ON STRAIN MEASURES AND THE GEODESIC DISTANCE TO $\text{SO}_n$ IN THE GENERAL LINEAR GROUP

RAZ KUPFERMAN AND ASAF SHACHAR

Institute of Mathematics
The Hebrew University
Jerusalem 91904, Israel

(Communicated by Reuven Segev)

ABSTRACT. We consider various notions of strains—quantitative measures for the deviation of a linear transformation from an isometry. The main approach, which is motivated by physical applications and follows the work of [12], is to select a Riemannian metric on $\text{GL}_n$, and use its induced geodesic distance to measure the distance of a linear transformation from the set of isometries. We give a short geometric derivation of the formula for the strain measure for the case where the metric is left-$\text{GL}_n$-invariant and right-$\text{O}_n$-invariant. We proceed to investigate alternative distance functions on $\text{GL}_n$, and the properties of their induced strain measures. We start by analyzing Euclidean distances, both intrinsic and extrinsic. Next, we prove that there are no bi-invariant distances on $\text{GL}_n$. Lastly, we investigate strain measures induced by inverse-invariant distances.

1. Introduction. In various physical and mathematical contexts, a natural question arises: how to quantify the distortion of an invertible linear transformation $A \in \text{GL}_n$? That is, how far is $A$ from being an isometry? In material science, the local distortion of a map between two manifolds is known as a strain measure.

One can investigate various notions of strain measures. A natural approach is to choose a distance function $d$ on $\text{GL}_n$, and define the strain measure as follows:

$$\text{Strain}(A) = \text{dist}(A, \text{SO}_n) = \inf_{Q \in \text{SO}_n} d(A, Q).$$

Since $\text{SO}_n$ is compact, the distance is realized for some $Q \in \text{SO}_n$.

In material science, the configuration of a body is a map $f$ from a body manifold $B$ to a space manifold $S$. If both manifolds are endowed with Riemannian metrics, then one can define a local strain measure at every point $p$ of the body manifold,

$$\text{Strain}(df) = \text{dist}(df, \text{SO}_n),$$

where $\text{SO}_n$ here refers to the space of pointwise orientation-preserving isometries. By choosing orthonormal frames at both $p$ and $f(p)$, invertible linear maps between tangent spaces can be identified with $\text{GL}_n$, whence the relevance of the proposed framework to general Riemannian settings.

The notion of strain measure depends on the choice of a distance function $d$. In physical applications, one expects this distance to satisfy certain symmetries with respect to left- and right-multiplication—the former is related to symmetries of...
the ambient space whereas the latter is related to material symmetries. The most common symmetry assumptions are frame-indifference, which is left-\(O_n\)-invariance, and material isotropy, which is right-\(O_n\)-invariance.

Left- and right-\(O_n\)-invariance do not determine a unique distance on \(GL_n\), nor do they determine a unique strain measure. The most common distance \(d\) is the so-called Frobenius, or Euclidean distance,

\[ d_{\text{Euc}}(A, B) = \| A - B \|_F, \]

where \(\| A \|_2^2 = \text{tr}(A^T A)\). The Euclidean distance results in a strain measure given by

\[ \text{Strain}_{\text{Euc}}(A) = \| \sqrt{A^T A} - I \|_F. \]

The Euclidean strain measure suffers from well-known drawbacks. From a physical point of view, the main drawback is that \(\text{Strain}_{\text{Euc}}(A)\) remains finite as \(A\) tends toward singularity. The Euclidean strain “penalizes” extreme expansions, but does not “penalize” extreme contractions.

The space \(GL_n\) is a smooth submanifold of the space \(M_n\) of \(n \times n\) matrices. Thus, a natural way to define a distance \(d\) on \(GL_n\) is via a Riemannian metric \(g\). In this context, the Euclidean distance \(d_{\text{Euc}}\) is induced by the Euclidean metric on \(M_n\),

\[ g_{\text{Euc}}(X, Y) = \text{tr}(X^T Y), \quad \text{where } X, Y \in T_Z M_n \simeq M_n \]

For \(A, B \in M_n\), \(d_{\text{Euc}}(A, B)\) is the length of the segment \([A, B]\) with respect to the metric \(g_{\text{Euc}}\).

A note about terminology: to avoid confusion, we will use the term “distance”, rather than “metric” in the context of a metric space. The term “metric” will be reserved for Riemannian metrics.

From a mathematical point of view, a drawback of \(d_{\text{Euc}}|_{GL_n}\) as a distance function on \(GL_n\) is that it is not an intrinsic distance. Since \(GL_n\) is not convex, segments \([A, B]\), \(A, B \in GL_n\) may not be contained in \(GL_n\).

The drawbacks of the Euclidean strain measure are at the heart of a series of papers by Neff and co-workers \([14, 7, 9, 12]\). They endow \(GL_n\) with a metric that possesses an additional symmetry: in addition to the bi-\(O_n\)-invariance, they assume left-\(GL_n\)-invariance; this is perhaps the most symmetric choice, as it is well-known that there are no bi-invariant metrics on \(GL_n\). This additional symmetry restricts drastically the set of possible metrics. The left-\(GL_n\) invariance implies that the metric is fully determined by its value at the identity. The addition of right-\(O_n\)-invariance yields a family of metrics depending only on three parameters.

In a prior related work by Mielke and co-workers \([11, 6, 10]\), the authors analyze the same family of left-invariant metrics on \(SL_n\). In particular they prove a formula for the geodesics on \(SL_n\), which is identical to \(1\).

Another work on the geodesics appears in \([1]\), where the authors derive a formula for the geodesics for a family of left-invariant metrics intersecting the family mentioned above, which is the main focus of this paper. (They took the standard \(p\)-norm on \(T_I GL_n\) and used its left-translation as the metric. This includes only the standard metric from our family, however it also contains some Finsler metrics that are not Riemannian).

It was shown in \([12]\) that the unique matrix in \(SO_n\) that is the closest to \(A \in GL_n^+\) is its orthogonal polar factor \(O\), where \(A = OP\), with \(O \in SO_n\) and \(P \in \text{Psym}_n\). Moreover, a closed formula for the strain measure was derived,

\[ \text{Strain}(A) = \text{dist}(A, SO_n) = \| \log \sqrt{A^T A} \|, \]
where the logarithm of a symmetric positive-definite matrix is its unique symmetric logarithm, and the norm \(|\cdot|\) depends on two out of the three parameters mentioned above (see (4) below). This strain measure diverges in singular limits. In particular, it is inverse-invariant, i.e.

\[
\text{Strain}(A) = \text{Strain}(A^{-1}).
\]

In this paper we provide an elementary derivation of formula (1) for the strain measure. Using geometric insights, the set of all possible minimizing paths from a given \(A \in \text{GL}_n^+\) to \(\text{SO}_n\) is narrowed considerably. This helps determining the minimal distance in an elementary way. In particular, our analysis clarifies the different roles played by the various symmetries of the metrics.

In section 2, we introduce the family of left-\(\text{GL}_n^\text{- right-}\text{O}_n\)-invariant metrics. We state a key property satisfied by these metrics—orthogonality relations—which play a central role in the forthcoming analysis. We also describe the form of the geodesics. Section 3 contains the derivation of the corresponding strain measure. In Subsection 8, we shed light on the reasons for assuming \(\text{O}_n\)-invariance, rather than \(\text{SO}_n\)-invariance, which might have seemed a more natural assumption.

In Section 4, we turn to analyze extrinsic versus intrinsic distances, first in a general Riemannian setting and then applied to the case of \(\text{GL}_n\) viewed as a submanifold of \(M_n\) endowed with the Euclidean metric. The main result is that while the intrinsic distance differs from the extrinsic distance, the strain measures are the same in both cases. In particular, we give a very short derivation of Grioli’s optimality theorem [5] (see also [13]), which says that for a given \(A \in \text{GL}_n^+\), its orthogonal polar factor is the closest matrix to \(A\) in \(\text{SO}_n\) with respect to the Frobenius norm.

In Section 5, we investigate how other natural symmetries on distance functions affect the strain measure. We start by showing there are no bi-invariant distance functions on \(\text{GL}_n\), hence there is an “upper limit” to the amount of symmetries a distance function can possess (see Subsection 5.1).

Next, we show that an inverse-invariant strain measure is obtained if the distance/metric is inverse-invariant. We then describe two different techniques for obtaining such distances/metrics via symmetrization, and analyze the resulting strain measures. In the case of symmetrizing a distance, we investigate the two families of distances considered thus far: the Euclidean (intrinsic and extrinsic) distance, and the (intrinsic) distances induced by the metrics considered in Section 2. In the case of the Euclidean distance, the result is an improved strain measure, which penalizes expansions and contractions equally. In the other cases, the strain measure is essentially the same as without the symmetrization.

Finally, we discuss the symmetrizations of all the metrics considered in Section 2. The resulting strain measure is also essentially the same as the original. The proof contains an analysis of metrics that are expressed as sums of two metrics, and also sheds light on the key ingredients in the derivation of the strain measure in Subsection 3.2.

2. Symmetries and geodesics.

2.1. Left-\(\text{GL}_n\)- and right-\(\text{O}_n\)-invariant metrics. Throughout this paper, we use the following notations:

\(\text{GL}_n\) is the group of \(n \times n\) invertible real matrices, \(\text{GL}_n^+\) and \(\text{GL}_n^-\) are the connected components of \(\text{GL}_n\), i.e., \(\text{GL}_n^+\) is the subgroup of \(n \times n\) invertible matrices with
positive determinant, and $\operatorname{GL}^-_n$ is the subset of matrices with negative determinant. We denote by 
$$
\mathcal{O}_n = \{ Q \in \operatorname{GL}_n \mid Q^T Q = I \} \subset \operatorname{GL}_n
$$
the subgroup of orthogonal matrices, whereas $\operatorname{SO}_n \subset \operatorname{GL}^+_n$ is the subgroup of special orthogonal matrices, i.e those with determinant 1.

We will denote by $M_n$ the vector space of $n \times n$ real matrices, and by $\operatorname{Psym}_n \subset M_n$ the cone of symmetric positive-definite matrices.

For readability, we will try to stick to the following choice of symbols:
$$
A, B \in \operatorname{GL}_n \\
O, U \in \mathcal{O}_n \\
Q \in \operatorname{SO}_n \\
X, Y \in M_n \\
P \in \operatorname{Psym}_n.
$$

Let $g$ be a left-$\operatorname{GL}_n$- and right-$\mathcal{O}_n$-invariant metric on $\operatorname{GL}_n$. A left-invariant metric $g$ on a Lie group $G$ is determined by its restriction at the identity. For $A \in \operatorname{GL}_n$, let $L_A : \operatorname{GL}_n \rightarrow \operatorname{GL}_n$ denote left multiplication by $A$, i.e $L_A(B) = AB$. $L_A$ is a diffeomorphism and its differential $(dL_A)_I : T_I \operatorname{GL}_n \rightarrow T_A \operatorname{GL}_n$ is a vector space isomorphism. For all $X, Y \in T_I \operatorname{GL}_n$,
$$
g_I(X,Y) = g_A((dL_A)_I X, (dL_A)_I Y). \quad (2)
$$

Since $\operatorname{GL}_n$ is an open subset of $M_n$, its tangent space at each point is canonically identified with $M_n$ as follows: Given $A \in \operatorname{GL}_n$, the identification $i_A : M_n \rightarrow T_A \operatorname{GL}_n$ is
$$
i_A(X) = [t \mapsto A + tX].
$$

Using the above identification,
$$
(dL_A)_B X = AX,
$$
where the dependence of the right-hand side on $B$ is implicit via the identification of $M_n$ with $T_{AB} \operatorname{GL}_n$.

Substituting this last identity for $B = I$ into (2) we obtain that left-$\operatorname{GL}_n$-invariance implies,
$$
g_I(X,Y) = g_A(AX,AY) \quad \forall A \in \operatorname{GL}_n.
$$

Similarly, right-$\mathcal{O}_n$-invariance implies
$$
g_A(X,Y) = g_{AO}(XO,YO) \quad \forall O \in \mathcal{O}_n.
$$

An immediate consequence of both left- and right-$\mathcal{O}_n$-invariance, is that $g_I$ is isotropic. For every $U \in \mathcal{O}_n$:
$$
g_I(X,Y) = g_U(U^T X, U^T Y) = g_{U^T} (U^T XU, U^T YU) = g_I(U^T XU, U^T YU). \quad (3)
$$

In fact, the same argument shows that for any Lie group $G$ and subgroup $H \subseteq G$, a left-invariant metric $g$ is right-$H$-invariant if and only if $g_e$ is invariant under conjugation with elements in $H$.

From a representation theorem for isotropic operators [2], it follows that there exist constants $\alpha, \beta \geq 0$ and $\gamma \leq 0$, such that
$$
g_I(X,Y) = \alpha \operatorname{tr}(X) \operatorname{tr}(Y) + \beta \operatorname{tr}(\operatorname{sym} X \operatorname{sym} Y) + \gamma \operatorname{tr} (\operatorname{skew} X \operatorname{skew} Y), \quad (4)
$$
where $\text{sym} X$ and $\text{skew} X$ denote respectively the symmetric and skew-symmetric parts of $X$.

Let $\text{sym} \subset M_n$ and $\mathfrak{D}_n \subset M_n$ denote the subspaces of symmetric and anti-symmetric matrices in $M_n \cong T_I \text{GL}_n$. The following lemma asserts that these sets are orthogonally complementary with respect to $g_I$:

**Lemma 2.1.** Let $g_I$ satisfy the isotropy condition (3). Then, $\text{sym}$ and $\mathfrak{D}_n$ are orthogonally complementary.

*Proof.* The orthogonality of $\text{sym}$ and $\mathfrak{D}_n$ can be shown by an explicit substitution in the form (4) of the metric, hence $\text{sym} \subseteq \mathfrak{D}_n^\perp$. The fact that these subspaces are complementary follows from a dimensional argument,

$$\dim \text{sym} + \dim \mathfrak{D}_n = n^2 = \dim T_I \text{GL}_n.$$

$\square$

### 2.2. Geodesics

In this section we review the properties of geodesic curves in $(\text{GL}_n, g)$.

**Proposition 1** ($g$-geodesics starting at the identity). Let $g$ be left-$\text{GL}_n$, right-$\text{O}_n$-invariant. Let $g_I$ be given by (4) and denote $\kappa = (\beta - \gamma)/2\beta$. Let $\gamma : I \to \text{GL}_n$ be the $g$-geodesic, satisfying the initial conditions

$$\gamma(0) = I \quad \text{and} \quad \dot{\gamma}(0) = X_0.$$  

Then,

$$\gamma(t) = \exp((1 - \kappa)tX_0 + \kappa tX_0^T) \exp(\kappa t(X_0 - X_0^T)).$$

*Proof.* This was proved in [9] using an argument based on variations of energy. A shorter alternative proof using Cartan’s moving frame method is given in Appendix A.

**Corollary 1** ($g$-geodesics). Under the same assumptions as above, let $\gamma : I \to \text{GL}_n$ be the $g$-geodesic satisfying the initial conditions

$$\gamma(0) = A \quad \text{and} \quad \dot{\gamma}(0) = AX_0.$$  

Then,

$$\gamma(t) = A \exp((1 - \kappa)tX_0 + \kappa tX_0^T) \exp(\kappa t(X_0 - X_0^T)).$$

*Proof.* This follows from the fact that left multiplication is an isometry of $(\text{GL}_n, g)$. It is a general property of Riemannian manifolds that isometries map geodesics into geodesics.

**Corollary 2.** Let $\gamma : I \to \text{GL}_n$ be the $g$-geodesic satisfying the initial conditions

$$\gamma(0) = Q \quad \text{and} \quad \dot{\gamma}(0) = QV,$$

where $Q \in \text{O}_n$ and $V \in \text{sym}$. Then,

$$\gamma(t) = Q \exp(tV).$$
3. Geodesic distance from $\text{SO}_n$. Every Lie group endowed with a left-invariant metric is complete as a Riemannian manifold. That is, every geodesic extends indefinitely. This follows from the fact that its isometry group acts transitively; see [3, p. 154, Example 12]. By the Hopf-Rinow theorem, [3, p. 146] the length-distance between any two points is realized by a minimizing geodesic.

Generally, there doesn’t seem to exist any explicit expression for the (possibly many) geodesics connecting any two elements $A, B \in \text{GL}_n^+$, nor for the resulting distance between these elements. Yet, we are only interested in the distance of an element $A \in \text{GL}_n^+$ from the subgroup $\text{SO}_n$ of isometries. As demonstrated in [12], an explicit expression can be derived for that distance. In this section we offer a simplified derivation of that expression.

3.1. Reduction to diagonal positive-definite matrices. The first step in calculating the distance of $A \in \text{GL}_n^+$ from $\text{SO}_n$ is to show that it is sufficient to obtain a formula for diagonal matrices. The following proposition holds for any bi-$\text{SO}_n$-invariant distance on $\text{GL}_n^+$—not necessarily a distance induced by a Riemannian metric.

**Proposition 2.** Let $d$ be a bi-$\text{SO}_n$-invariant distance on $\text{GL}_n^+$; we denote the corresponding distance between sets by $\text{dist}$. Let $A \in \text{GL}_n^+$. If $A = U\Sigma V^T$ is a singular value decomposition (SVD) of $A$ with $U, V \in \text{SO}_n$, then

$$\text{dist}(A, \text{SO}_n) = \text{dist}(\Sigma, \text{SO}_n).$$

Moreover, if $Q$ is a matrix closest to $\Sigma$ in $\text{SO}_n$, then $UQV^T$ is a matrix closest to $A$ in $\text{SO}_n$.

**Proof.** We first note that for every $A \in \text{GL}_n^+$, there exists an SVD such that $U, V \in \text{SO}_n$ (see the comment after the proof of Corollary 3). Moreover, $\Sigma$ is unique (up to permutation), i.e., the singular values do not depend on the particular decomposition.

Assuming $U, V \in \text{SO}_n$ and using the bi-$\text{SO}_n$-invariance,

$$\text{dist}(A, \text{SO}_n) = \min_{Q \in \text{SO}_n} d(A, Q) = \min_{Q \in \text{SO}_n} d(U\Sigma V^T, Q) = \min_{Q \in \text{SO}_n} d(\Sigma, U^T Q V) = \text{dist}(\Sigma, \text{SO}_n).$$

(5)

The last equality holds since $\{U^T Q V \mid Q \in \text{SO}_n\} = \text{SO}_n$. Equation (5) implies that $Q \in \text{SO}_n$ is a matrix closest to $\Sigma$ in $\text{SO}_n$ if and only if $UQV^T$ is a matrix closest to $A$ in $\text{SO}_n$.

3.2. Geodesic distance for diagonal matrices. By Proposition 2, we can focus our attention on finding the distance from $\text{SO}_n$ for diagonal positive-definite matrices, $\Sigma$. Since $\text{GL}_n^+$ is complete, we look for a minimizing geodesic from $\Sigma$ to $\text{SO}_n$. To do so, we are going to exploit the fact that any geodesic minimizing the distance of a point to a submanifold intersects that submanifold perpendicularly. More precisely:

**Lemma 3.1.** Let $M$ be a complete Riemannian manifold. Let $S \subseteq M$ be a submanifold, and let $p \in M \setminus S$. Assume $q \in S$ is a point on $S$ satisfying $d(p, q) = \text{dist}(p, S)$ (there is always such a point $q$ if $S$ is compact). Let $\alpha$ be a minimizing geodesic connecting $p$ and $q$. Then $\alpha$ is orthogonal to $S$ at $q$.

See C.1 for a proof.
Proposition 3. Let $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ be a diagonal matrix with positive entries. Then,

$$\text{dist}(\Sigma, SO_n) = d(\Sigma, I).$$

Moreover, $I$ is the unique element of $SO_n$ minimizing the distance from $\Sigma$.

Proof. Let $Q \in SO_n$ satisfy

$$d(\Sigma, Q) = \text{dist}(\Sigma, SO_n).$$

By the completeness of $GL_n^+$, there exists a minimizing geodesic $\alpha : [0, 1] \rightarrow GL_n^+$ from $Q$ to $\Sigma$, i.e.,

$$\alpha(0) = Q, \quad \alpha(1) = \Sigma \quad \text{and} \quad L(\alpha) = \|\dot{\alpha}(0)\|_Q = d(\Sigma, Q).$$

Lemma 3.1 implies that $\dot{\alpha}(0) \perp T_QSO_n$.

Denoting $\dot{\alpha}(0) = QV$ (where we think of $V$ as an element of $T_QGL_n^+ \cong M_n$, and $QV$ is identified with $d(L_Q)I(V)$), we obtain for any antisymmetric matrix $X \in O_n = T_I SO_n$:

$$g_I(X, V) = g_Q(QX, QV) = g_Q(QX, \alpha(0)) = 0,$$

where the last equality is valid since $d(L_Q)I(T_QSO_n) = T_QSO_n$, hence $QX = d(L_Q)I(X) \in T_QSO_n$.

Thus, $V \in T_IGL_n^+$ is orthogonal to every anti-symmetric matrix, and by Lemma 2.1, $V \in \text{sym}$. It follows from Corollary 2 that

$$\alpha(t) = Q e^{tV}.$$  

By the definition of $\alpha$, $\alpha(1) = \Sigma = Q e^{V}$. Since $V$ is symmetric, $e^{V}$ is symmetric positive-definite, hence we obtain two polar decompositions of $\Sigma$,

$$\Sigma = I \Sigma \quad \text{and} \quad \Sigma = Q e^{V}.$$  

By the uniqueness of polar decomposition for invertible matrices, we conclude that $Q = I$, which completes the proof.

We proceed to derive an explicit formula for the distance of $\Sigma$ from $SO_n$. Since $Q = I$,

$$e^{V} = \Sigma = e^{\log \Sigma},$$

where $\log \Sigma = \text{diag}(\log \sigma_i)$. Since $V$ and $\log \Sigma$ are symmetric, and since the matrix exponential is injective on the space of symmetric matrices, it follows that $V = \log \Sigma$. Hence,

$$\text{dist}(\Sigma, SO_n) = \|\dot{\alpha}(0)\|_I = \|V\|_I = \|\log \Sigma\|_I.$$  

Substituting the explicit form (4) of the metric $g_I$,

$$\text{dist}(\Sigma, SO_n) = \sqrt{\alpha \left( \sum \log \sigma_i \right)^2 + \beta \sum (\log \sigma_i)^2}. \quad (6)$$

As a corollary, we get that $\alpha(t) = e^{t \log \Sigma}$ is the unique minimizing geodesic connecting $I$ to $\Sigma$.  

3.3. Geodesic distance for arbitrary matrices. Let \( A \in GL_n^+ \) be an arbitrary matrix. If \( A = U \Sigma V^T \) is an SVD of \( A \), then \( \sqrt{A^T A} = V \Sigma V^T \), hence \( \log \sqrt{A^T A} = V \log \Sigma V^T \). By Proposition 2,

\[
\text{dist}(A, SO_n) = \| \log \Sigma \|_I = \| V \log \Sigma V^T \|_I = \| \log \sqrt{A^T A} \|_I, \tag{7}
\]

where the second equality follows from the invariance (3). Whenever we write \( \log \) of a symmetric positive-definite matrix, we refer to its unique symmetric logarithm. Since the exponential map is a diffeomorphism from \( \text{sym}_n \) to \( \text{Psym}_n \), there is no ambiguity here. We have thus obtained an explicit expression for the distance of any matrix \( A \in GL_n^+ \) from \( SO_n \) by elementary means.

We have shown that for a diagonal positive-definite matrix \( \Sigma, Q = I \) is the unique element in \( SO_n \) satisfying dist(\( \Sigma, SO_n \)) = \( d(\Sigma, Q) \). By Proposition 2, if \( A = U \Sigma V^T \) is an SVD of \( A \in GL_n^+ \) with \( U, V \in SO_n \), then \( UV^T \) is the unique matrix in \( SO_n \) that is closest to \( A \).

Moreover:

Corollary 3 (The orthogonal polar factor is the minimizer). Let \( A \in GL_n^+ \). Let \( A = OP \) be the polar decomposition of \( A, O \in SO_n \) and \( P \in \text{Psym}_n \). Then \( O \) is the matrix closest to \( A \) in \( SO_n \).

Proof. By orthogonally diagonalizing \( P \) with \( P = \tilde{U} \Sigma \tilde{U}^T \), we obtain an SVD,

\[
A = O \tilde{U} \Sigma \tilde{U}^T = U \Sigma V^T,
\]

where \( U = O \tilde{U} \) and \( V = \tilde{U} \). Note that by interchanging two columns if necessary, we can assume \( \tilde{U} \in SO_n \), hence \( V, U \in SO_n \). By the above discussion, \( UV^T = O \tilde{U} \tilde{U}^T = O \) is the matrix closest to \( A \).

Please note: the above argument shows that for \( A \in GL_n^+ \) there always exists an SVD where both orthogonal matrices are in \( SO_n \).

3.4. \( O_n \) versus \( SO_n \)-invariance. The analysis presented in Sections 2 and 3 assumes that the metric \( g \) is left-\( GL_n^- \)- and right-\( O_n \)-invariant. Since we are interested in intrinsic distances in \( GL_n^+ \) from the subgroup \( SO_n \), it may seem as if we could perform the whole analysis in \( GL_n^+ \) rather than in \( GL_n \). In such case, it only makes sense to require the Riemannian metric to be left-\( GL_n^+ \) and right-\( SO_n \)-invariant—that \( O_n \)-invariance, for example, is meaningless. A natural question is the following: would we obtain the same geodesic distances and the same strain measures if we considered left-\( GL_n^+ \) and right-\( SO_n \)-invariant metric on \( GL_n^+ \)?

An inner-product \( g_I \) satisfying condition (3) is called isotropic. In contrast, an inner-product \( g_I \) satisfying

\[
g_I(X,Y) = g_I(S^TXS, S^TYS), \quad \forall S \in SO_n \tag{8}
\]

is called hemitropic. If every hemitropic inner-product is isotropic, then our entire analysis extends as is to \( SO_n \)-invariant metrics on \( GL_n^+ \). If, however, isotropy and hemitropy are not equivalent, then our analysis has to be revisited, as the representation of the inner-product (4) relies explicitly on the isotropic nature of the inner-product \( g_I \).

Our analysis relies on the specific form (4) of the inner-product \( g_I \) in two crucial aspects: (i) in the derivation of an explicit formula for the geodesics, and (ii) in obtaining the orthogonality of symmetric and anti-symmetric matrices. Since an inner-product is of the form (4) if and only if it is isotropic, any hemitropic, but non-isotropic inner-product is not of that form, hence our analysis is not applicable.
It turns out that for all dimensions \( n \neq 4 \), there are no hemitropic non-isotropic inner-products. For odd \( n \) this is trivial to see since \(-I \in O_n \setminus SO_n\) commutes with every other matrix. The analysis for even dimensions is less trivial. A proof can be found in [4]. Thus, our work holds as is with isotropy replaced by hemitropy in any dimension other than 4. Understanding the implications of an hemitropy assumption for \( n = 4 \) remains an open question.

4. Intrinsic versus extrinsic distances. Endowing \( GL_n^+ \) with distances induced by Riemannian metrics is one type of choice for measuring the distortion of a linear map. Another popular choice is the distance induced by the Frobenius inner-product on \( M_n \), or equivalently, the Euclidean metric on \( M_n \) identified with \( \mathbb{R}^{n^2} \). In fact, one of the motivations in [12] for considering distances induced by Riemannian metrics was the claimed inadequacy of the Euclidean metric. Note that the Euclidean metric gives rise to two distinct distances on \( GL_n^+ \): (i) an extrinsic distance, obtained by restricting the Euclidean distance function to the subset \( GL_n^+ \) of \( M_n \), and (ii) an intrinsic length-distance determined by paths in \( GL_n^+ \).

In this section we explore the distinction between extrinsic and intrinsic distances, first in a general Riemannian context, and second, in the original Euclidean context.

4.1. The general Riemannian case. Let \((M, g)\) be a Riemannian manifold. Denote the induced Riemannian distance function by \( d^M \). Let \( S \subset M \) be an embedded connected submanifold.

There are two natural ways to induce a distance on \( S \):

1. Intrinsic: Consider \( S \) as a Riemannian submanifold of \( M \), i.e., endow \( S \) with the pullback metric \( i^*g \) along the inclusion \( i : S \to M \). Denote by \( d_S^{\text{int}} \) the Riemannian distance function induced by \( i^*g \).
2. Extrinsic: Consider \( S \) as a subspace of the metric space \((M, d^M)\). Denote by \( d_S^{\text{ext}} \) the restriction of \( d^M \) to \( S \times S \).

An immediate observation is that \( d_S^{\text{int}} \geq d_S^{\text{ext}} \). The question we pose is under what conditions, \( d_S^{\text{int}} = d_S^{\text{ext}} \).

In general, both equality and inequality may hold: For \( M = \mathbb{S}^2 \) endowed with the round metric and \( S \) a great circle, \( d_S^{\text{int}} = d_S^{\text{ext}} \). For \( M = \mathbb{R}^2 \) with the standard Euclidean metric and \( S = \mathbb{S}^1 \), \( d_S^{\text{int}} > d_S^{\text{ext}} \).

To state our results we need the following classical definitions:

**Definition 4.1.** Let \((M, g)\) be a Riemannian manifold. A subset \( C \) of \( M \) is said to be a geodesically convex if, given any two points in \( C \), there is a minimizing geodesic (in \( M \)) contained within \( C \) joining these two points.

**Definition 4.2.** A submanifold \( S \) of a Riemannian manifold \((M, g)\) is called totally geodesic if any geodesic on the submanifold \( S \) with its induced Riemannian metric is also a geodesic on the Riemannian manifold \((M, g)\). (This condition is equivalent to the vanishing of the second fundamental form of \( S \) in \( M \).)

To prove our results we shall need the following lemma which roughly says that paths that are close to being length-minimizers are within a narrow tubular neighborhood of a (minimizing) geodesic.

**Lemma 4.3** (Nearly length-minimizing paths are close to geodesics). Let \((M, g)\) be a complete Riemannian manifold, and let \( p, q \in M \). Then, for any \( \epsilon > 0 \) there...
exists a $\delta > 0$ (possibly dependent on $p$ and $q$) such that any path $\alpha$ joining $p$ and $q$ satisfying
\[ L(\alpha) < d(p, q) + \delta \]
is in an $\epsilon$-neighborhood of a minimizing geodesic $\gamma : I \to M$ joining $p$ and $q$. In particular, there exists a reparametrization $\alpha \circ \varphi : I \to M$ of $\alpha$ satisfying
\[ \sup_{t \in I} d(\alpha \circ \varphi(t), \gamma(t)) < \epsilon. \]

In this lemma there is no submanifold, so the distance has only one possible meaning— the Riemannian distance on $M$.

**Proof.** Assume by contradiction that the claim is false. Denote $r = d(p, q)$. Then, there exists an $\epsilon > 0$ and a sequence of paths $\alpha_n : I \to M$ joining $p$ and $q$, satisfying
\[ L(\alpha_n) \leq d(p, q) + \frac{1}{n}, \]
and $\alpha_n$ is not in an $\epsilon$-neighborhood of any minimizing geodesic.

Since $L(\alpha_n) \to r$, we can assume $L(\alpha_n) \leq 2r$, thus $\text{Image}(\alpha_n) \subseteq B^M(p, 2r)$ (the closed ball of radius $2r$ around $p$). By the completeness of $M$, it follows from the Hopf-Rinow theorem that $B^M(p, 2r)$ is compact.

Reparametrize $\alpha_1$ by arclength, i.e. assume $I = [0, L(\alpha_1)]$, and that $\alpha_1 : I \to M$ has a constant speed. For every $n \in \mathbb{N}$, reparametrize $\alpha_n : I \to M$ such that it has a constant speed $c_n = \|\dot{\alpha}_n\|$. Then
\[ 2L(\alpha_1) \geq 2r \geq L(\alpha_n) = c_n L(\alpha_1) \quad \Rightarrow \quad c_n \leq 2, \]
which implies that the $\alpha_n$ are equicontinuous, since
\[ d(\alpha(t), \alpha(s)) \leq L(\alpha_{[t,s]}) \leq 2(s - t). \]

By the Arzela-Ascoli theorem, there exists a subsequence (also denoted $\alpha_n$) converging uniformly to a path $\alpha : I \to M$. By the lower-semicontinuity of the length functional we deduce:
\[ L(\alpha) \leq \lim_{n \to \infty} L(\alpha_n) = d(p, q). \]
This implies that $\alpha$ is a length-minimizing curve between $p$ and $q$, hence its reparametrization by arclength $\alpha \circ \varphi$ is a geodesic.

Finally, the uniform convergence $\alpha_n \to \alpha$ yields a contradiction: there exists an $N$ such that for all $n > N$,
\[ \sup_{t \in I} d((\alpha_n \circ \varphi)(t), (\alpha \circ \varphi)(t)) < \epsilon. \]

\[ \square \]

We next prove the following:

**Proposition 4.** Let $S$ be a submanifold $S$ of a Riemannian manifold $(M, g)$. Then:
1. If $S$ is a geodesically convex subset of $M$, then $d^\text{int}_S = d^\text{ext}_S$. The reverse implication does not hold in general. The next assertion shows that the only obstruction for the reverse direction to hold, is topological.
2. If $S$ is topologically closed in $M$, then $d^\text{int}_S = d^\text{ext}_S$ if and only if $S$ is a geodesically convex subset of $M$.
3. If $d^\text{int}_S = d^\text{ext}_S$ then $S$ is a totally geodesic submanifold of $M$. The reverse implication does not hold in general.
4. Let $p, q \in S$. Assume there exists a unique minimizing geodesic $\gamma : I \to M$ connecting $p$ and $q$. If $\gamma \cap (M \setminus S) \neq \emptyset$. Then $d^\text{int}_S(p, q) > d^\text{ext}_S(p, q)$.
The last two statements hold also if we replace the existence of a unique geodesic with the existence of finitely many geodesics.

**Proof.** 1. The fact that geodesic convexity implies equality of the distances is immediate. A counter-example for the reverse implication is $M = \mathbb{R}^2, S = \mathbb{R}^2 \setminus (0,0)$.

2. Assume $d_S^{Int} = d_S^{Ext}$. Let $p, q \in S$. Let $\alpha_n : I \to S$ be a sequence of paths satisfying $L(\alpha_n) \to d(p,q)$. Then, by a similar argument to the one in the proof of Lemma 4.3, there is a subsequence $\alpha_n$ converging uniformly to a path $\alpha$. By the lower-semicontinuity of the length, $L(\alpha) = d(p,q)$, so $\alpha$ is minimizing, hence it is a reparametrization of a geodesic. Closedness of $S$ implies $\Image(\alpha) \subseteq S$.

3. Suppose that $d_S^{Int} = d_S^{Ext}$, and let $\alpha$ be a geodesic in $S$. Then, it is locally length-minimizing in $S$, and for small enough $t$,

$$L(\alpha|_{[0,t]}) = d_S^{Int}(\alpha(0), \alpha(t)) = d_S^{Ext}(\alpha(0), \alpha(t)) = d_M(\alpha(0), \alpha(t)).$$

So $\alpha|_{[0,t]}$ is length-minimizing path between $\alpha(0), \alpha(t)$ in $M$ with parameter proportional to arc length, hence it is a geodesic in $M$.

A counter-example for the reverse implication is $M = S^1, S = S^1 \setminus \{p\}$ (where $p$ is an arbitrary point in $S^1$). We will see another counter-example in the next section: $M = M_n$ and $S = \text{GL}_n^+$.

4. By assumption, there exists $t_0 \in I$ such that $\gamma(t_0) \in M \setminus \bar{S}$. Since $M \setminus \bar{S}$ is open, there is some open ball of $d^M$-radius $\epsilon$, $\gamma(t_0) \in B_\epsilon \subseteq M \setminus \bar{S}$.

By Lemma 4.3, $\exists \delta > 0$ such that if $\alpha$ is a path between $p, q$, $L(\alpha) < d^M(p,q) + \delta$ then $\alpha$ is in an $\epsilon$-neighborhood of some minimizing geodesic joining $p$ and $q$. By our assumption, there is only one minimizing geodesic between $p$ and $q$ in $M$, namely $\gamma$.

Thus, there exists a reparametrization of $\alpha$, $\alpha \circ \varphi : I \to M$, such that for every $t$,

$$d((\alpha \circ \varphi)(t), \gamma(t)) < \epsilon.$$  

In particular, $d((\alpha \circ \varphi)(t_0), \gamma(t_0)) < \epsilon$ implies that $(\alpha \circ \varphi)(t_0) \in B_\epsilon \subseteq M \setminus \bar{S}$.

This shows that any path $\alpha$ which is $\delta$-close to being a minimizer intersect $M \setminus \bar{S}$. Hence $d_S^{Int}(p,q) \geq d_S^{Ext}(p,q) + \delta$.

\[ \square \]

**4.2. Euclidean distances in $\text{GL}_n^+$.** Next, we consider the particular case where the Riemannian manifold is the vector space of $n \times n$ matrices endowed with the Euclidean metric. This is the case considered classically in the context of elastic strain measures, and whose shortcomings has motivated, in part, the consideration of alternative measures of strain.

We start with a few definitions:

**Definition 4.4 (Euclidean metric on $M_n$).** We denote by $(M_n, d^{\text{Euc}})$ the space of $n \times n$ real matrices endowed with the Euclidean distance. Note that the distance $d^{\text{Euc}}$ can be derived from a Riemannian metric $g$ given by

$$g_Z(X,Y) = \text{tr}(X^TY).$$

**Definition 4.5 (Extrinsic Metric on $\text{GL}_n^+$).** We denote by $(\text{GL}_n^+, d^{\text{ext}})$ the metric space of $n \times n$ invertible matrices with positive determinant, where $d^{\text{ext}}$ is the restriction of $d^{\text{Euc}}$, with $\text{GL}_n^+$ viewed as a subset of $\mathbb{R}^{n^2}$. 
Definition 4.6 (Intrinsic Metric on $GL_n^+$). Consider $(GL_n^+, g|_{GL_n^+})$ as an open submanifold of the Riemannian manifold $(M_n, g)$. That is, we endow $GL_n^+$ with the pullback metric $i^*g$ of the Euclidean metric $g$ along the inclusion $i : GL_n^+ \to (M_n, g)$. We denote by $d^\text{int}$ the distance function induced by the Riemannian metric $i^*g$.

We first observe that $d^\text{int} > d^\text{ext}$ for some pair of matrices. Indeed, for any $X, Y \in M_n$, the unique minimizing geodesic is the segment 
\[ [X, Y] = \{ X + t (Y - X) : t \in [0, 1] \}. \]
Since the sub-manifold $GL_n^+$ is not convex, there exist $A, B \in GL_n^+$, such that the segment $[A, B]$ intersects $GL_n^-$. By Item 4 in Proposition 4, 
\[ d^\text{ext}(A, B) < d^\text{int}(A, B). \]

However, we note the following:

Lemma 4.7. Let $A, B \in GL_n^+$. If $[A, B] \subset GL_n^+$, then $d^\text{int}(A, B) = d^\text{ext}(A, B)$.

Proof. This is obvious, since $[A, B]$ is an extrinsic length-minimizing path which stays in the submanifold. \qed

We next observe that both right- and left-multiplications by elements of $SO_n$ are isometries of $(GL_n^+, i^*g)$. Hence, they are isometries of the metric space $(GL_n^+, d^\text{int})$ (any Riemannian isometry is an isometry of the induced distance function). It follows that $d^\text{int}$ is both left- and right-$SO_n$ invariant.

In particular, let $A = U\Sigma V^T$ be an SVD of $A \in GL_n^+$. By Proposition 2, 
\[ d^\text{int}(A, SO_n) = d^\text{int}(\Sigma, SO_n), \]
hence, as before, the problem of computing the distance of $A \in GL_n^+$ from $SO_n$ (and finding the minimizer) can be reduced to positive-definite diagonal matrices $\Sigma$. We now give a short proof that $I$ is the unique matrix closest to $\Sigma$ with respect to the extrinsic distance, that is, 
\[ d^\text{ext}(\Sigma, SO_n) = d^\text{ext}(\Sigma, I) = \| \Sigma - I \|_F. \]
Indeed, since 
\[ \| A - Q \|_F^2 = \| A \|^2_F + \| Q \|^2_F - 2 \langle A, Q \rangle_F = \| A \|^2_F + n - 2 \text{tr}(A^T Q), \] (9)
it follows that given $A \in M_n$, minimizing $\| A - Q \|_F^2$ over $Q \in SO_n$ is equivalent to maximizing the linear functional $\varphi_A(Q) = \langle A, Q \rangle_F = \text{tr}(A^T Q)$. For diagonal and positive-definite $\Sigma$, 
\[ \varphi_\Sigma(Q) = \text{tr}(\Sigma^T Q) = \sum_{i=1}^n \sigma_i Q_{ii} \leq \sum_{i=1}^n \sigma_i = \text{tr}(\Sigma^T) = \varphi_\Sigma(I), \]
where the inequality follows from the fact that $Q$ is orthonormal, hence $|Q_{ij}| \leq 1$. The unique maximizer is $Q = I$, hence $I$ is the unique matrix closest to $\Sigma$.

Theorem 4.8. Let $\Sigma \in GL_n^+$ be a diagonal matrix with strictly positive entries on the diagonal. Then the unique minimizer of the intrinsic distance of $\Sigma$ from $SO_n$ is $I$, and 
\[ d^\text{int}(\Sigma, SO_n) = d^\text{ext}(\Sigma, SO_n). \]
Proof. Since \( \text{Psym}_n \) is closed under convex combinations, it follows that \( [\Sigma, I] \subseteq GL^+_n \). By Lemma 4.7, it follows that \( d^{\text{int}}(\Sigma, I) = d^{\text{ext}}(\Sigma, I) \).

The inequality \( d^{\text{int}} \geq d^{\text{ext}} \) implies that

\[
\text{dist}^{\text{int}}(\Sigma, SO_n) \geq \text{dist}^{\text{ext}}(\Sigma, SO_n).
\]

Hence,

\[
\text{dist}^{\text{ext}}(\Sigma, SO_n) = d^{\text{ext}}(\Sigma, I) = d^{\text{int}}(\Sigma, I) \geq d^{\text{int}}(\Sigma, SO_n) \geq d^{\text{ext}}(\Sigma, SO_n).
\]

So, \( I \) is a minimizer and \( \text{dist}^{\text{int}}(\Sigma, SO_n) = \text{dist}^{\text{ext}}(\Sigma, SO_n) \).

The uniqueness of the minimizer for \( \text{dist}^{\text{int}} \) follows from the uniqueness of the minimizer for \( \text{dist}^{\text{ext}} \): Let \( Q \in SO_n \) be a minimizer for the distance \( \text{dist}^{\text{int}} \) of \( \Sigma \) from \( SO_n \). Then,

\[
\text{dist}^{\text{ext}}(\Sigma, SO_n) = \text{dist}^{\text{int}}(\Sigma, SO_n) = d^{\text{int}}(\Sigma, Q) \geq d^{\text{ext}}(\Sigma, Q),
\]

hence \( Q \) is also a minimizer for \( \text{dist}^{\text{ext}} \), and by the uniqueness of the extrinsic minimizer, \( Q = I \).

Corollary 4. Let \( A \in GL^+_n \). Then

\[
d^{\text{int}}(A, SO_n) = d^{\text{ext}}(A, SO_n),
\]

and the closest matrix to \( A \) in \( SO_n \) with respect to both distances is the same—it is the orthogonal polar factor of \( A \). Moreover,

\[
d^{\text{int}}(A, SO_n) = \| \sqrt{A^T A} - I \|_F. \tag{10}
\]

Proof. This is an immediate consequence of Proposition 2 applied to both distances, together with Theorem 4.8. The fact that the orthogonal polar factor is the unique minimizer follows from the same considerations as in the proof of Corollary 3. \( \square \)

5. Additional results concerning invariant distances.

5.1. Bi-invariance. As always in mathematics, the most symmetric structures are the most easy to handle. In [12], the authors consider either left- and right-\( GL_n \)-invariance as natural requirements on a metric on \( GL_n \). Their choice is motivated by physical considerations.

A natural question is: does there exist a distance function on \( GL_n \) that is more symmetric than the ones we have considered? In this section we show here that there are no bi-invariant distance functions on \( GL_n \) that are compatible with the standard topology. (The fact there is no bi-invariant Riemannian metric is common knowledge.)

Theorem 5.1. There is no bi-invariant distance function on \( GL_n \) generating the standard topology on \( GL_n \) (the subspace topology induced by the inclusion \( GL_n \to \mathbb{R}^n^2 \)).

Proof. The essential point is the existence of a non-trivial conjugacy class whose closure contains the identity. Assume, by contradiction, there is a bi-invariant distance function \( d \) compatible with the standard topology. Consider the following
matrices:
\[
D = \begin{pmatrix}
2^{-1} & 0 & \cdots & 0 \\
0 & 2^{-2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 2^{-n}
\end{pmatrix}
\quad \text{and} \quad
A = \begin{pmatrix}
1 & 1 & 0 & \cdots & 0 \\
0 & 1 & 1 & \cdots & 0 \\
0 & 0 & 1 & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}.
\]

An explicit calculation yields,
\[
D^{-n}AD^n = \begin{pmatrix}
1 & 2^{-n} & 0 & \cdots & 0 \\
0 & 1 & 2^{-n} & \cdots & 0 \\
0 & 0 & 1 & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}.
\]

Since \(d\) generates the standard topology on \(GL_n\), it follows that
\[
\lim_{n \to \infty} d(D^{-n}AD^n, I) = 0.
\]
On the other hand, bi-invariance implies that for every \(n\),
\[
d(D^{-n}AD^n, I) = d(D^{-n}AD^n, D^{-n}ID^n) = d(A, I) \neq 0,
\]
which is a contradiction.

As an immediate corollary we obtain the classical result:

**Corollary 5.** There is no bi-invariant Riemannian metric on \(GL_n\).

### 5.2. Inverse-invariance

When it comes to physical applications, a major draw-back of the Euclidean strain measure is that it does not diverge in the limit where the linear map is singular. From a physical viewpoint, we expect a strain measure \(\text{dist}(A, SO_n)\) to diverge when \(A\) either tends to infinity (expansion), or when it tends toward singularity (contraction). The strain measure is said to be inverse-invariant if

\[
\text{Strain}(A) = \text{Strain}(A^{-1}),
\]

for every \(A \in GL_n^+\).

As noted in the Introduction, the strains obtained via the metrics considered in Section 2 are all inverse-invariant. The essential reason behind this phenomenon, is the extreme symmetry of these metrics. Specifically, we have the following very general assertion:

**Proposition 5.** Let \(G\) be a group and \(H \subseteq G\) a subgroup. Let \(d\) be a left-\(G\)-right-\(H\)-invariant distance function on \(G\). Then,

\[
\text{dist}(g, H) = \text{dist}(g^{-1}, H).
\]

Moreover, if \(h\) is a closest element to \(g\) in \(H\), then \(h^{-1}\) is a closest element to \(g^{-1}\) in \(H\).

**Proof.** Using the assumed invariances,

\[
\text{dist}(g, H) = \inf_{h \in H} d(g, h) = \inf_{h \in H} d(e, g^{-1}h) = \inf_{h \in H} d(h^{-1}, g^{-1}) = \text{dist}(g^{-1}, H)
\]

The equality \(d(g, h) = d(g^{-1}, h^{-1})\) (for \(h \in H\)) implies the correspondence between closest elements to \(g\) and \(g^{-1}\). \(\square\)
Note how this observation implies, without any computation, that certain strain measures are inverse-invariant (the actual form of the strain measures is irrelevant).

Another means for obtaining inverse-invariant strain measures is to consider distance functions \( d \) on \( \text{GL}_n \) that are inverse-invariant, i.e.,

\[
d(A, B) = d(A^{-1}, B^{-1}).
\]

Indeed, the inverse-invariance of the distance implies the inverse-invariance of the strain measure, as

\[
dist(A, \text{SO}_n) = dist(A^{-1}, \text{SO}_n^{-1}) = dist(A^{-1}, \text{SO}_n).
\]

As we will see, the requirement for inverse-invariance of the distance is far less restrictive than left-\( \text{GL}_n \)- and right \( \text{O}_n \)-invariance.

In a search for maximally-symmetric distance functions, we first look for inverse-invariant distances (or metrics) possessing additional symmetries. In Propositions 9–10 we prove that in any Lie group \( G \), a left-invariant metric (distance) is inverse-invariant if and only if it is bi-invariant. Since there are no bi-invariant metrics/distances on \( \text{GL}_n \), it follows that no left-\( \text{GL}_n \)- and inverse-invariant metric/distance exists either.

In particular, it follows that the distance functions induced by the metrics considered in Sections 2 and 3 are not inverse-invariant. That is, the inverse-invariance of the strain measure does not result from the inverse-invariance of the distance, but rather from the left-\( \text{GL}_n \) and right-\( \text{SO}_n \) invariance of the metrics, as shown in Proposition 5 (also observed in [12, Section 3.2 Eq. (23)]).

There is a systematic way of constructing inverse-invariant distances from arbitrary distances. Denote the inverse automorphism by \( i \). Since \( i \) is a diffeomorphism of finite order, given any distance \( d \) on \( \text{GL}_n \), it is possible to construct an inverse-invariant distance via symmetrization,

\[
\hat{d}(A, B) = d(A, B) + d(A^{-1}, B^{-1}).
\]

It is easy to see that \( \hat{d} \) generates the same topology as \( d \), and it is of course inverse-invariant.

A similar construction can be carried out for Riemannian metrics on \( \text{GL}_n \). Given any metric \( g \), the metric \( g + i^*(g) \) is inverse-invariant, and induces the standard topology on \( \text{GL}_n \), as does any Riemannian metric.

In the following subsections, we analyze how these two different methods for generating inverse-invariant distances affect the strain measure. We will see that if we start from distances/metrics having certain symmetries, then their symmetrizations possess corresponding symmetries as well.

5.2.1. **Inverse-invariant distances.** The next lemma considers a setting that generalizes our treatment of left-\( \text{GL}_n \)- and right-\( \text{O}_n \) invariant distance functions.

**Lemma 5.2.** Let \( G \) be a group and let \( H \subseteq G \) be a subgroup. Let \( d \) be a left-\( G \)- and right-\( H \)-invariant distance function on \( G \). Let \( \hat{d} \) be its symmetrization. Then,

\[
\text{dist}_{\hat{d}}(g, H) = 2 \text{dist}_d(g, H).
\]

Moreover, an element \( h \in H \) is a closest element in \( H \) to \( g \) with respect to \( d \) if and only if it is a closest element with respect to \( \hat{d} \).
Proof. The assumed symmetries imply $d(g,h) = d(g^{-1}, h^{-1})$ for any $g \in G$ and $h \in H$ (see the proof of Proposition 5). Hence,

$$
\text{dist}_{\tilde{d}}(g, H) = \inf_{h \in H} \tilde{d}(g, h)
= \inf_{h \in H} (d(g, h) + d(g^{-1}, h^{-1}))
= \inf_{h \in H} 2d(g, h)
= 2 \text{dist}_d(g, H).
$$

To prove the second part, note that for $h \in H$, $\tilde{d}(g, h) = \text{dist}_{\tilde{d}}(g, H)$ if and only if $2d(g, h) = 2 \text{dist}_d(g, H)$.

The above lemma implies that there is not much interest in symmetrizing left-$G$- and right-$H$-invariant distance functions, since the symmetrizations give rise to essentially identical notions of distance from $H$.

We turn to analyze symmetrization within the context of intrinsic versus extrinsic Euclidean distances. As above, we provide a slightly more general treatment that considers the relevant symmetries. Since this setting possesses less symmetries than the one considered above, we will have to use properties that are specific to the Euclidean distance; in particular, SVD plays an important role.

**Lemma 5.3.** Let $G$ be a group and let $H \subseteq G$ be a subgroup. Let $d$ be a bi-$H$-invariant distance on $G$. Then, its symmetrization $\tilde{d}$ is also bi-$H$-invariant.

**Proof.** For every $h \in H, x, y \in G$,

$$
\tilde{d}(hx, hy) = d(hx, hy) + d(x^{-1}h^{-1}, y^{-1}h^{-1}) = d(x, y) + d(x^{-1}, y^{-1}) = \tilde{d}(x, y).
$$

Right-$H$-invariance is proved similarly.

In our context, since both intrinsic and extrinsic Euclidean distances are bi-SO$_n$ invariant, their symmetrizations are also bi-SO$_n$ invariant.

By Proposition 2, $\text{dist}_{\tilde{d}}(A, \text{SO}_n) = \text{dist}_{\tilde{d}}(\Sigma, \text{SO}_n)$, where $A = U\Sigma V^T$ is any SVD of $A$. By the results in Section 4.2, $I$ is the closest matrix to both $\Sigma$ and $\Sigma^{-1}$ with respect to both intrinsic and extrinsic Euclidean distances.

Hence, for every $Q \in \text{SO}_n$,

$$
\tilde{d}(\Sigma, Q) = d(\Sigma, Q) + d(\Sigma^{-1}, Q^{-1}) \geq d(\Sigma, I) + d(\Sigma^{-1}, I) = \tilde{d}(\Sigma, I),
$$

from which follows that $I$ is the matrix in $\text{SO}_n$ that is the closest to $\Sigma$ with respect to $\tilde{d}$, and

$$
\text{dist}_{\tilde{d}}(A, \text{SO}_n) = \text{dist}_{\tilde{d}}(\Sigma, \text{SO}_n) = \tilde{d}(\Sigma, I) = d(\Sigma, I) + d(\Sigma^{-1}, I) = \text{dist}_d(A, \text{SO}_n) + \text{dist}_d(A^{-1}, \text{SO}_n).
$$

Again, we obtain that the matrix in $\text{SO}_n$ that is the closest to $A$ is the orthogonal polar factor of $A$.

The symmetrization of the Euclidean distance gives a truly different notion of strain measure, as it penalizes equally both expansions and contractions. Thus, it can be considered an improved strain measure. At the same time, it preserves the symmetries pertinent to the Euclidean metric—frame invariance and material isotropy.
5.2.2. Inverse-invariant metrics. We now turn to the symmetrization of Riemannian metrics on $\text{GL}_n$. Given any metric $g$, the metric $\tilde{g} = g + i^*(g)$ is inverse-invariant. It is common knowledge (see sketch of proof in [15]) that the space of inverse-invariant metrics for any Lie group is infinite-dimensional, hence, we have to address the question of finding natural inverse-invariant metrics. We note that in general, the distance function induced by the symmetrized metric $\tilde{g}$ is not the symmetrization of the distance function induced by $g$. Hence, the analysis of the previous subsection is not applicable.

We start by showing that the symmetrized metric inherits some of the symmetries of the original metric.

Lemma 5.4. Let $G$ be a Lie group and let $H \subseteq G$ be a subgroup. Let $g$ be a bi-$H$-invariant metric on $G$. Its symmetrization $\tilde{g}$ is also bi-$H$-invariant.

Proof. Let $h \in H$. Then

$$L_h^* (g + i^* g) = L_h^* g + L_h^* (i^* g) = g + (i \circ L_h)^* g.$$ 

Since $i \circ L_h = R_{h^{-1}} \circ i$,

$$L_h^* (g + i^* g) = g + (R_{h^{-1}} \circ i)^* g = g + i^* (R_{h^{-1}}^* g) = g + i^* g.$$ 

The proof that $R_h^* (g + i^* g) = g + i^* g$ is similar. \qed

It follows that the symmetrizations of all the metrics considered in sections 2 and 4.2 are bi-$\text{O}_n$ invariant.

In the remaining part of this section we study the symmetrization of the metrics considered in Section 2.

As mentioned above, the symmetrized metrics, unlike the original metrics, are not left-$\text{GL}_n$ invariant, hence, our analysis of the geodesics is not applicable. However, the symmetrized metrics share three important properties with the original metrics:

1. $(\text{GL}_n^+, \tilde{g})$ is complete.
2. The symmetric and the skew-symmetric matrices are orthogonal with respect to $\tilde{g}$.
3. $\alpha(t) = e^{tV}$ is a $\tilde{g}$-geodesic for any symmetric matrix $V$.

We start by showing the orthogonality of symmetric and skew-symmetric matrices. For any $A \in \text{sym}_n$ and $B \in \text{O}_n$,

$$(i^* g)_{1}(A, B) = g_{1}(di_{1}(A), di_{1}(B)) = g_{1}(-A, -B) = 0,$$ 

hence

$$\tilde{g}_{1}(A, B) = g_{1}(A, B) + (i^* g)_{1}(A, B) = 0.$$ 

We proceed to prove the completeness of $(\text{GL}_n^+, \tilde{g})$. First note that completeness of $g$ implies completeness of $i^* g$, since $i : (\text{GL}_n, i^* g) \to (\text{GL}_n, g)$ is an isometry. So, it suffices to prove that if two Riemannian manifolds $(M, g_1)$ and $(M, g_2)$ are complete then so is $(M, g_1 + g_2)$.

Lemma 5.5. Let $M$ be a smooth manifold, and let $g_1, g_2$ be Riemannian metrics on $M$. If either $(M, g_1)$ or $(M, g_2)$ is complete, then $(M, g_1 + g_2)$ is complete.
Proof. Let \( p, q \in M \). For any path \( \alpha \) from \( p \) to \( q \),
\[
L_{g_1+g_2}(\alpha) = \int \sqrt{g_1(\dot{\alpha}(t), \dot{\alpha}(t)) + g_2(\dot{\alpha}(t), \dot{\alpha}(t))} \\
\geq \int g_1(\dot{\alpha}(t), \dot{\alpha}(t)) \\
= L_{g_1}(\alpha).
\]
Similarly \( L_{g_1+g_2}(\alpha) \geq L_{g_2}(\alpha) \). Without loss of generality, Assume that \((M, g_1)\) is complete. For any \( p, q \in M \),
\[
d^{g_1}(p, q) = \inf_{\alpha: p \to q} L_{g_1}(\alpha) \leq \inf_{\alpha: p \to q} L_{g_1+g_2}(\alpha) = d^{g_1+g_2}(p, q)
\]
By the Hopf-Rinow theorem, a Riemannian manifold \((M, g)\) is complete if and only if closed and \( g \)-bounded sets are compact. Let \( A \subseteq M \) be a closed and \((g_1 + g_2)\)-bounded set. Boundedness implies that there exists a point \( p \in M \) and a number \( R > 0 \), such that \( d^{g_1+g_2}(a, p) \leq R \) for every \( a \in A \).
Since \( d^{g_1}(a, p) \leq d^{g_1+g_2}(a, p) \leq R \), it follows that \( A \) is also \( g_1 \)-bounded. Since \((M, g_1)\) is complete, \( A \) is compact, hence \( g_1 + g_2 \) is complete as well.

It remains to show that \( \alpha(t) = e^{tV} \) is a geodesic for symmetric \( V \). We first note that \( \alpha(t) = i(e^{tV}) = e^{-tV} \) is a geodesic of \( i^*g \), since \( i : (\mathop{GL_n},i^*g) \to (\mathop{GL_n},g) \) is an isometry. Reversing time, we get that \( \alpha(t) = e^{tV} \) is a geodesic of both \( g \) and \( i^*g \).
Note that every geodesic is parametrized by a parameter proportional to arclength, i.e., its speed \( \|\dot{\alpha}(t)\| \) is constant. In this particular case, the speeds are the same when measured with respect to both metrics, since \( g_I = (i^*g)_I \), hence \( \|\dot{\alpha}(0)\| \) is independent of the metric chosen. It turns out that in this particular situation, \( \alpha \) is a geodesic with respect to the metric \( g + i^*g \).

**Lemma 5.6.** Let \( M \) be a smooth manifold, and let \( g_1, g_2 \) be Riemannian metrics on \( M \). Assume \( \beta(t) \) is a geodesic for both \( g_1 \) and \( g_2 \), and that its speed is the same with respect to both metrics. Then \( \beta(t) \) is also a \((g_1 + g_2)\)-geodesic.

Proof. The inequality: \( \sqrt{a + b} \geq \frac{1}{\sqrt{2}}(\sqrt{a} + \sqrt{b}) \) implies that for any path \( \alpha \) in \( M \),
\[
L^{g_1+g_2}(\alpha) = \int \sqrt{g_1(\dot{\alpha}(t), \dot{\alpha}(t)) + g_2(\dot{\alpha}(t), \dot{\alpha}(t))} \\
\geq \frac{1}{\sqrt{2}} \cdot \left( \int \sqrt{g_1(\dot{\alpha}(t), \dot{\alpha}(t))} + \int \sqrt{g_2(\dot{\alpha}(t), \dot{\alpha}(t))} \right) \\
= \frac{1}{\sqrt{2}} \cdot (L^{g_1}(\alpha) + L^{g_2}(\alpha)).
\]
Since \( \beta \) is a geodesic with respect to both \( g_1 \) and \( g_2 \), it is locally length-minimizing with respect to both metrics; for small enough \( t \), \( L^g(\beta|_{[0,t]}) = d^{g}(\beta(0), \beta(t)) = tc \), for some constant \( c \). Let \( \alpha \) be any path connecting \( \beta(0), \beta(t) \). By our assumption, \( \sqrt{g_1(\beta(t), \dot{\beta}(t))} = \sqrt{g_2(\dot{\beta}(t), \ddot{\beta}(t))} = c \). Note that
\[
L^{g_1+g_2}(\beta|_{[0,t]}) = \int_0^t \sqrt{g_1(\dot{\beta}(t), \dot{\beta}(t)) + g_2(\dot{\beta}(t), \dot{\beta}(t))}
\]
where the last inequality uses (13). Thus, $\beta$ locally minimizes length with respect to $g_1 + g_2$, hence it is a geodesic. (Note it is parametrized proportional to arclength with respect to $g_1 + g_2$.)

Next, we imitate from Section 3 the argument for finding the geodesic distance from $\mathrm{SO}_n$. By Lemma 5.4, we can use Proposition 2 to reduce again the question to diagonal matrices. Since the derivation of the strain measure uses only the three properties of the metric mentioned above, it works in exactly the same manner for the symmetrized metric.

There is just one delicacy. The proof hinges on the fact that $\alpha(t) = Qe^{tV}$ is a $\tilde{g}$-geodesic for any symmetric $V$ and $Q \in \mathrm{SO}_n$ (note that at this stage we don’t yet know that $Q = I$). From Lemma 5.6 follows that $e^{tV}$ is a $\tilde{g}$-geodesic. Since (Lemma 5.4 again) $\tilde{g}$ is bi-$O_n$-invariant it follows that $\alpha$ is also a $\tilde{g}$ geodesic. Following the rest of the proof, the only difference is at the final stage, when evaluating the speed $\|\dot{\alpha}(0)\|_I$, where $\tilde{g}_I$-is scaled by $\sqrt{2}$,

$$\|V\|_I = \sqrt{g_I(V,V) + (i* g_I)(V,V)} = \sqrt{2} \cdot g_I(V,V).$$

Hence, the strain measure is multiplied by a factor of $\sqrt{2}$.

**Appendix A. Calculating the geodesics.**

**A.1. Analysis of the geodesic equations.** Let $g_I$ be an inner-product on $M_n = T_I \mathrm{GL}_n$ given by the form (4), and let $g$ be the Riemannian metric on $\mathrm{GL}_n$ which is the left-translation of $g_I$.

First, we need the following result.

**Proposition 6.** For every $X,Y \in T_I \mathrm{GL}_n$,

$$g_I([X,Y],X) = \frac{\beta - \gamma}{2\beta} g_I([X,X^T],Y).$$

(Here $[X,Y]$ denotes the standard matrix commutator, i.e $[X,Y] = XY - YX$.)

**Proof.** Note first that

$$g_I([X,Y],X) = \beta tr(symX sym[Y,X]) + \gamma tr(skew X skew[Y,X]),$$

and

$$g_I([X,X^T],Y) = \beta tr(sym[X,X^T]symY).$$

Now,

$$tr(symX sym[Y,X]) = \frac{1}{4} tr((X + X^T)(YX - XY + X^TY^T - Y^TX^T))$$

$$= \frac{1}{2} tr([X,X^T]Y) = \frac{1}{2\beta} g_I([X,X^T],Y),$$

where

$$= \sqrt{2}ct = \frac{1}{\sqrt{2}}(tc + tc)$$

$$= \frac{1}{\sqrt{2}} \cdot (d^{g_1}(\beta(0),\beta(t)) + d^{g_2}(\beta(0),\beta(t))),$$

$$\leq \frac{1}{\sqrt{2}} \cdot (L^{g_1}(\alpha) + L^{g_2}(\alpha)) \leq L^{g_1 + g_2}(\alpha),$$

where the last inequality uses (13). Thus, $\beta$ locally minimizes length with respect to $g_1 + g_2$, hence it is a geodesic. (Note it is parametrized proportional to arclength with respect to $g_1 + g_2$.)
and
\[
\text{tr}(\text{skew } X \text{ skew}[Y, X]) = \frac{1}{4} \text{tr}((X - X^T)(YX - XY - XT^T + YT^T))
\]
\[
= -\frac{1}{2} \text{tr}([X, X^T]Y) = -\frac{1}{2\beta} g_I([X, X^T], Y),
\]

hence
\[
g_I(X, [Y, X]) = \frac{\beta - \gamma}{2\beta} g_I([X, X^T], Y).
\]

\[\square\]

Let \{a_\alpha\} be a \(g_I\)-orthonormal basis for \((T_1 \text{GL}_n, g)\). Since \(g\) is left-invariant, \(e_\alpha = d(L_g)_{e}(a_\alpha)\) is an orthonormal frame for \((T_\text{GL}_n, g)\). Let \{\(\vartheta^\alpha\)\} be the orthonormal co-frame,
\[
\vartheta^\alpha|_A = g_A(e_\alpha|A, \cdot) = g_A(i_A(Aa_\alpha), \cdot),
\]

which implies that
\[
\vartheta^\alpha|_A(i_A(AX)) = g_I(a_\alpha, X).
\]

The Riemannian connection is represented by an anti-symmetric matrix of 1-forms, \{\(\omega^{\alpha \beta}\)\}, defined by
\[
\nabla_{e_\alpha}e_\beta = \omega^{\gamma \beta}(e_\alpha)e_\gamma,
\]

and satisfying Cartan’s first structural equation,
\[
d\vartheta^\alpha + \omega^{\alpha \beta} \wedge \vartheta^\beta = 0.
\]

Noting that,
\[
d\vartheta^\alpha(e_\mu, e_\nu) = (\vartheta^\alpha(e_\nu))e_\mu - (\vartheta^\alpha(e_\mu))e_\nu - \vartheta^\alpha([e_\mu, e_\nu])
\]
\[
= \delta^\gamma_\mu e_\nu - \delta^\gamma_\nu e_\mu - \vartheta^\alpha([e_\mu, e_\nu]) - g(e_\alpha, [e_\mu, e_\nu]),
\]

and that
\[
g(e_\alpha, [e_\mu, e_\nu]) = g_I(a_\alpha, [a_\mu, a_\nu]), \quad (14)
\]

we get
\[
-g_I(a_\alpha, [a_\mu, a_\nu]) + \omega^{\alpha \nu}(e_\mu) - \omega^{\alpha \mu}(e_\nu) = 0.
\]

Equality (14) holds because Lie brackets of left-invariant vector fields are left-invariant, together with the well-known fact that the Lie algebra commutator of \(\text{GL}_n\) is merely the standard matrix commutator in \(M_n\), i.e \([e_\mu, e_\nu](e) = [a_\mu, a_\nu]\) \[8, p. 193\].

Rotating the indexes,
\[
-g_I(a_\mu, [a_\alpha, a_\nu]) + \omega^{\mu \nu}(e_\alpha) - \omega^{\mu \alpha}(e_\nu) = 0
\]
\[
-g_I(a_\nu, [a_\mu, a_\alpha]) + \omega^{\nu \alpha}(e_\mu) - \omega^{\nu \mu}(e_\alpha) = 0.
\]

Adding the three equations and renaming the indexes,
\[
2\omega^{\gamma \beta}(e_\alpha) = g_I(a_\alpha, [a_\gamma, a_\beta]) + g_I(a_\gamma, [a_\alpha, a_\beta]) + g_I(a_\beta, [a_\gamma, a_\alpha]).
\]

That is,
\[
\nabla_{e_\alpha}e_\beta = \Gamma^\gamma_{\alpha \beta} e_\gamma,
\]

where
\[
\Gamma^\gamma_{\alpha \beta} = \frac{1}{2} (g_I(a_\alpha, [a_\gamma, a_\beta]) + g_I(a_\gamma, [a_\alpha, a_\beta]) + g_I(a_\beta, [a_\gamma, a_\alpha]))
\]

are constant coefficients.

Consider now a geodesic curve, \(\gamma : I \rightarrow \text{GL}_n\), where
\[
\dot{\gamma}(t) = p^\alpha(t) e_\alpha|_{\gamma(t)}.
\]
The geodesic equation for the coefficients \( p^\alpha(t) \) is
\[
\dot{p}^\alpha(t) + \Gamma^\alpha_{\beta\gamma}(t) p^\beta(t) p^\gamma(t) = 0.
\]
Exploiting the symmetries of \( \Gamma \) and the symmetry of the geodesic equation,
\[
\dot{p}^\alpha(t) + g_I(a_\alpha, [a_\gamma, a_\beta]) p^\alpha(t) p^\beta(t) = 0.
\]
This is a set of \( n^2 \) quadratic equations with constant coefficients.

Multiplying this equation by \( a_\gamma \), setting \( X(t) = p^\alpha(t) a_\alpha \), which is a curve in \( T_I GL_n \),
\[
\dot{X}(t) + g_I(X(t), [a_\gamma, X(t)]) a_\gamma = 0,
\]
which by Proposition 6,
\[
0 = \dot{X} + g_I(X, [a_\gamma, X]) a_\gamma = \dot{X} + \frac{\beta - \gamma}{2\beta} g_I([X, X^T], a_\gamma) a_\gamma,
\]

namely,
\[
\dot{X} = \kappa (X^T X - XX^T), \quad (15)
\]
where
\[
\kappa = \frac{\beta - \gamma}{2\beta}.
\]

Equation (15) is an ordinary differential system in the vector space \( M_n \).

The geodesic \( \gamma : I \to GL_n \) is related to \( X(t) \) via,
\[
\dot{\gamma}(t) = i_{\gamma(t)}(\dot{\gamma}(t), \gamma(t) X(t)). \quad (16)
\]

A.2. Solution of geodesic equations. The factor \((\beta - \gamma)/2\beta\) has for effect to rescale time. We start by ignoring it.

**Proposition 7.** The solution to
\[
\dot{X} = X^T X - XX^T \quad X(0) = X_0,
\]
where \( X : I \to M_n \) is
\[
X(t) = \exp(t(X_0^T - X_0))X_0 \exp(t(X_0 - X_0^T)).
\]

**Proof.** Clearly, the initial conditions are satisfied. Differentiating with respect to \( t \) we get
\[
\dot{X}(t) = \exp(t(X_0^T - X_0))(X_0^T - X_0)X_0 \exp(t(X_0 - X_0^T))
+ \exp(t(X_0^T - X_0))X_0(X_0 - X_0^T) \exp(t(X_0 - X_0^T))
= \exp(t(X_0^T - X_0))(X_0^T X_0 - X_0 X_0^T) \exp(t(X_0 - X_0^T)).
\]
It only remains to insert \( \exp(t(X_0 - X_0^T)) \) \( \exp(t(X_0^T - X_0)) \) inside the products in the middle term to get
\[
\dot{X}(t) = X^T(t) X(t) - X(t) X^T(t).
\]

Please note that this exponential is the “standard” matrix exponential, i.e., the one obtained from integral curves of left-invariant vector fields. It is not the exponential map of the \( g \)-geodesics.

**Corollary 6.** The solution to (15) is
\[
X(t) = \exp(\kappa t(X_0^T - X_0))X_0 \exp(\kappa t(X_0 - X_0^T)).
\]
Proposition 8. Let $\gamma : I \to GL_n$ be the $g$-geodesic, 
\[ \dot{\gamma}(t) = i_{\gamma(t)}(\gamma(t), \gamma(t)X(t)). \]
satisfying the initial conditions 
\[ \gamma(0) = e \quad \text{and} \quad \dot{\gamma}(0) = X_0. \]

Then, 
\[ \gamma(t) = \exp((1 - \kappa)tX_0 + \kappa tX_0^T) \exp(\kappa t(X_0 - X_0^T)). \]

Proof. Clearly, the initial conditions are satisfied. Differentiating with respect to $t$ we get 
\[ \dot{\gamma}(t) = (T\gamma)_t(\partial_t) \]
\[ = (T\gamma)_t([t + s]) \]
\[ = [\gamma(t) + \exp((1 - \kappa)tX_0 + \kappa tX_0^T)X_0 \exp(\kappa t(X_0 - X_0^T))]s \]
\[ = i_{\gamma(t)}(\gamma(t), \exp((1 - \kappa)tX_0 + \kappa tX_0^T)X_0 \exp(\kappa t(X_0 - X_0^T))) \]
\[ = i_{\gamma(t)}(\gamma(t), \gamma(t)X(t)). \]

\[ \square \]

Appendix B. Inverse-invariant metrics on Lie groups.

Proposition 9. Let $G$ be a Lie group. A left- (or right-)invariant metric on $G$ is inverse-invariant if and only if it is bi-invariant.

Proof. Note that $\text{inv} = R_{s^{-1}} \circ \text{inv} \circ L_{s^{-1}}$, 
\[ R_{s^{-1}} \circ \text{inv} \circ L_{s^{-1}}(g) = R_{s^{-1}} \circ \text{inv}(s^{-1}g) = R_{s^{-1}}(g^{-1}s) = g^{-1}. \]

By the chain rule, 
\[ (d\text{inv})_s = (dR_{s^{-1}})_e \circ (d\text{inv})_e \circ (dL_{s^{-1}})_s. \]

Since $(d\text{inv})_e : T_eG \to T_eG$ is the additive inverse operation ($v \mapsto -v$), 
\[ (d\text{inv})_s = -(dR_{s^{-1}})_e \circ (dL_{s^{-1}})_s. \]

It follows at once that a bi-invariant metric is inverse-invariant.

Conversely, assume the metric is both left- and inverse-invariant. Then $\forall s \in G$ , 
$(dR_s)_e$ is an isometry, hence the metric is right-invariant as well. \[ \square \]

Proposition 10. Let $G$ be a Lie group. A left- (or right-)invariant distance function $d$ on $G$ is inverse-invariant if and only if it is bi-invariant.

Proof. Assume $d$ is left- and inverse-invariant. Since $\text{inv} = R_{s^{-1}} \circ \text{inv} \circ L_{s^{-1}}$, $d$ is also right-invariant.

Conversely, assume $d$ is bi-invariant. Then, 
\[ d(x, y) = d(1, x^{-1}y) = d(y^{-1}, x^{-1}) = d(x^{-1}, y^{-1}). \]

In fact, Proposition 10 implies Proposition 9 by virtue of the Myers-Steenrod theorem whereby every isometry of $d$ is a Riemannian isometry.
Appendix C. A minimizing geodesic from a point to a submanifold is orthogonal to the submanifold. We prove the following lemma, which plays a central role in the derivation of the geodesic distance.

**Lemma C.1.** Let $M$ be a complete Riemannian manifold. Let $S \subseteq M$ be a submanifold, and let $p \in M \setminus S$. Assume $q \in S$ is a point on $S$ satisfying $d(p,q) = \text{dist}(p,S)$ (there is always such a point $q$ if $S$ is compact). Let $\alpha$ be a minimizing geodesic connecting $p$ and $q$. Then $\alpha$ is orthogonal to $S$ at $q$.

**Proof.** The proof is based upon a “variation of energy” technique, and is essentially taken from a discussion in [16].

Assume $\alpha : [0,a] \to M$ is a minimizing geodesic joining the points $p$ and $q$. Let $v \in T_qS$ and $\phi : (-\epsilon, \epsilon) \to S$ be a differentiable curve such that $\phi(0) = q$ and $\phi'(0) = v$.

Consider a variation $f : (-\epsilon, \epsilon) \times [0,a] \to M$, such that $f(0, \cdot) = \alpha$, $f(\cdot, 0) = p$ and $f(\cdot, a) = \phi$. Let

$$E(s) = \int_0^a \left| \frac{\partial f}{\partial t}(s,t) \right|^2 \, dt$$

be the energy associated with $f$. We are going to use the following inequality, relating the length of a curve $c : [0,a] \to M$ and its energy (see [3, p. 194]),

$$L^2(c) \leq a E(c).$$

Equality holds if and only if $c$ is parametrized proportional to arclength.

In particular,

$$L^2(f_s) \leq a E(f_s). \tag{17}$$

Note that since $\text{Image}(\phi) \subseteq S$, then for any $s \in (-\epsilon, \epsilon)$,

$$L(f_0) = L(\alpha) = d(p,q) = d(p,S) \leq d(p,\phi(s)) = d(f_s(0), f_s(a)) \leq L(f_s). \tag{18}$$

Finally, since $\alpha = f_0$ is a geodesic, its parametrization is proportional to arclength, so by the above comment,

$$L^2(f_0) = a E(f_0). \tag{19}$$

Combining (17),(18) and (19), we obtain

$$a E(f_0) = L^2(f_0) \leq L^2(f_s) \leq a E(f_s).$$

Thus $E(f_s)$ as a function of $s$ has a minimum at $s = 0$, therefore $E'(0) = 0$. By the first variation of energy formula (see [3, p. 195]),

$$\frac{d}{ds} \left| \frac{1}{2} \frac{dE^2(f_s)}{ds} \right|_{s=0} = (v, \alpha'(a)).$$

Since $v$ is an arbitrary vector in $T_qS$, the proof is complete. \qed

**Acknowledgments.** We are grateful to Patrizio Neff for clarifying the physical motivation behind the invariance assumptions and for pointing out relevant historic references. We are very much indebted to Amotz Oppenheim, Pavel Giterman and Amitay Yuval for useful discussions. We also thank Pietro Majer for useful comments on mathoverflow.

This research was partially supported by the Israel-US Binational Foundation (Grant No. 2010129), by the Israel Science Foundation (Grant No. 661/13), and by a grant from the Ministry of Science, Technology and Space, Israel and the Russian Foundation for Basic Research, the Russian Federation.
REFERENCES


Received March 2016; revised August 2016.

*E-mail address: raz@math.huji.ac.il*

*E-mail address: shacharasaf@gmail.com*