + (1 + \omega) \mu_c \langle \text{skew } M, \text{skew } T \rangle \, ds \]
\[ = \int_0^1 \mu (\text{dev } \text{sym } M, \text{dev } \text{sym } T) + \mu_c \langle \text{skew } M, -\omega \text{ skew } T \rangle \]
\[ + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) + (\mu_c + \omega \mu_c) \langle \text{skew } M, \text{skew } T \rangle \, ds \]
\[ = \int_0^1 \mu (\text{dev } \text{sym } M, \text{dev } \text{sym } T) + \mu_c \langle \text{skew } M, \text{skew } T \rangle + \frac{\kappa}{n} (\text{tr } M) (\text{tr } T) \, ds = \langle M, T \rangle_{\mu, \mu, n} . \]

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 \( \Phi_A \) 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], \( \Phi_A(0) \) 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 \( \gamma \) 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 [7].
Lemma 4.3. Let $\gamma \in C^0([0, 1]; B_c(0))$ be a piecewise differentiable curve in $B_c(0) \subseteq \mathfrak{gl}(n)$ with $\gamma(0) = 0$, $\gamma(1) = M$. Then

\begin{align}
L(\Phi_A \circ \gamma) & \geq \|M\|_{\mu, \mu_c, \kappa}, \\
E(\Phi_A \circ \gamma) & \geq \|M\|_{\mu, \mu_c, \kappa}.
\end{align}

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

\begin{align}
L(\Phi_A \circ \gamma) &= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(d\Phi_A(\gamma)(t), d\Phi_A(\gamma)(t))} \, dt \\
&= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], \gamma, D\Phi_A[\gamma], \gamma)} \, dt.
\end{align}

Note that since $\gamma$ and $\dot{\gamma}$ are generally not equal, the previous lemma cannot be applied directly. Therefore we decompose $\dot{\gamma}$ into the sum of $\xi_1$ and $\xi_2$, where $\xi_1$ is a multiple of $\gamma$ and $\xi_2$ is orthogonal to $\gamma$ with regard to the inner product on $\mathfrak{gl}(n)$: define $r, \xi_1, \xi_2$ by

\begin{align}
r(t) &:= \begin{cases} 
\frac{(\gamma(t), \gamma)}{\|\gamma(t)\|_{\mu, \mu_c, \kappa}} : & \gamma(t) \neq 0 \\
0 & : \gamma(t) = 0 \end{cases}, \\
\xi_1(t) &:= r(t) \gamma(t), \\
\xi_2(t) &:= \gamma(t) - \xi_1(t).
\end{align}

Then

\begin{align}
\dot{\gamma} = \xi_1 + \xi_2, \quad \langle \gamma, \xi_2 \rangle_{\mu, \mu_c, \kappa} = 0
\end{align}

and thus

\begin{align}
L(\Phi_A \circ \gamma) &= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], \gamma, D\Phi_A[\gamma], \gamma)} \, dt \\
&= \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], (\xi_1 + \xi_2), D\Phi_A[\gamma], (\xi_1 + \xi_2))} \, dt \\
&\geq \int_0^1 \sqrt{g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], (\xi_1 + \xi_2), D\Phi_A[\gamma], (\xi_1 + \xi_2))} \, dt \\
&\geq \int_0^1 \sqrt{r(t)^2 g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], \gamma, D\Phi_A[\gamma], \gamma) + 2 r(t) g_{\Phi_A(\gamma)}(D\Phi_A[\gamma], \gamma, D\Phi_A[\gamma], \gamma)} \, dt \\
&\geq \int_0^1 \sqrt{r(t)^2 \langle \gamma, \gamma \rangle_{\mu, \mu_c, \kappa} + 2 r(t) \langle \gamma, \xi_2 \rangle_{\mu, \mu_c, \kappa}} \, dt.
\end{align}

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

\begin{align}
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 \\
&\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{align}
The curve \( \gamma \) is piecewise differentiable by assumption, so choose \( a = a_1 < \ldots < a_n < a_{n+1} = 1 \) such that \( \gamma \) is continuously differentiable on \((a, 1) \setminus \{a_1, \ldots, a_n\}\).

We compute

\[
L(\Phi \circ \gamma) \geq \int_a^1 \frac{d}{dt} \|\gamma(t)\|_{\mu, \mu, \kappa} \, dt \geq \sum_{i=1}^{m} \int_{a_i}^{a_{i+1}} \frac{d}{dt} \|\gamma(t)\|_{\mu, \mu, \kappa} \, dt = \sum_{i=1}^{m} (\|\gamma(a_{i+1})\|_{\mu, \mu, \kappa} - \|\gamma(a_i)\|_{\mu, \mu, \kappa})
\]

\[
= \|\gamma(a_{m+1})\|_{\mu, \mu, \kappa} - \|\gamma(a_1)\|_{\mu, \mu, \kappa} = \|\gamma(1)\|_{\mu, \mu, \kappa} - 0 = \|M\|_{\mu, \mu, \kappa}.
\]

Finally, Lemma 1.7 yields

\[
E(\Phi \circ \gamma) \geq L(\Phi \circ \gamma)^2 \geq \|M\|^2_{\mu, \mu, \kappa}.
\]

We can now give a lower bound for the length of a curve \( Z \in A^B_A \mid [0, 1] \to GL^+(n) \) with \( B \notin \Phi_A(B_\varepsilon(0)) \): with \( t_0 = \min\{t \in [0, 1] \mid Z(t) \notin \Phi_A(B_\varepsilon(0))\} \), 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

\[
\lim_{t \to t_0} \|\gamma(t)\|_{\mu, \mu, \kappa} = \varepsilon.
\]

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

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

Furthermore, for \( \|M\|_{\mu, \mu, \kappa} < \varepsilon \), we can directly compute the distance between \( A \) and \( B = \Phi_A(M) \in B_\varepsilon(A) \):

For the geodesic curve \( X \in A^B_A \mid [0, 1] \to GL^+(n) \) with

\[
X(t) = \Phi_A(tM)
\]

we find \( X(0) = A, X(1) = \Phi_A(M) = B \) and, according to equation (3.10), \( L(X) = \|M\|_{\mu, \mu, \kappa} \).

Thus \( \text{dist}_{\text{geod}}(A, B) \leq \|M\|_{\mu, \mu, \kappa} \).

For any other curve \( Y \in A^B_A \mid [0, 1] \to GL^+(n) \), we find that \( Y \) is either contained in \( B_\varepsilon(A) \) in which case (4.10) implies \( L(Y) \geq \|M\|_{\mu, \mu, \kappa} \); or \( Y \) leaves \( B_\varepsilon(A) \). But then, as shown above, \( L(Y) \geq \varepsilon > \|M\|_{\mu, \mu, \kappa} \), and therefore

\[
\text{dist}_{\text{geod}}(A, B) = \|M\|_{\mu, \mu, \kappa}.
\]

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

\[
\text{dist}_{\text{geod}}(A, N) > 0.
\]

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

\[
g_{\Phi_A}(\gamma)(D\Phi_A[\gamma] \cdot \xi_2, D\Phi_A[\gamma] \cdot \xi_2) = 0
\]

on \([0, 1]\). Since \( g_{\Phi_A}(\gamma) \) is positive definite, this is equivalent to \( 0 = D\Phi_A[\gamma] \cdot \xi_2 = AD\Phi[\gamma] \cdot \xi_2 \), and since \( D\Phi \) is non-singular in a neighbourhood of 0, this equality holds (for small enough \( \varepsilon \)) if and only if

\[
0 = \xi_2(t) = \dot{\gamma}(t) - \xi_1(t) = \dot{\gamma}(t) - r(t) \gamma(t)
\]

for all \( t \in [0, 1] \). But then \( \gamma \) and \( \dot{\gamma} \) are linearly dependent, and thus \( \gamma \) 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 lemma:

**Lemma 4.4.** Let \( A \in GL^+(n) \). Then there exists \( \varepsilon(A) > 0 \) such that for all \( 0 < \varepsilon < \varepsilon(A) \):
where and (3.10) yields $L$

To simplify notation, denote by $L$

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

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

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

$$
\text{gl}(n) \ni S_\varepsilon(0) \to \mathbb{R}^+, \quad N \mapsto \text{dist}_{GL}(\Phi_A(N), B)
$$

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 $X_A \in C^\infty([0, r_A]; GL^+(n))$ by

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

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 $L_t(X_A) = t \varepsilon^{-1} \|M_A\|_{\mu, \mu^*, \kappa} = t$. Thus we find $\text{dist}_{GL}(X_A(t), A) \leq t$, and therefore

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

We will denote by

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

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

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

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].

**Corollary 4.5.** The geodesic distance $\text{dist}_{GL}$ is a metric on $GL^+(n)$.
Since $\text{dist}_{\text{geod}}(A, B) \geq \varepsilon$ by definition of $\varepsilon$, the curve $Y$ intersects $S_{\varepsilon}(A)$, so there exist $s \in (a, b)$ and, according to Lemma 4.4, $N \in S_{\varepsilon}(0) \subset \mathfrak{gl}(n)$ with $Y(s) = \Phi_A(N)$. Since $L(Y) = L(Y|_{[a,s]}) + L(Y|_{[s,b]})$, we obtain

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

(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 $\hat{t} := \max I_A \neq r_A$ and, for $\hat{A} := X_A(\hat{t})$, choose $\tilde{\varepsilon} < \varepsilon(\hat{A})$, $M_{\hat{A}}$, $X_{\hat{A}}$, $I_{\hat{A}}$ accordingly. We find $\varepsilon \in I_{\hat{A}}$, and thus $\text{dist}_{\text{geod}}(X_{\hat{A}}(\varepsilon), B) = \text{dist}_{\text{geod}}(\hat{A}, B) - \tilde{\varepsilon}$. Let $\hat{X}$ denote the piecewise smooth curve

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

(4.30)

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

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

(4.31)

and

$$\text{dist}_{\text{geod}}(A, \hat{X}(\hat{t} + \tilde{\varepsilon})) = \text{dist}_{\text{geod}}(A, X_{\hat{A}}(\tilde{\varepsilon})) \geq \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(X_{\hat{A}}(\tilde{\varepsilon}), B)$$

$$= \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(\hat{A}, B) + \tilde{\varepsilon}$$

$$= \text{dist}_{\text{geod}}(A, B) - \text{dist}_{\text{geod}}(X_A(\hat{t}), B) + \tilde{\varepsilon} \geq \hat{t} + \tilde{\varepsilon},$$

we infer that $\hat{X}$ is a length minimizer between $A$ and $\hat{X}(\hat{t} + \tilde{\varepsilon})$. Since $X_A$, $X_{\hat{A}}$ have (the same) constant speed, $\hat{X}$ is a piecewise smooth energy minimizer in $A_{\hat{A}}^{\hat{X}(\hat{t} + \tilde{\varepsilon})}$. Then according to Proposition 2.4, $\hat{X}$ is smooth and
satisfies (2.26) everywhere. Finally, because $X_A$ satisfies the differential equation with the same initial values as $\hat{X}$, the uniqueness from Proposition 2.8 yields

$$\hat{X} = X_A$$

on $[0, \tilde{t} + \tilde{\varepsilon}]$, \hspace{1cm} (4.32)

and therefore

$$\text{dist}_{\text{geod}}(X_A(\tilde{t} + \tilde{\varepsilon}), B) = \text{dist}_{\text{geod}}(X_A(\tilde{\varepsilon}), B) = \text{dist}_{\text{geod}}(\hat{A}, B) - \tilde{\varepsilon}$$

$$= \text{dist}_{\text{geod}}(X_A(\tilde{t}), B) - \tilde{\varepsilon}$$

$$= \text{dist}_{\text{geod}}(A, B) - (\tilde{t} + \tilde{\varepsilon}) \Rightarrow \tilde{t} + \tilde{\varepsilon} \in I_A,$$ \hspace{1cm} (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 Corollary 4.5, we find $X(r_A) = B$. Since $L(X_A) = r_A$, this concludes the proof. \hspace{1cm} \blacksquare

4.3 Conclusion

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

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

$$\langle M, N \rangle_{\mu, \mu, \kappa} = \mu(\text{dev sym } M, \text{dev sym } N) + \mu_v(\text{skew } M, \text{skew } N) + \frac{\kappa}{n} (\text{tr } M)(\text{tr } N),$$

$$\langle M, N \rangle = \text{tr}(M^T N)$$

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 $X : [0, 1] \to \text{GL}(n)$ (Lemmas 1.5 and 1.8) and the representation formula of energy minimizers from Theorem 3.1, we obtain our main result.

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

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

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

$$\|M\|_{\mu, \mu, \kappa} = L(X) = \text{dist}_{\text{geod}}(A, B) = \inf_{Y \in A^B} L(Y).$$ \hspace{1cm} (4.35)

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

$$\text{dist}_{\text{geod}}(A, B) = \min\{\|M\|_{\mu, \mu, \kappa} | M \in \mathfrak{gl}(n), A \exp(\text{sym } M - \omega \text{ skew } M) \exp((1 + \omega) \text{ skew } M) = B\}.$$ \hspace{1cm} (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

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

let alone one that is minimal with regard to the norm $\|.\|_{\mu, \mu, \kappa}$. We will therefore consider some easier special cases.
5.1 The geodesic distance on $GL^+(1)$

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

$$\langle x, y \rangle_\kappa = \kappa x \cdot y, \quad \kappa > 0,$$

with the corresponding left-invariant Riemannian metric

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

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

$$L(X) = \int_0^1 \sqrt{g_{X(t)}(\dot{X}(t), \dot{X}(t))} \, dt = \int_0^1 \sqrt{\dot{X}(t) \cdot \dot{X}(t)} \, dt = \kappa \int_0^1 \frac{\dot{X}(t)}{X(t)} \, dt,$$

$$E(X) = \int_0^1 g_{X(t)}(\dot{X}(t), \dot{X}(t)) \, dt = \kappa^2 \int_0^1 \left( \frac{\dot{X}(t)}{X(t)} \right)^2 \, dt.$$  

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 $X \in A_p^q$ 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 $X \in A_p^q$, and therefore any $X$ with

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

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

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

and check

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

as well as

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

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

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

and

$$\min_{Y \in A_p^q} L(Y) = L(X) = \kappa \int_0^1 \frac{\dot{X}(t)}{X(t)} \, dt = \kappa \int_0^1 |\ln(\frac{q}{p})| \, dt = \kappa |\ln(\frac{q}{p})| = \kappa |\ln(q) - \ln(p)|.$$  

We conclude:

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

$$\text{dist}_{\text{geo}}(p, q) = \kappa |\ln(\frac{q}{p})| = \kappa |\ln(q) - \ln(p)|,$$

and a shortest geodesic connecting $p$ and $q$ is given by

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

\[
\begin{align*}
\text{(sym} \ M)(\text{skew} \ M) &= (\text{skew} \ M)(\text{sym} \ M) \\
&\iff \frac{1}{2}(M + M^T) \cdot \frac{1}{2}(M - M^T) = \frac{1}{2}(M - M^T) \cdot \frac{1}{2}(M + M^T) \\
&\iff \frac{1}{4}(M^2 + M^T M - MM^T - (M^T)^2) = \frac{1}{4}(M^2 - M^T M + MM^T - (M^T)^2) \\
&\iff M^T M - MM^T = -M^T M + MM^T \\
&\iff M^T M = MM^T. \quad (5.12)
\end{align*}
\]

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). A proof of iii) can be found in [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 \( X : [0, 1] \to \text{GL}(n) \) 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

\[
\begin{align*}
\dot{U} &= X^{-1} \dot{X}, \\
\ddot{U} &= \frac{1 + \frac{t\omega}{\nu}}{2} (U^T U - U U^T) \quad (5.14)
\end{align*}
\]

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 thus \( \exp(\lambda_1 \text{sym} \ N + \lambda_2 \text{skew} \ N) = \exp(\lambda_1 \text{sym} \ N) \exp(\lambda_2 \text{skew} \ N) \) according to (A.6). This allows us to simplify \( X \): we find

\[
X(t) = \exp(t \text{sym} \ N - t \omega \text{skew} \ N) \exp(t(1 + \omega) \text{skew} \ N) \\
= \exp(t \text{sym} \ N) \exp(-t \omega \text{skew} \ N) \exp(t(1 + \omega) \text{skew} \ N) \\
= \exp(t \text{sym} \ N) \exp(-t \omega \text{skew} \ N + t(1 + \omega) \text{skew} \ N) \\
= \exp(t \text{sym} \ N) \exp(t \text{skew} \ N) = \exp(t \text{sym} \ N + t \text{skew} \ N) = \exp(tN). \quad (5.16)
\]
Thus for normal matrices $A$

positive definite symmetric (and therefore normal) matrix $F$

is indeed the length of $A$ for the distance of a normal matrix and $1$

and hence

$\log(N) = \log(M) + \log(1)$.

Let locally

i.e. there exists a normal matrix $N$.

But, again due to Lemma 5.2, a normal matrix $A$

and

1

of a normal matrix.

X length of $A$

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

$$L(X) = \int_0^1 \sqrt{\langle X^{-1} X', X^{-1} X' \rangle_{\mu,\mu,\kappa}} \, dt$$

$$= \int_0^1 \sqrt{\langle (\exp(tN)^{-1} \exp(tN) N) \rangle_{\mu,\mu,\kappa}} \, dt$$

$$= \int_0^1 \sqrt{\langle (N, N) \rangle_{\mu,\mu,\kappa}} \, dt = \|N\|_{\mu,\mu,\kappa} = \sqrt{\mu \|\text{dev sym} N\|^2 + \mu \|\text{skew} N\|^2 + \frac{\kappa}{n} \|\text{tr}(N)\|^2}, \quad (5.17)$$

and $X(0) = 1$, $X(1) = \exp(N)$. 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 $1$ with a normal matrix

Now let $A \in \text{GL}^+(n)$ be normal. To find a geodesic $X$ connecting $1$ 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 due to Lemma 5.2, a normal matrix $A$ has a (generally not uniquely determined) normal logarithm, i.e. there exists a normal matrix $\log(A) \in \mathfrak{gl}(n)$ such that $\exp(\log(A)) = A$. According to (5.16), with $N = \log(A)$, the geodesic $X$ with initial tangent $\log(A)$ has the form

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

and hence

$$X(1) = \exp(\log(A)) = A.$$

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

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

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

Note carefully that this does not immediately imply $\text{dist}_{\text{geod}}(1, A) = \|\log(A)\|_{\mu,\mu,\kappa}$: While $\text{dist}_{\text{geod}}(1, A)$ is indeed the length of a geodesic curve connecting $1$ 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 $\mu,\mu,\kappa > 0$. Then there exists $\varepsilon > 0$ such that for every normal $A \in \text{GL}^+(n)$ with $\|A - 1\| < \varepsilon$, the geodesic distance between $1$ and $A$ is given by

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

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

$$\text{dist}^2(F, F^{-T}) = \text{dist}^2(F^T F, 1) = \min\{\|M\|^2_{\mu,\mu,\kappa} \mid \exp(M) = F^T F\}$$

$$= \mu \|\text{dev} \log(F^T F)\|^2 + \frac{\kappa}{n} \|\text{tr} \log(F^T F)\|^2, \quad (5.20)$$

where $\log(F^T F)$ is the symmetric principal matrix logarithm of $F^T F$ on $\text{Sym}^+(n)$ [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 $1$ and $F^T F$. 

28
5.3 Application to nonlinear elasticity

An example for the application of geodesic distance measures is the theory of nonlinear elasticity (and, more specifically, hyperelasticity), where the deformation of a solid body is modelled via an energy functional \( W \) which depends on the gradient \( F \in \text{GL}^+(n) \) of the deformation (see e.g. [5, Chapter 4] for an introduction to hyperelasticity). Similarly, Mielke’s work on the geodesic distance in \( \text{SL}(n) \) was primarily motivated by elasto-plastic applications [18].

Among the energy functions considered in nonlinear elasticity is the isotropic Hencky strain energy \( W : \text{GL}^+(n) \to \mathbb{R}, \ W(F) = \mu \| \text{dev} \log \sqrt{F^TF} \|_2^2 + \frac{\kappa}{2} [\text{tr}(\log \sqrt{F^TF})]^2 \), (5.21)

which was introduced in 1929 by Heinrich Hencky [9]. Note that, because \( \log \sqrt{F^TF} \) is symmetric, \( W \) can be written as

\[
W(F) = \| \log \sqrt{F^TF} \|_{\mu,\mu_c,K}
\]

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 \( O(n) \)-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 \( \langle \cdot, \cdot \rangle_{\mu,\mu_c,K} \) on \( \mathfrak{gl}(n) \) with \( \mu_c \geq 0 \), and let \( F \in \text{GL}^+(n) \). Then

\[
\text{dist}_{\text{geod}}(F, \text{SO}(n)) = \text{dist}_{\text{geod}}(F, R) = \| \log \sqrt{F^TF} \|_{\mu,\mu_c,K} = \| \log U \|_{\mu,\mu_c,K},
\]

(5.23)

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

\[
\text{dist}_{\text{geod}}(F, \text{SO}(n)) = \inf_{Q \in \text{SO}(n)} \text{dist}_{\text{geod}}(F, Q)
\]

(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 \( \| F - \mathbf{1} \| \).

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 \( A, B \in \text{SO}(n) \) regarded as elements of \( \text{GL}(n) \) [17]. Note that although the geodesic distance on \( \text{SO}(n) \) with respect to the canonical bi-invariant metric is already well known [19], 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 matrix spaces as well as to Prof. Alexander Mielke (Weierstraß-Institut, Berlin) for pertinent discussions on geodesics in \( \text{GL}(n) \).
6 References

[1] E. Andruchow, G. Larotonda, L. Recht, and A. Varela. “The left invariant metric in the general linear group”. Journal of Geometry and Physics 86.0 (2014), pp. 241–257.

[2] D. S. Bernstein. Matrix Mathematics: Theory, Facts, and Formulas (Second Edition). Princeton reference. Princeton University Press, 2009.

[3] A. M. Bloch, P. E. Crouch, N. Nordkvist, and A. K. Sanyal. “Embedded geodesic problems and optimal control for matrix Lie groups”. Journal of Geometric Mechanics 3.2 (2011), pp. 197–223.

[4] M. P. do Carmo. Riemannian Geometry. Birkhäuser Basel, 1992.

[5] P. G. Ciarlet. Three-Dimensional Elasticity. Studies in Mathematics and its Applications 1. Elsevier Science, 1988.

[6] C. De Boor. “A naive proof of the representation theorem for isotropic, linear asymmetric stress-strain relations”. English. Journal of Elasticity 15.2 (1985), pp. 225–227. ISSN: 0374-3535. DOI: 10.1007/BF00041995

[7] J.-H. Eschenburg and J. Jost. Differentialgeometrie und Minimalflächen. Springer, 2007.

[8] S. Gallot, D. Hulin, and J. Lafontaine. Riemannian geometry. 3rd ed. Springer, 1990.

[9] H. Hencky. “Welche Umstände bedingen die Verfestigung bei der bildsamen Verformung von festen isotropen Körperrn?” Zeitschrift für Physik 55 (1929). available at www.uni-due.de/imperia/md/content/mathematik/ag_neff/hencky1929.pdf, pp. 145–155.

[10] N. J. Higham. Functions of Matrices: Theory and Computation. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2008.

[11] R. A. Horn and C. R. 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. Nakatsuoka. “The minimization of matrix logarithms: On a fundamental property of the unitary polar factor”. Linear Algebra and its Applications 449.0 (2014), pp. 28–42.

[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 (2007), pp. 2215–2224.

[16] J. E. Marsden and T. S. Ratiu. Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems. Vol. 17. Springer, 1999.

[17] R. J. Martin and P. Neff. “The GL(n)-geodesic distance on SO(n)”. in preparation (2016).

[18] A. Mielke. “Finite Elastoplasticity, Lie Groups and Geodesics on SL(d)”. Geometry, Mechanics, and Dynamics - Volume in Honor of the 60th Birthday of J.E. Marsden. Ed. by P. Newton, P. Holmes, and A. Weinstein. Springer New York, 2002, pp. 61–90.

[19] M. Moakher. “Means and Averaging in the Group of Rotations”. SIAM Journal on Matrix Analysis and Applications 24.1 (Jan. 2002), pp. 1–16. ISSN: 0895-4798. DOI: 10.1137/S0895479801383877

[20] P. Neff. Convexity and Coercivity in nonlinear, anisotropic elasticity and some useful relations. Tech. rep. available at https://www.uni-due.de/-hm0014/Download_files/cism_convexity08.pdf. Technische Universität Darmstadt, 2008.

[21] P. Neff, B. Eidel, and R. J. Martin. “Geometry of logarithmic strain measures in solid mechanics”. to appear in Archive for Rational Mechanics and Analysis (2016).

[22] P. Neff, B. Eidel, F. Osterbrink, and R. Martin. “A Riemannian approach to strain measures in nonlinear elasticity”. Comptes Rendus Mécanique 342.4 (2014), pp. 254–257.

[23] P. Neff, Y. Nakatsuoka, and A. Fischle. “A logarithmic minimization property of the unitary polar factor in the spectral norm and the Frobenius matrix norm”. SIAM Journal on Matrix Analysis and Applications 35.3 (2014). available at arXiv:1302.3235, pp. 1132–1154.

[24] C. H. Taubes. Differential geometry: bundles, connections, metrics and curvature. Oxford Graduate Texts in Mathematics. Oxford: Oxford University Press, 2011.

[25] B. Vandereycken, P.-A. 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 (2013), pp. 481–514.
A Appendix

A.1 Basic properties of the matrix exponential

A proof for the following elementary properties of the matrix exponential can be found in [10].

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

\[
\begin{align*}
&i) \quad \det(\exp(M)) = e^{\lambda M}, \\
&ii) \quad \exp(\frac{\lambda}{n} \mathbb{1}) = e^{\lambda} \mathbb{1}, \\
&iii) \quad \exp(M^T) = \exp(M)^T, \\
&iv) \quad \exp(-M) = \exp(M)^{-1}, \\
v) \quad MN = NM \Rightarrow \exp(M + N) = \exp(M) \exp(N) = \exp(N) \exp(M), \\
vi) \quad MN = NM \Rightarrow M \exp(N) = \exp(N)M, \\
vii) \quad T \in \text{GL}(n) \Rightarrow \exp(T^{-1}MT) = T^{-1} \exp(M)T, \\
viii) \quad M \in \mathfrak{so}(n) \Rightarrow \exp(M) \in \text{SO}(n), \\
ix) \quad M \in \text{Sym}(n) \Rightarrow \exp(M) \in \text{Sym}^+(n), \\
x) \quad \exp \left( \begin{array}{cc}
B_1 & 0 \\
\vdots & \ddots \\
0 & B_k
\end{array} \right) = \left( \begin{array}{cc}
\exp B_1 & 0 \\
\vdots & \ddots \\
0 & \exp B_k
\end{array} \right), \\
xii) \quad D \exp[M]T = \int_0^1 \exp(sM)T \exp((1 - s)M) \\ds.
\end{align*}
\]

If \( M \) and \( T \) commute, formula xi) simplifies to

\[
D \exp[M]T = \exp(M)T = T \exp(M).
\] (A.12)

Proposition A.2. The function

\[
\exp: \text{Sym}(n) \to \text{Sym}^+(n)
\]

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

\[
\exp^{-1}|_{\text{Sym}^+(n)}: \text{Sym}^+(n) \to \text{Sym}(n),
\]

is called the principal logarithm on \( \text{Sym}^+(n) \) and is denoted by \( \log \).

Proof. We refer to [2] Proposition 11.4.5] for a proof that \( \exp: \text{Sym}(n) \to \text{Sym}^+(n) \) 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 \( \text{Sym}^+(n) \) 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 \( \text{Sym}^+(n) \), it is differentiable. For further information on the matrix exponential, the matrix logarithm and matrix functions in general we refer to [10].

A.2 Additional proofs

Lemma A.3. 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 \( a = a_0 < \cdots < a_{m+1} = b \) such that \( X \) is continuously differentiable on \([a, b] \setminus \{a_0, \ldots, a_{m+1}\} \). Then the function

\[
l: [a, b] \to [0, L(X)], \quad l(t) := \int_a^t \|\dot{X}(t)\|_{X(t)} \, dt = L(X)_{[a, t]} \quad (A.13)
\]

is continuously differentiable on \([a, b] \setminus \{a_0, \ldots, a_{m+1}\} \) as well with \( l'(t) = \|\dot{X}(t)\|_{X(t)} \). By definition of regular curves (Definition 14.1), \( \dot{X}(t) \neq 0 \) for all \( t \in [a, b] \setminus \{a_0, \ldots, a_{m+1}\} \) 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

\[
\hat{\varphi}: [0, L(X)] \to [a, b], \quad \hat{\varphi}(s) = l^{-1}(s) \quad (A.14)
\]

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

\[
\dot{\tilde{X}}(t) = \frac{d}{dt}(X \circ \hat{\varphi}(t)) = \varphi'(t)X(\hat{\varphi}(t)) = \frac{1}{l'(\hat{\varphi}(t))} \dot{X}(\hat{\varphi}(t)) = \frac{\dot{X}(\hat{\varphi}(t))}{\|X(\hat{\varphi}(t))\|_{X(\hat{\varphi}(t))}} \quad (A.15)
\]
Lemma A.5. Let \( \dot{X} \) be differentiable at \( t \) and thus \( \|\dot{X}(t)\|_{\dot{X}(t)} \equiv 1 \), hence the transformation

\[
\varphi : [a, b] \to [a, b], \quad \varphi(t) := \tilde{\varphi}\left( (t - a) \frac{L(y)}{b - a} \right)
\]

has the desired properties.

To see the uniqueness of \( \varphi \), note that if both \( X \) and \( X \circ \tilde{\varphi} \) have constant speed, then

\[
\frac{d}{dt}(X \circ \tilde{\varphi})(t) \bigg|_{X \circ \tilde{\varphi}(t)} \equiv c \implies \|\dot{\varphi}'(t) \cdot \dot{X}(\tilde{\varphi}(t))\|_{X(\tilde{\varphi}(t))} \equiv c
\]

\[
\|\dot{\varphi}'(t)\| \equiv \frac{c}{\| X(\tilde{\varphi}(t))\|_{X(\tilde{\varphi}(t))}} \equiv \text{constant}
\]

for \( t \in [a, b] \setminus \{a_0, \ldots, a_{m+1}\} \). Then the restrictions \( \tilde{\varphi}(a) = a, \tilde{\varphi}(b) = b \) and \( \dot{\varphi} > 0 \) only allow for \( \tilde{\varphi} \equiv 1 \). \( \blacksquare \)

Lemma A.4. 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 \( \varphi' > 0 \). Then

\[
L(X \circ \varphi) = L(X).
\]

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

Then

\[
\begin{align*}
L(X \circ \varphi) &= \int_c^d \frac{d}{dt}(X \circ \varphi)(t) \|_{X \circ \varphi(t)} dt = \int_c^d \sum_{i=0}^{m} \int_{a_i}^{a_{i+1}} \|\dot{\varphi}'(t) \cdot \dot{X}(\varphi(t))\|_{X(\varphi(t))} dt dt \\
&= \sum_{i=0}^{m} \int_{a_i}^{a_{i+1}} \|\dot{\varphi}'(t)\|_{X(\varphi(t))} dt = \sum_{i=0}^{m} \int_{a_i}^{a_{i+1}} \|\dot{X}(t)\|_{X(t)} dt = \int_a^b \|\dot{X}(t)\|_{X(t)} dt = L(X).
\end{align*}
\]

\( \blacksquare \)

Lemma A.5. 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) \( \langle Q^T MQ, Q^T NQ \rangle = \langle M, N \rangle \) for all \( Q \in \mathfrak{O}(n) \), \hspace{1cm} (A.19)

ii) \( \langle M, \mathbb{I} \rangle = \text{tr}(M) \), \hspace{1cm} (A.20)

iii) \( \langle S, W \rangle = 0 \) for all \( S \in \text{Sym}(n) \), \( W \in \mathfrak{s}(n) \), \hspace{1cm} (A.21)

Proof.

i) \( \langle Q^T MQ, Q^T NQ \rangle = \text{tr}((Q^T MQ)^T Q^T NQ) = \text{tr}(Q^T M^T N Q Q^T) = \text{tr}(Q^T M^T N Q) = \text{tr}(M^T N Q) = \langle M, N \rangle \). \hspace{1cm} (A.22)

ii) \( \langle M, \mathbb{I} \rangle = \text{tr}(M^T \mathbb{I}) = \text{tr}(M^T) = \text{tr}(M) \). \hspace{1cm} (A.23)

iii) \( \langle S, W \rangle = \text{tr}(S^T W) = \text{tr}(SW) = \text{tr}(WS) = -\text{tr}(W^T S) = -\langle W, S \rangle = -\langle S, W \rangle \implies \langle S, W \rangle = 0 \). \hspace{1cm} (A.24)

\( \blacksquare \)

Lemma A.6. Let \( M, N \in \mathfrak{gl}(n) \) and \( \mu, \mu_c, \kappa \geq 0 \). Then

i) \( \langle Q^T MQ, Q^T NQ \rangle_{\mu,\mu_c,\kappa} = \langle M, N \rangle_{\mu,\mu_c,\kappa} \) for all \( Q \in \mathfrak{O}(n) \), \hspace{1cm} (A.25)

ii) \( \langle M, N \rangle_{\mu,\mu_c,\kappa} = \mu(\text{dev sym} M, N) + \mu_c(\text{skew} M, N) + \kappa \text{tr}(M) \text{tr}(N) \)

\[ = \langle \mu \text{dev sym} M + \mu_c \text{skew} M + \kappa \text{tr}(M) \cdot \mathbb{I}, N \rangle, \]

iii) \( \langle M, N \rangle_{1,1,1} = \langle M, N \rangle \), \hspace{1cm} (A.27)

iv) \( \langle S, W \rangle_{\mu,\mu_c,\kappa} = 0 \) for all \( S \in \text{Sym}(n) \) \( W \in \mathfrak{s}(n) \), \hspace{1cm} (A.28)

where \( \text{Sym}(n) \) and \( \mathfrak{s}(n) \) 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.5 yields \( \langle \text{dev sym} M, \text{skew} M \rangle = 0 \) (note that \( \text{dev sym} M \) is symmetric) as well as

\[
\langle \text{skew} M, \mathbb{I} \rangle = \text{tr}(\text{skew} M^T \mathbb{I}) = -\text{tr}(\text{skew} M) = -\langle \text{skew} M, \mathbb{I} \rangle = -\langle \text{skew} M, \mathbb{I} \rangle \implies \langle \text{skew} M, \mathbb{I} \rangle = 0,
\]

\[
\langle \text{dev sym} M, \mathbb{I} \rangle = \text{tr}(\text{dev sym} M) = \text{tr}(M - \frac{\text{tr} M}{n} \mathbb{I}) = \text{tr} M - \frac{\text{tr} M}{n} \text{tr} \mathbb{I} = 0;
\]
Thus
\[
\langle \mu \, \text{dev sym } M + \mu_c \, \text{skew } M + \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ N \rangle \\
= \langle \mu \, \text{dev sym } M + \mu_c \, \text{skew } M + \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ \text{dev sym } N + \text{ skew } N + \frac{\text{ tr} N}{n} \cdot I \rangle \\
= \langle \mu \, \text{dev sym } M , \ \text{dev sym } N \rangle + \langle \mu \, \text{dev sym } M , \ \text{skew } N \rangle + \langle \mu \, \text{dev sym } M , \ \frac{\text{ tr} N}{n} \cdot I \rangle \\
+ \langle \mu_c \, \text{skew } M , \ \text{dev sym } N \rangle + \langle \mu_c \, \text{skew } M , \ \text{skew } N \rangle + \langle \mu_c \, \text{skew } M , \ \frac{\text{ tr} N}{n} \cdot I \rangle \\
+ \langle \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ \text{dev sym } N \rangle + \langle \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ \text{skew } N \rangle + \langle \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ \frac{\text{ tr} N}{n} \cdot I \rangle \\
= \langle \mu \, \text{dev sym } M , \ \text{dev sym } N \rangle + \langle \mu_c \, \text{skew } M , \ \text{skew } N \rangle + \langle \frac{\kappa}{n} \text{ tr}(M) \cdot I , \ \frac{\text{ tr} N}{n} \cdot I \rangle \\
= \mu \langle \text{dev sym } M , \ \text{dev sym } N \rangle + \mu_c \langle \text{skew } M , \ \text{skew } N \rangle + \frac{\kappa}{n} \langle \text{tr}(M) \cdot I , \ \frac{\text{ tr} N}{n} \rangle \\
\]
which proves ii). Then iii) follows directly by
\[
\langle M , N \rangle_{1,1,1} = \langle \text{dev sym } M + \text{ skew } M + \frac{\text{ tr} M}{n} \cdot I , \ N \rangle = \langle M , N \rangle ,
\]
and iv) from Lemma [A.5] by
\[
\langle S , W \rangle_{\mu,\mu_c,\kappa} = \langle W , S \rangle_{\mu,\mu_c,\kappa} = \langle \mu \, \text{dev sym } W + \mu_c \, \text{skew } W + \frac{\kappa}{n} \text{ tr}(W) \cdot I , \ S \rangle = \mu_c \langle W , S \rangle = 0 .
\]
Finally, since
\[
\text{ tr}(Q^T MQ) = \text{ tr}(MQQ^T) = \text{ tr}(M) ,
\]
\[
\text{ sym}(Q^T MQ) = \frac{1}{2} (Q^T MQ + (Q^T MQ)^T) = \frac{1}{2} (Q^T (M + M^T)Q) = Q^T (\text{sym } M)Q ,
\]
\[
\text{ skew}(Q^T MQ) = \frac{1}{2} (Q^T MQ - (Q^T MQ)^T) = \frac{1}{2} (Q^T (M - M^T)Q) = Q^T (\text{skew } M)Q ,
\]
\[
\text{ dev sym}(Q^T MQ) = \text{ dev}(Q^T (\text{sym } M)Q) = Q^T (\text{sym } M)Q - \frac{\text{ tr}(Q^T (\text{sym } M)Q)}{n} \cdot I \\
= Q^T (\text{sym } M)Q - \frac{\text{ tr}(\text{sym } M)}{n} \cdot I = Q^T (\text{sym } M - \frac{\text{ tr}(\text{sym } M)}{n} I)Q = Q^T (\text{dev sym } M)Q ,
\]
i) follows from
\[
\langle Q^T MQ , Q^T NQ \rangle_{\mu,\mu_c,\kappa} = \mu \langle \text{dev sym}(Q^T MQ) , \text{dev sym}(Q^T NQ) \rangle + \mu_c \langle \text{skew}(Q^T MQ) , \text{skew}(Q^T NQ) \rangle + \frac{\kappa}{n} \text{ tr}(Q^T MQ) \text{ tr}(Q^T NQ) \\
= \mu \langle Q^T (\text{dev sym } M)Q , Q^T (\text{dev sym } N)Q \rangle + \mu_c \langle Q^T (\text{skew } M)Q , Q^T (\text{skew } N)Q \rangle + \frac{\kappa}{n} \langle \text{tr}(M) \cdot I \rangle \text{ tr}(N) \\
= \mu \langle \text{dev sym } M , \text{dev sym } N \rangle + \mu_c \langle \text{skew } M , \text{skew } N \rangle + \frac{\kappa}{n} \langle \text{tr}(M) \cdot I \rangle \text{ tr}(N) = \langle M , N \rangle_{\mu,\mu_c,\kappa} .
\]
\]