SIAM J. IMAGING SCIENCES
Vol. 7, No. 3, pp. 1503–1557

Left-Invariant Riemannian Geodesics on Spatial Transformation Groups*

Ernesto Zacur†, Matias Bossa†, and Salvador Olmos†

Abstract. Spatial transformations are mappings between locations of a d-dimensional space and are commonly used in computer vision and image analysis. Many of the spatial transformation sets have a group structure and can be represented by matrix groups. In the computer vision and image analysis fields there is a recent and growing interest in performing analyses on spatial transformations data. Differential and Riemannian geometry have been used as a framework to endow the set of spatial transformations with a metric space structure, allowing the extension of the standard analysis techniques defined on vector spaces. This paper presents a review of the concepts and an overview of approaches to computing Riemannian geodesics on spatial transformation groups. The paper is aimed at providing a bridge for researchers from computer vision and image analysis fields to fill in the gap between differential geometry and computer vision and imaging disciplines. Some application examples are shown to illustrate the use of invariant Riemannian geodesics, such as interpolation of spatial transformations and filtering of matrix-valued images.

Key words. left-invariant geodesics, Riemannian exponential, Lie groups, spatial transformation groups

AMS subject classifications. 58E10, 49K15, 53C22

DOI. 10.1137/130928352

1. Introduction.

Spatial transformations are bijective mappings between locations of a d-dimensional space, where d is usually 2 or 3. They play an important role in the fields of computer graphics, computer vision, and image analysis. For example, spatial transformations can be used for object alignment as well as to define an interpolation scheme that provides intermediate object positions [9, 157, 106, 132]. In shape analysis, spatial transformations are key ingredients for the definition of pose and shape [52]. Image registration and pose recovery are often defined as optimization problems on the space of a given spatial transformation model [22, 120, 140]. Spatial transformations are also broadly used in computer graphics for object modeling [123, 11, 59]. In the field of computational anatomy [69, 117, 165] the anatomical information is encoded in the spatial transformation that maps a given atlas to a target image.

Many of the spatial transformation sets have a group structure and are continuously parameterized with a finite number of parameters. Most of these sets can be represented by continuous matrix groups which are also Lie groups [164] and therefore differentiable manifolds. Then, a suitable distance on these manifolds can be defined from a Riemannian metric assigned to the manifold. Additionally, it is common that computer vision and image analysis

*Received by the editors July 9, 2013; accepted for publication (in revised form) April 25, 2014; published electronically July 24, 2014. This work was funded by research grants TEC2009-14587-C03-01 from CICYT, AMIT project CEN-20101014 from the CENIT program, and CIM project IPT-2011-1638-900000 from the INNPACTO program, Spain.

http://www.siam.org/journals/siims/7-3/92835.html

†GTC, I3A, Universidad de Zaragoza, Zaragoza 50018, Spain (zacur@unizar.es, bossa@unizar.es, olmos@unizar.es).

problems present a symmetry related to a set of spatial transformations. A wise selection of the metric should take into account these symmetries. Different problems can lead to symmetries under left-actions, right-actions, or from both sides. Left-invariant metrics are usually preferred for applications where measurements depend on an external coordinate system [136, 129], and right-invariant metrics are usually appropriate for applications where a reference system is used to normalize the results [178]. However, the definition of the symmetries is completely application-dependent and must be studied carefully.

Many applications have been proposed where the choice of a metric is crucial. Within the field of statistics on manifold data, applications include mean and variance estimation [67, 14, 133, 28, 119, 136], regression and curve fitting [55, 110, 83], dimensionality reduction [152, 160], hypothesis testing for group comparison [178], classification [34, 65], and statistical sampling [93, 112]. Interpolation of spatial transformations [21, 87, 106, 9, 132, 157, 35] and filtering of manifold data [146, 54, 56, 100, 71] also make use of a distance function. Intrinsic differential calculus on groups, which is essential in geometric optimization, is defined for a given metric [60, 4, 159, 3]. Moreover, the choice of a metric defines spatial transformation neighborhoods that are commonly used in global optimization techniques [105, 128, 78], as well as in derivative free optimization methods [44, 98, 72]. A distance between spatial transformations can also be used to evaluate the performance of image registration algorithms when a ground truth transformation is available [138, 147, 158]. For free and transitive group actions, distance between transformations can be used to define a distance between objects. For example, in robotics, rigid transformations were used to define distances between configurations [180]; in medical imaging, transformations between subcortical structures within the brain [34, 65] and spinal vertebrae [6, 28] were analyzed. Furthermore, a distance between deformations on the plane was used to measure distances between two-dimensional (2D) curves [176]. This distance between spatial deformations was used as a regularizer in ill-posed optimization problems [20]. Other applications are in the area of computer vision, such as object tracking [95, 166], visual servoing of robot manipulators [19, 51], and dynamical analysis of visual flows [108]. The particular cases of rotations and rigid transformations have been deeply analyzed within the fields of robotics [129, 130, 181, 179, 126, 124] and geometric mechanics [131, 84, 12].

When a Riemannian metric is assigned to the manifold of spatial transformations, geodesics are locally length-minimizing paths on the manifold. Geodesics play an important role in many applications of computer vision and image analysis. In this framework, the length of the shortest curve connecting two transformations defines the Riemannian distance and endows a metric structure to the set of spatial transformations [50, 64]. The shortest curve segment between two given spatial transformations belongs to the set of geodesics.

Riemannian geodesics can be expressed via the Riemannian exponential function in terms of a departing element and an initial velocity. The Riemannian exponential function maps vectors from a tangent space to the group. On the other hand, the group exponential function also maps vectors from the tangent space at the identity to the group. It is well known that for matrix groups this mapping is given by the matrix exponential. Although in a few particular groups the Riemannian exponential function for some invariant metrics and translated group exponentials provides the same curve [118, 50, 62], the Riemannian exponential and the group exponential functions are totally different concepts. Unluckily, it should be noted that these concepts have not always been used with sufficient accuracy in some publications. In particular, group exponential curves were interpreted as shortest curves in [162, 19, 167, 66, 108, 107, 146, 141, 166, 32, 95, 94]. In general, several local coordinate systems can be used to describe neighborhoods of spatial transformations with increasing levels of mathematical structure (and usually also with increasing computational complexity): charts, which make use of the topological structure of the spatial transformations; geodesics, which consider an affine connection on the group, and retractions, which provide local approximation of geodesics; translated group exponentials, which take into account the group structure of the spatial transformations; and invariant Riemannian geodesics, which make use of the symmetries of the group together with a metric structure on it. Also, the concept of geodesics is often found in computer graphics and image analysis disciplines for other applications, such as implicit surfaces [49], triangle meshes [163, 17, 29, 139], and images [36, 41, 46, 156], but those will not be considered in this work.

The aim of this paper is to provide an overview of the approaches that describe geodesic evolutions on spatial transformation groups and to bridge the gap between disciplines—on one hand, differential geometry, Lie group theory, and geometric mechanics, and on the other hand, computer vision, computer graphics, and image analysis. This paper will help the reader understand these concepts. It provides several resolution methods for computing geodesics and details of many spatial transformation groups and illustrates the application of geodesics with several examples from computer vision and image analysis.

1.1. Outline.

Section 1.2 presents a definition of transformation models with a group structure. A description of the algebraic operations and the action over geometric objects is also given together with an enumeration of the spatial transformations within the scope of the paper.

A summary of several concepts from Lie group and Riemannian geometry theories is presented in section 2 with the aim of providing a self-contained paper written for computer graphic and imaging researchers. At the same time this section introduces the notation used throughout the paper. Some footnotes are included to generalize for groups that are not matrix groups, such as quotient groups or infinite dimensional groups.

Section 3 gathers several approaches to computing left-invariant Riemannian geodesics on matrix groups. The left-invariance case was arbitrarily chosen, but it is straightforward to obtain right-invariant geodesics, as will be indicated in section 2.7. Whenever possible, the resolution methods were written using matrix operations, allowing straightforward numerical implementations. The methods include the integration of ordinary differential equations (ODEs) described in different ways: on a specific chart coordinate in section 3.1, on a chart coordinate valid for any matrix group in section 3.2, in solving the Euler–Poincaré equation in section 3.3, in a coordinate-free formulation of Euler–Poincaré equation in section 3.4, in a coupled state-costate equation system derived from optimal control in section 3.5, and in a symmetric representation of geodesics in section 3.6.

Section 4 presents the derivation of two well-known conserved quantities along left-invariant geodesics, using results from section 3.5. These quantities allow formulation in section 5 of the geodesic evolution as an equivalent initial value problem (IVP) expressed only in terms of the transformation elements. Finally, the derivatives of the Riemannian exponential function with respect to the initial velocity are described in section 6. These are known as *Jacobi fields* and are used in practical applications to compute the inverse of the Riemannian exponential function. Section 7 shows how to find geodesics in a broad family of Riemannian metrics using a known geodesic defined for a particular metric.

To provide a detailed presentation of invariant metrics on spatial transformation groups commonly involved in computer vision and image analysis applications, section 8 includes various closed-form solutions of the Riemannian exponential function.

As geodesics on spatial transformation groups play a key role in many applications, a set of illustrative examples is given in section 9. This section starts presenting simple numerical schemes for computation of the Riemannian exponential function and its inverse. These numerical schemes are given only as basic and illustrative examples because this paper is not aimed at exploring the vast variety of numerical integrators. Finally, three application examples are shown: invariant interpolation of spatial transformations; Gaussian filtering of a field of spatial transformations; and an extension of the closed-form solution of left-invariant geodesics for the $\mathcal{ST}(1)$ group under any invariant metric.

1.2. Spatial transformation models. In computer graphics, computer vision, and image analysis, *geometrical objects* are typically defined by subsets of points of an ambient space, such as, for example, the Euclidean space $\mathbb{R}^d$. The dimension d of the ambient space where objects lie is usually 2 or 3.

Spatial transformations are one-to-one mappings of the whole ambient space onto itself. The action of spatial transformations on geometrical objects typically consists in applying the mapping to each constituent point [153, 77].

Let O be the set of geometrical objects, and let $\mathcal{G}$ be a nonempty set of transformations; the *action* of $\chi \in \mathcal{G}$ on $\mathsf{A} \in \mathsf{O}$ will be denoted by $\chi \cdot \mathsf{A} = \mathsf{B} \in \mathsf{O}$. The *identity transformation*, e, fulfills $e \cdot \mathsf{A} = \mathsf{A}$ for all A. The *inverse transformation* always exists because χ is defined as a bijection. The corresponding inverse is denoted by χ^{-1} and satisfies $\chi^{-1} \cdot \mathsf{B} = \mathsf{A} \Leftrightarrow \chi \cdot \mathsf{A} = \mathsf{B}$ for all A and $\mathsf{B} \in \mathsf{O}$.

A binary operation called *composition* on the set $\mathcal{G}$, denoted by $(\cdot\bullet\cdot)$, is given by $(y\bullet\chi)\cdot\mathsf{A} = y \cdot (\chi \cdot \mathsf{A})$ for all objects A. When $(y \bullet \chi) = z$ belongs to the set $\mathcal{G}$ for all $\chi, y \in \mathcal{G}$ and the inverse elements χ^{-1} are also members of $\mathcal{G}$, then $(\mathcal{G}, \bullet)$ defines a *group structure*. It is standard notation to omit the $\bullet$ symbol and use the juxtaposition of group elements as shorthand for the composition operation $\chi y \equiv \chi \bullet y$. The *identity transformation*, e, belongs to $\mathcal{G}$ and satisfies $\chi e = e\chi = \chi$ for all $\chi \in \mathcal{G}$, the *identity element* of the group $\mathcal{G}$. Moreover, the *inverse group operation*, denoted by $\mathbf{inv}(\cdot)$, is automatically assigned to $\mathbf{inv}(\chi) = \chi^{-1}$, satisfying $\mathbf{inv}(\chi)\chi = \chi\mathbf{inv}(\chi) = e$.

The action of spatial transformations preserves some geometrical properties of the objects [136, 23, 123]. A set of transformations can be categorized as a *transformation model* according to the geometrical properties that the transformations preserve. For example, rigid motion transformations maintain invariant Euclidean distances between any pair of points; orientation preserving transformations preserve the handedness of a frame, i.e., they do not generate inversions of the space; affine transformations preserve parallelism; conformal transformations keep local angles unchanged.

Examples of finite dimensional transformation models which are commonly used in computer graphics and imaging applications are given below (section 8 gives more details on each model):

1. *Translation group*, $\mathcal{T}(d)$
2. *Isotropic Scale group*, $\mathcal{S}^+(1)$
3. *Non-Isotropic Scale group*, $\mathcal{S}^+(d)$
4. *Isotropic Scale + Translation group*, $\mathcal{ST}(d)$
5. *Rotation group*, $\mathcal{SO}(d)$
6. *Special Euclidean group*, $\mathcal{SE}(d)$
7. *Similarity group*, $\mathcal{SIM}(d)$
8. *Orientation Preserving General Linear group*, $\mathcal{GL}^+(d)$
9. *Special Linear group*, $\mathcal{SL}(d)$
10. *Centered Transformation group*, $(\mathcal{G}(d) \times \mathcal{T}(d))$
11. *Projective group*, $\mathcal{PGL}(d)$
12. *Möbius Transformation group*, $\mathcal{M}(d)$
13. *Orientation Preserving General Affine group*, $\mathcal{GA}^+(d)$
14. *Special Affine group*, $\mathcal{SA}(d)$

There are other transformation models in the literature that are not considered in this work for different reasons:

- Transformations that do not ensure the invertibility of the spatial mappings such as *polynomial deformations* [174], *free-form deformations* [155, 99, 96], and *sparse radial basis deformations* (such as thin-plate splines (TPS) [30, 148] or Gaussian elastic body splines (GEBS) [90]).
- Transformations not closed under composition, and therefore without a group structure such as the *weak affine* [103, 58] or *nine-parameter affine model* [18, Chapter 29], *perspectivities* [77, Appendix A7.4], and *polyaffine transformations* [15].
- Transformations that are sets of isolated elements without a continuous parametrization such as *Inversion groups* and *Dihedral groups*.
- Transformation models which are infinite dimensional groups and cannot be identified, not even locally, with a matrix group such as 2*D Conformal Deformation groups* and *Diffeomorphism groups* [177].

2. Background.

This section gathers some concepts and properties involved in the definition of geodesics on transformation groups. In particular, Riemannian metrics and the Riemannian exponential function are briefly explained. Also, the concept of group exponential is revisited in order to avoid its misleading use. Knowledge of these concepts and properties is essential for dealing with the computation of geodesics and distances. Further details can be found in classical textbooks [50, 64, 74, 164, 37, 4].

2.1. Matrix groups, Lie groups.

Typically, the points of the ambient space can be mapped to $\mathbb{R}^n$ for some $n \geqslant d$, where the spatial transformation model acts linearly. This mapping induces a *matrix representation* of the spatial transformations in $\mathbb{M}_n$, the set of all real $n \times n$ matrices. A matrix representation is an identification of a group element with a finite dimensional linear transform. This matrix representation enables the use of matrix algebra operations providing a way to make numerical computations. Using this representation the group composition operation results in the matrix-matrix product, the identity element is identified with the identity matrix, $e \equiv I$, and the group inverse operation $\mathbf{inv}(\cdot)$ is computed as the matrix inversion. Examples of mappings from the d-dimensional ambient space $\mathbb{R}^n$ to $\mathbb{R}^n$ and matrix representations of transformation groups are as follows:

- For the Rotation group, $\mathcal{SO}(d)$, a spatial transformation can be represented as a matrix in $\mathbf{SO}(d)$ (the group of $d\times d$ orthogonal matrices with positive determinant) which acts linearly on $\mathbb{R}^d$.
- For the Orientation Preserving General Affine group, $\mathcal{GA}^+(d)$, spatial points can be mapped to vectors in $\mathbb{R}^{d+1}$ of the form $\left(c\mathsf{y}^T, c\right)^T$ for any $c \neq 0$ and $\mathsf{y} \in \mathbb{R}^d$. A matrix representation of an affine transformation is a $(d+1)\times(d+1)$ matrix of the form

$$\left(\begin{array}{c|c} X & \bar{t} \\ \hline 0 & 1 \end{array}\right), \tag{2.1}$$

where $X \in \mathbf{GL}^+(d)$, the set of $d\times d$ matrices with positive determinant, and $\bar{t} \in \mathbb{R}^d$. The matrices of the form (2.1) form the matrix group $\mathbf{GA}^+(d)$.
- For the Möbius Transformation group, $\mathcal{M}(d)$, d-dimensional spatial points can be mapped to $\mathbb{R}^{d+2}$, and a matrix representation of $\mathcal{M}(d)$ is given by $(d+2)\times(d+2)$ matrices. More details will be given in section 8.12.

Almost all classical spatial transformation groups of finite dimension allow a matrix representation ($\mathcal{G} \mapsto \mathbf{G}\colon \chi \mapsto X \in \mathbb{M}_n$ for some n). For simplicity only matrix groups are considered in this work, although some footnotes generalize the concepts to abstract structures.

As shown in [164], all matrix groups are *Lie groups*. This means that group elements can be identified with elements of a *smooth*[1] *manifold*, where the composition and inversion are differentiable operations.

2.2. Charts and manifolds. The concept of a manifold relies on the concepts of charts and atlases. Let $\mathcal{G}$ be a nonempty set of spatial transformations, and let χ be an element of $\mathcal{G}$. A *local coordinate chart* $(\mathcal{D}, \varphi)$ is a bijection that maps elements in $\mathcal{D} \subset \mathcal{G}$ to an open set $\Omega \subset \mathbb{R}^k$ for some k. The array $\varphi(\chi) = \bar{x} = \left(x^1, x^2, \ldots, x^k\right)^T \in \Omega$ (hereinafter the overline means that the corresponding quantity is a column array while superscripts refer to its components) contains the coordinates of χ in the chart $(\mathcal{D}, \varphi)$. An *atlas* for $\mathcal{G}$ is a countable collection of charts $\{(\mathcal{D}_i, \varphi_i)\}$ such that the union of all $\mathcal{D}_i$ covers the set $\mathcal{G}$. This construction allows definition of a *topology* on the set $\mathcal{G}$ since open sets in $\mathbb{R}^k$ (or, more precisely, in each Ω_i) define *open sets in* $\mathcal{G}$ via the functions φ_i^{-1}. If for any pair of charts $(\mathcal{D}_i, \varphi_i)$ and $(\mathcal{D}_j, \varphi_j)$ with a nonempty intersection $\mathcal{D}_i \cap \mathcal{D}_j$, the function $\varphi_i \circ \varphi_j^{-1}\colon \mathbb{R}^k \to \mathbb{R}^k$ is smooth, then the set $\mathcal{G}$ becomes a *smooth manifold* with *intrinsic dimension* k where continuity and differentiability properties in $\mathcal{G}$ are inherited from the continuity and differentiability of $\mathbb{R}^k$. Let $\mathcal{G}$ be a smooth manifold with the atlas $\{(\mathcal{D}_i, \varphi_i)\}$, and let $\mathbb{V}$ be a vector space; a function $\mathbf{f}\colon \mathcal{G} \to \mathbb{V}$ is smooth if all the compositions $\mathbf{f} \circ \varphi_i^{-1}\colon \Omega_i \subset \mathbb{R}^k \to \mathbb{V}$ are smooth functions. Similarly, a function $\mathbf{g}\colon \mathbb{V} \to \mathcal{G}$ is smooth if all $\varphi_i \circ \mathbf{g}\colon \mathbb{V} \to \Omega_i$ are smooth.

In the case of a matrix group $\mathbf{G} \subset \mathbb{M}_n$, as the set of matrices $\mathbb{M}_n$ is isomorphic to $\mathbb{R}^{n^2}$, the manifold $\mathbf{G}$ can be thought of as a smooth "surface" of dimension k embedded in $\mathbb{R}^{n^2}$.

As we are interested in building continuous paths between elements of a group, we will avoid transformation models of isolated elements such as Inversion or Dihedral groups and

[1]In this work *smooth* refers to functions with derivatives up to any order, i.e., $\mathcal{C}^\infty$ functions.

only cases where $k \geqslant 1$ will be considered. Only elements belonging to the same *connected component* can be connected by continuous paths.

2.3. Tangent vectors, tangent space, and vector fields. A *curve* in a matrix group $\mathbf{G}$ is a smooth function $\gamma\colon \mathbb{R} \to \mathbf{G}\colon t \mapsto \gamma_{(t)}$. The *velocity*, alternatively called *tangent vector*, of the curve is defined as the usual limit $\lim_{\epsilon\to 0}(\gamma_{(t+\epsilon)} - \gamma_{(t)})/\epsilon$ resulting in a matrix[2] $V_{(t)} = \mathrm{d}_s\,\gamma_{(s)}\big|_{(s=t)} = \dot{\gamma}_{(t)}$. Like group elements, the velocities $\dot{\gamma}_{(t)}$ can also be represented by matrices in $\mathbb{M}_n$, but in general they do not belong to the set $\mathbf{G}$. The velocities of all curves passing through the point X at time $t = 0$ define the *tangent space* of $\mathbf{G}$ at X, and it is denoted by $T_X\mathbf{G} = \{\dot{\gamma}_{(0)}|$ such that γ is a curve in $\mathbf{G}$ with $\gamma_{(0)} = X\}$. Every tangent space $T_X\mathbf{G}$ is a vector space of dimension k.

A *vector field* $\mathcal{H}$ on the group is an assignment from each $X \in \mathbf{G}$ to a velocity in $T_X\mathbf{G}$. $\mathcal{H}$ acts on a smooth scalar function $\mathbf{f}\colon \mathbf{G} \to \mathbb{R}$ providing the directional derivative $(\mathcal{H}\mathbf{f})(X) = \mathcal{H}(X)\mathbf{f} = \lim_{\epsilon\to 0}(\mathbf{f}(X + \epsilon\mathcal{H}(X)) - \mathbf{f}(X))/\epsilon$ (whenever the limit can be evaluated; otherwise the limit can be evaluated in chart coordinates). A vector field $\mathcal{H}$ is smooth if for any smooth scalar function $\mathbf{f}$, the scalar function $(\mathcal{H}\mathbf{f})\colon \mathbf{G} \to \mathbb{R}$ is smooth.

2.4. Translating elements, curves, and tangent vectors. For every group element Y, *left-translation* is the map $\mathbf{L}_Y\colon \mathbf{G} \to \mathbf{G}$ defined by $\mathbf{L}_Y(X) = YX$ for all $X \in \mathbf{G}$. Analogously, *right-translation* is defined by $\mathbf{R}_Y(X) = XY$.

Let γ be a curve through X with velocity V at $t = 0$; then the curve ξ defined by $\xi_{(t)} = \mathbf{L}_Y(\gamma_{(t)})$ passes through YX at $t = 0$ ($\xi_{(0)} = YX$). Its velocity $W = \dot{\xi}_{(0)}$ belongs to $T_{YX}\mathbf{G}$. The corresponding action on tangent vectors of $\mathbf{L}_Y$ is a linear operation denoted by $T_X\mathbf{L}_Y$ given by[3] $W = \mathrm{d}_s(Y\gamma_{(s)})\big|_{(s=0)} = T_X\mathbf{L}_Y\,V = YV$.

As velocities can be translated along the group, it is very useful and convenient to define the properties of velocities at a single point. Algebraically, the most convenient point is the identity I. The tangent space at the identity $T_I\mathbf{G}$ is called *group algebra* and is denoted by $\mathfrak{g}$. Any velocity V in any tangent space $T_X\mathbf{G}$ can be associated with an element $U \in \mathfrak{g}$ via $V = XU$.

The algebra $\mathfrak{g}$ is a k-dimensional vector space; therefore, a set of bases $\{B_1, B_2, \ldots, B_k\}$, with $B_i \in \mathbb{M}_n$, can be selected such that $\mathfrak{g} \equiv \operatorname{span}(\{B_i\})$. In some disciplines these bases are referred to as *infinitesimal group generators*, and they span all group perturbations that can be done at the identity. Matrices representing vectors of $\mathfrak{g}$ are written as $U = \sum_i^k B_i\mu^i \in \mathfrak{g}$, where $\bar{\mu} = \left(\mu^1, \mu^2, \ldots, \mu^k\right)^T \in \mathbb{R}^k$ are called the *algebra coordinates of* U. Let $\overline{U}$ denote

[2]On an abstract manifold $\mathcal{G}$ without a vector space structure, the usual limit can only be evaluated in terms of coordinates. Alternatively tangent vectors can be defined as differential operators. Let $\mathbf{f}$ be a smooth real valued function defined on $\mathcal{G}$; then $\mathbf{f} \circ \boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}\colon t \mapsto \mathbf{f}(\boldsymbol{\gamma}_{(t)})$ and the velocity $\dot{\boldsymbol{\gamma}}_{(t)}$ is the linear operator satisfying (see, for example, [5, 62])

$$\dot{\boldsymbol{\gamma}}_{(t)}\mathbf{f} := \mathrm{d}_s\,\mathbf{f}(\boldsymbol{\gamma}_{(s)})\big|_{(s=t)}\,.$$

[3]For abstract structures, $\mathcal{V} \in T_\chi\mathcal{G}$ and $\mathcal{W} \in T_{y\chi}\mathcal{G}$ are related by a linear operator called *tangent-lift*, denoted by $T_\chi\mathbf{L}_y\colon T_\chi\mathcal{G} \to T_{y\chi}\mathcal{G}$ or, in a more general form, for a function $\mathbf{F}\colon \mathcal{G} \to \mathcal{G}$, $T_\chi\mathbf{F}\colon T_\chi\mathcal{G} \to T_{\mathbf{F}(\chi)}\mathcal{G}$ (see, for example, [84]). There are different notations for this operator. For example, [62] denotes it as $(\mathrm{d}\mathbf{F})_\chi$; [12] and [13] use $\mathbf{F}_*$ or $\mathbf{F}_*|_\chi$; [14] uses $D_\chi\mathbf{F}$; [43, 91] simply denote $\mathbf{F}$, leaving to the context its meaning; $\mathrm{D}\mathbf{F}(\chi)$ was used in [4].

the n^2-dimensional array obtained by stacking the columns of the matrix U. It is practical to write $\overline{U} = \boldsymbol{B}\bar{\mu}$, where $\boldsymbol{B}$ is an $n^2 \times k$ matrix built from the vectorization of each basis, $\boldsymbol{B} = \left(\overline{B}_1, \overline{B}_2, \ldots, \overline{B}_k\right)$. Moreover, as $\boldsymbol{B}$ is full column rank, the algebra coordinates can be obtained by $\bar{\mu} = \boldsymbol{B}^{\dagger}\overline{U}$, where $(\cdot)^{\dagger}$ denotes a left generalized inverse, i.e., any matrix satisfying $\boldsymbol{B}^{\dagger}\boldsymbol{B} = I_k$.

The group algebra $\mathfrak{g}$ has a bilinear operation called the *Lie bracket*, $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. By representing the elements of the algebra as matrices, the Lie bracket operation results in the matrix commutation, $[U, W] = UW - WU$ for U and $W \in \mathfrak{g}$.

Left-translations of velocities from the group algebra to the whole group allow the generation of *left-invariant vector fields*[4] $\widetilde{\mathcal{U}}(X) = XU$ for any $U \in \mathfrak{g}$ and any $X \in \mathbf{G}$. It is remarkable that any smooth vector field (not necessarily invariant) can be written locally as a linear combination of left-invariant vector fields, $\mathcal{H} = \sum_i \mathbf{h}_i \widetilde{\mathcal{U}}_i$, with $\mathbf{h}_i$ smooth scalar functions on $\mathbf{G}$.

Another important operation defined on the group algebra is the *Adjoint action* of $\mathbf{G}$ on $\mathfrak{g}$, given by[5] $\mathbf{Ad}_Y\colon \mathfrak{g} \to \mathfrak{g}\colon U \mapsto YUY^{-1}$ for any $Y \in \mathbf{G}$. For a curve γ passing through I with velocity U at $t = 0$, the *conjugated* curve $\zeta_{(t)} = Y\gamma_{(t)}Y^{-1}$ passes through $\zeta_{(0)} = I$ with velocity $\dot{\zeta}_{(0)} = \mathbf{Ad}_Y U$.

Now, let η be a curve in $\mathbf{G}$ such that $\eta_{(0)} = I$ and $\dot{\eta}_{(0)} = W \in \mathfrak{g}$. For a given $U \in \mathfrak{g}$, the operation given by $\mathrm{d}_s(\mathbf{Ad}_{\eta_{(s)}}U)\Big|_{(s=0)}$ is called the *adjoint operator* of W on $\mathfrak{g}$. It is denoted by $\mathbf{ad}_W\colon \mathfrak{g} \to \mathfrak{g}$ and is a linear operation on $\mathfrak{g}$ for any $W \in \mathfrak{g}$. Moreover, it can be proved that the adjoint operator coincides with the Lie bracket on the algebra vectors, i.e., $\mathbf{ad}_W(U) = [W, U]$ for any $U, W \in \mathfrak{g}$.

2.5. Group exponential. Given a smooth vector field $\mathcal{H}$, a curve γ is called an *integral curve* of $\mathcal{H}$ if $\dot{\gamma}_{(t)} = \mathcal{H}(\gamma_{(t)})$. For any smooth vector field $\mathcal{H}$ and any $X \in \mathbf{G}$ there is a unique integral curve γ of $\mathcal{H}$ such that $\gamma_{(0)} = X$.

The *group exponential* is a function $\exp\colon \mathfrak{g} \to \mathbf{H} \subseteq \mathbf{G}\colon U \mapsto X = \exp(U)$ defined by $X = \gamma_{(t=1)}$, where γ is the integral curve of the left-invariant vector field $\widetilde{\mathcal{U}}$ starting at the identity. The group exponential is a *one-parameter subgroup* of $\mathbf{G}$, i.e., $\exp(\alpha U)\exp(\upsilon U) = \exp((\alpha + \upsilon)U)$ for any $U \in \mathfrak{g}$. In particular, $\exp(0) = I$ and $\exp(-U) = (\exp(U))^{-1}$.

For matrix groups, integral curves are given by the solution of the following autonomous ODEs:

$$\begin{cases} \dot{X}_{(t)} = \widetilde{\mathcal{U}}(X_{(t)}) = X_{(t)}U, \\ X_{(t=0)} = I. \end{cases} \tag{2.2}$$

The solution of (2.2) up to time $t = 1$ is given by the *matrix exponential* function [7, 122, 81] defined as $X_{(1)} = \exp(U) = e^U = \sum_m U^m/m!$.

The inverse of the group exponential function is given by the *matrix logarithm* [81], satisfying $\exp(\log(X)) = X$ for any $X \in \mathbf{H} \subseteq \mathbf{G}$. For some groups, such as the Rotation

[4] Equivalently, a left-invariant vector field is a vector field which satisfies $\widetilde{\mathcal{H}}(y\chi) = T_\chi \mathbf{L}_y \widetilde{\mathcal{H}}(\chi)$ for any $\chi, y \in \mathcal{G}$. Alternatively, the group algebra can be defined as the set of all left-invariant vector fields, and the Lie bracket is $[\widetilde{\mathcal{U}}, \widetilde{\mathcal{W}}] = \widetilde{\mathcal{U}} \circ \widetilde{\mathcal{W}} - \widetilde{\mathcal{W}} \circ \widetilde{\mathcal{U}}$ for any left-invariant vector fields $\widetilde{\mathcal{U}}$ and $\widetilde{\mathcal{W}}$.

[5] In abstract notation, $\mathbf{Ad}_y = (T_{y^{-1}}\mathbf{L}_y \circ T_e \mathbf{R}_{y^{-1}})\colon \mathfrak{g} \to \mathfrak{g}$.

group, the matrix logarithm function is multivalued, and in general it cannot be ensured that $\log(\exp(U)) = U$.

Remarks. For finite dimensional groups, the group exponential is a diffeomorphism from an open neighborhood of the algebra to an open neighborhood of the group at the identity. Even though the group exponential is defined for all algebra vectors, not all group elements are the group exponential of an algebra vector[6] (the conditions for a real matrix to have a real logarithm can be found in [61]).

2.6. Riemannian metrics, length of curves, Riemannian geodesics, and Riemannian exponential. Let $\mathbb{V}$ be a vector space; an *inner product* is a bilinear, symmetric, and positive definite form denoted by $\langle\!\langle \cdot,\cdot \rangle\!\rangle\colon \mathbb{V}\times\mathbb{V}\to\mathbb{R}$. The inner product is uniquely identified with a linear map to the dual space $\boldsymbol{\mathcal{M}}\colon \mathbb{V}\to\mathbb{V}^*$ which is called the *inertia operator* (a.k.a. *inertia tensor*). Let us denote the pairing between $\mathbb{V}$ and its dual $\mathbb{V}^*$ by $\langle\cdot,\cdot\rangle : \mathbb{V}^*\times\mathbb{V}\to\mathbb{R}\colon \langle \mathfrak{w}, v\rangle = \mathfrak{w}(v)$, where $\mathfrak{w}\in\mathbb{V}^*$ is a *linear form*, also called *covectors*. Then, the inner product and the inertia operator satisfy $\langle\!\langle w,v\rangle\!\rangle = \langle \boldsymbol{\mathcal{M}}(w), v\rangle$ (see [91, 12]).

A *Riemannian metric* on a manifold $\mathbf{G}$ is a smooth assignment of inner products to the tangent space at every element of $\mathbf{G}$, or, equivalently, each $X\in\mathbf{G}$ is associated with an inertia operator $\boldsymbol{\mathcal{M}}_X$ such that $\langle \boldsymbol{\mathcal{M}}_X(W), V\rangle = \langle\!\langle W,V\rangle\!\rangle_X$ is the inner product on $T_X\mathbf{G}$. With the assignment $(\mathbf{G}, \langle\!\langle\cdot,\cdot\rangle\!\rangle.)$ the *length* of a curve segment $\gamma\colon [t_0,t_1]\subset\mathbb{R}\to\mathbf{G}$ is defined by

$$\mathbf{Length}(\gamma; t_0, t_1) = \int_{t_0}^{t_1} \langle\!\langle \dot\gamma_{(s)}, \dot\gamma_{(s)}\rangle\!\rangle_{\gamma_{(s)}}^{1/2}\,\mathrm{d}s\,. \tag{2.3}$$

The *Riemannian distance* between two elements A and $B\in\mathbf{G}$ that can be connected by a curve is

$$\mathbf{distance}(A,B) = \begin{cases} \underset{\gamma,\text{ a curve in }\mathbf{G}}{\text{infimum}} & \mathbf{Length}(\gamma;t_0,t_1),\\ \text{subject to} & \gamma_{(t_0)} = A \quad ; \quad \gamma_{(t_1)} = B. \end{cases} \tag{2.4}$$

Endowed with the function $\mathbf{distance}(\cdot,\cdot)$, the group $\mathbf{G}$ acquires a *metric space* structure with the same topology as the original topology on $\mathbf{G}$ given by the atlas $\{(\mathcal{D}_i,\varphi_i)\}$ as in section 2.2.

The length of a curve segment is independent of its parametrization, which means that $\mathbf{Length}(\gamma;t_0,t_1) = \mathbf{Length}(\gamma\circ h;t_0,t_1)$ for any smooth and invertible time warping $h\colon [t_0,t_1]\to[t_0,t_1]$ with $h_{(t_0)} = t_0$ and $h_{(t_1)} = t_1$. If the *arc-length* parametrization is used, $\left\|\dot\gamma_{(s)}\right\|_{\gamma_{(s)}} = \langle\!\langle \dot\gamma_{(s)},\dot\gamma_{(s)}\rangle\!\rangle_{\gamma_{(s)}}^{1/2} = \text{constant}$, then the length of the segment is equal to $(t_1-t_0)\left\|\dot\gamma_{(0)}\right\|_{\gamma_{(0)}}$.

In order to avoid multiple solutions of minimal curve length (2.3), the following *Riemannian energy* (also called *action* in the area of geometric mechanics) is used:

$$\mathbf{E}(\gamma;t_0,t_1) = \frac{1}{2}\int_{t_0}^{t_1}\langle\!\langle \dot\gamma_{(s)},\dot\gamma_{(s)}\rangle\!\rangle_{\gamma_{(s)}}\mathrm{d}s = \frac{1}{2}\int_{t_0}^{t_1}\left\|\dot\gamma_{(s)}\right\|_{\gamma_{(s)}}^2\mathrm{d}s\,, \tag{2.5}$$

[6]For example, it can be checked that the 2D spatial transformation obtained by the composition of a rotation of $3/4\pi$ around the origin followed by a non-isotropic scale of 6 along the horizontal axis cannot be connected to the identity with a one-parameter subgroup.

where the factor 1/2 is included for later algebraic convenience. It is easy to observe that $\mathbf{E}(\gamma)$ does depend on the time parametrization of γ, and this value is minimal when $\left\|\dot{\gamma}_{(s)}\right\|_{\gamma_{(s)}}$ is constant. Therefore,

$$\left.\begin{array}{rl} \underset{\gamma}{\text{infimum}} & \mathbf{E}(\gamma; t_0, t_1) \\ \text{subject to} & \gamma_{(t_0)} = A \quad ; \quad \gamma_{(t_1)} = B \end{array}\right\} = \frac{\mathbf{distance}(A, B)^2}{2(t_1 - t_0)} . \tag{2.6}$$

Finally, solving (2.6) is equivalent to minimizing **Length** (equation (2.4)) with the advantage that the solution is an arc-length parameterized curve.

The extremal curves of (2.5) are called *Riemannian geodesics*. In general, given two elements A, B belonging to the same connected component of $\mathbf{G}$, there may exist several geodesic segments γ satisfying $\gamma_{(0)} = A$ and $\gamma_{(1)} = B$, even having the same length or energy. The minimal length among all geodesic segments is the distance between the elements. Furthermore, given a geodesic segment $\gamma_{(t)}$ passing through A and B, an affine reparametrization $\gamma_{(\alpha t+\tau)}$ is also a geodesic segment connecting A and B. The lengths and energies of all affine reparametrized segments from A to B are the same. As a general rule, when considering geodesics between two points, we will restrict our attention to those which at $t = 0$ pass through the first point and at $t = 1$ pass through the second one.

Given a Riemannian metric $(\mathbf{G}, \langle\!\langle \cdot, \cdot \rangle\!\rangle_{\cdot})$, for every $X \in \mathbf{G}$ and every $V \in T_X\mathbf{G}$, there exist an open interval $(-\varepsilon, \varrho) \subset \mathbb{R}$ around 0 and a unique geodesic $\gamma(t; X, V)$ for t in the interval, such that $\gamma_{(t=0)} = X$ is its *starting point* and $\dot{\gamma}_{(t=0)} = V$ is its *initial velocity*. This construction of a geodesic satisfies the homogeneity property $\gamma(t; X, V) = \gamma(\alpha t; X, V/\alpha)$ for $\alpha \in \mathbb{R}^+$. With the previous construction, if $t = 1$ is included in the interval $(-\varepsilon, \varrho)$, the *Riemannian exponential* function is defined as the element of the resulting geodesic at $t = 1$:

$$\mathrm{Exp}_X \colon T_X\mathbf{G} \to \mathbf{G} \colon V \mapsto \mathrm{Exp}_X(V) = \gamma(1; X, V) .$$

The Riemannian exponential function provides geodesics in terms of a starting point and an initial velocity instead of a starting element and an ending element: $\gamma(t; X, V) = \mathrm{Exp}_X(tV)$ is the unique geodesic such that $\gamma_{(t=0)} = X$ and $\dot{\gamma}_{(t=0)} = V$. As it is arc-length parametrized, its length under the given metric is $\mathbf{Length}(\mathrm{Exp}_X(tV); 0, 1) = \langle\!\langle V, V \rangle\!\rangle_X^{1/2} = \|V\|_X$.

The set $\mathfrak{c}_X = \{\, V \in T_X\mathbf{G} \text{ such that } \|V\|_X = \mathbf{distance}(X, \mathrm{Exp}_X(V)) \,\}$ is star-shaped, and its boundary is called the *tangential cut locus*. The $\mathrm{Exp}_X(\cdot)$ function is a diffeomorphism from any open 0-neighborhood $\mathfrak{s}_X \subseteq \mathfrak{c}_X$ onto an open X-neighborhood $\mathbf{H}_X \subseteq \mathbf{G}$. Therefore, the inverse function $\mathrm{Exp}_X^{-1}(\cdot)$ is well defined on $\mathbf{H}_X$. In computer vision, $\mathrm{Exp}_X^{-1}(\cdot)$ is usually called the *Riemannian logarithm function* [133, 56, 57, 162, 160] and is denoted by $\mathrm{Log}_X(\cdot)$.

Let $A, B \in \mathbf{G}$ such that $B \in \mathbf{H}_A$; then a geodesic connecting both elements is given by $\gamma_{(t)} = \mathrm{Exp}_A(t\,\mathrm{Log}_A(B))$, being that $\gamma_{(0)} = A$ and $\gamma_{(1)} = B$. The corresponding length of the curve segment from $t = 0$ to $t = 1$ is $\mathbf{Length}(\gamma; 0, 1) = \|\mathrm{Log}_A(B)\|_A$. Moreover, this geodesic segment is the shortest one, and therefore its length is the Riemannian distance between A and B. Note that, while $\mathrm{Exp}(\cdot)$ identifies geodesics in terms of starting points and velocities, $\mathrm{Log}(\cdot)$ allows the identification of geodesics in terms of their initial and ending points.

2.7. Invariant Riemannian metrics. In the following, the group structure and group operations will be considered on the manifold $\mathbf{G}$ leading to metrics that preserve the group symmetries. Two families of Riemannian metrics on Lie groups arise naturally: *left-invariant* and *right-invariant* metrics. In this work, without loss of generality, left-invariant metrics are analyzed, but right-invariant ones can be likewise considered or eventually solved by means of a change of parametrization (see (2.11) and (2.12)).

A left-invariant metric is an isometry under left-translations; i.e., the inner products fulfill

$$\langle\!\langle V_1, V_2 \rangle\!\rangle_X = \langle\!\langle T_X \mathbf{L}_Y V_1 \, , \, T_X \mathbf{L}_Y V_2 \rangle\!\rangle_{YX} = \langle\!\langle YV_1 \, , \, YV_2 \rangle\!\rangle_{YX} \tag{2.7}$$

for any $X, Y \in \mathbf{G}$ and any $V_1, V_2 \in T_X\mathbf{G}$. In particular,

$$\langle\!\langle V_1, V_2 \rangle\!\rangle_X = \left\langle\!\!\left\langle X^{-1}V_1, X^{-1}V_2 \right\rangle\!\!\right\rangle_I = \langle\!\langle U_1, U_2 \rangle\!\rangle_I \, , \tag{2.8}$$

where U_i are the left-translated velocities V_i to the algebra $\mathfrak{g}$. Any inner product at the algebra can be propagated to the whole group using (2.8) defining a left-invariant Riemannian metric. Correspondingly, an inertia operator $\boldsymbol{\mathcal{M}}_I \colon \mathfrak{g} \to \mathfrak{g}^*$ (henceforth referred to as $\boldsymbol{\mathcal{M}}$) can be defined and properly propagated[7] to $\boldsymbol{\mathcal{M}}_X$. In terms of algebra coordinates, $\boldsymbol{\mathcal{M}}$ can be represented by a $k{\times}k$ symmetric and positive definite matrix, denoted by $\boldsymbol{m}$, such that

$$\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \langle \boldsymbol{\mathcal{M}}(U_1), U_2 \rangle = \sum_{i,j}^{k} \mu_1^i \mu_2^j \langle \boldsymbol{\mathcal{M}}(B_i), B_j \rangle = \bar{\mu}_1^T \boldsymbol{m} \bar{\mu}_2 \, , \tag{2.9}$$

where $U_1, U_2 \in \mathfrak{g}$. The entries $\boldsymbol{m}^{i,j}$ measure the energy of differential transformations around the identity. These values are application dependent, and some heuristic strategies for their selection are given in [129] and [173, Appendix 3].

The corresponding Riemannian distance is invariant to left-translations

$$\mathbf{distance}(X, Y) = \mathbf{distance}(ZX, ZY) = \mathbf{distance}(I, X^{-1}Y)$$

for any $X, Y, Z \in \mathbf{G}$. The same invariance property applies to geodesics, yielding

$$Y \mathrm{Exp}_X(t\,V) = \mathrm{Exp}_{YX}(t\,YV) \, , \tag{2.10}$$

whose length is

$$\mathbf{Length}(\mathrm{Exp}_X(t\,V); 0, 1) = \langle\!\langle V, V \rangle\!\rangle_X^{1/2} = \langle\!\langle U, U \rangle\!\rangle_I^{1/2} = \|U\|_I \, ,$$

with $U = X^{-1}V \in \mathfrak{g}$. Similarly, $Y\,\mathrm{Log}_X(Z) = \mathrm{Log}_{YX}(YZ)$.

Whenever right-invariance is required, all previous concepts can be derived using the same procedure but replacing left-translations $\mathbf{L}_{(\cdot)}$ with right-translations $\mathbf{R}_{(\cdot)}$. Alternatively, the following relations can be used:[8]

$$\mathrm{Exp}_I^{right}(U) = \left(\mathrm{Exp}_I^{left}(-U)\right)^{-1} , \tag{2.11}$$

$$\mathbf{distance}^{right}(X, Y) = \mathbf{distance}^{left}\left(X^{-1}, Y^{-1}\right) . \tag{2.12}$$

[7] In abstract notation, $\boldsymbol{\mathcal{M}}_e$ is left-invariantly propagated to the whole group as $\boldsymbol{\mathcal{M}}_\chi = ((T_\chi \mathbf{L}_{\chi^{-1}})^* \circ \boldsymbol{\mathcal{M}}_e \circ T_\chi \mathbf{L}_{\chi^{-1}}) \colon T_\chi \mathcal{G} \to T_\chi^* \mathcal{G}$, where $(T_\chi \mathbf{L}_{\chi^{-1}})^*$ is the cotangent-lift of the $\mathbf{L}_\chi$ which translate linear functions from $\mathfrak{g}^*$ to $T_\chi^*\mathcal{G}$ (see [12, Appendix B]).

[8] In general and for abstract groups, $\mathrm{Exp}_\chi^{right}(\mathcal{V}) = \mathbf{inv}(\mathrm{Exp}_{\chi^{-1}}^{left}(-(T_{\chi^{-1}}\mathbf{L}_\chi)^{-1} \circ T_\chi \mathbf{R}_{\chi^{-1}} \mathcal{V}))$.

Remarks. For a Lie group endowed with an invariant Riemannian metric, geodesics are complete; i.e., Riemannian geodesics can be extended to the whole real line, and $\mathrm{Exp}_I(\cdot)$ maps the whole algebra onto the group [62]. Moreover, according to the Hopf–Rinow theorem, any two elements $A, B \in \mathbf{G}$ belonging to the same connected component of the group can be connected with a Riemannian geodesic; i.e., there exists at least one velocity V such that $\mathrm{Exp}_A(V) = B$.

2.8. Bi-invariant metrics. When a Riemannian metric is both left- and right-invariant, it is called *bi-invariant*, and the Riemannian exponential function around the identity coincides with the group exponential function. Therefore, geodesics are translated one-parameter subgroups through the identity; that is,

$$\mathrm{Exp}_X(V) = X \exp(X^{-1}V) = \exp(VX^{-1})X\,.$$

For *commutative groups*, where $\mathbf{L}_X$ coincides with $\mathbf{R}_X$, any left-invariant metric is bi-invariant. Some noncommutative groups admit a bi-invariant metric, although not all left-invariant metrics are bi-invariant. Compact groups always admit bi-invariant metrics [14]. Some noncompact groups also admit bi-invariant metrics, but this is not the general case. For example, no bi-invariant metric exists for the group of rigid transformations $\mathcal{SE}(d)$ [180]. Additionally, if the group does not admit a bi-invariant Riemannian metric, then no bi-invariant $\mathcal{C}^2$ (with continuous second derivatives) distance metric can be defined on the group [129]. More information about sufficient and necessary conditions for the existence of bi-invariant metric can be found in [118, 50, 62].

2.9. Geodesics as "straightest" curves. For the sake of completeness, this section introduces geodesics defined via the concept of "straightest" curve in a manifold. While there is a natural way to move forward keeping a direction fixed in a Euclidean space, an additional structure is needed on the manifold allowing to map between nearby tangent spaces. This structure is defined by an affine connection which quantifies the derivative of vector fields on the manifold. Most textbooks about differential geometry define geodesics in terms of the affine connection structure.

We will outline in the following that, for a given Riemannian metric there always exist affine connections whose "straightest" geodesics coincide with Riemannian ("shortest") geodesics defined in section 2.6. Analogously, for a given group there always exist affine connections whose geodesics coincide with the translated group exponential (sections 2.5 and 2.8).

Affine connections, covariant derivative. An *affine connection* on a manifold, denoted by $\nabla_{\mathcal{H}}\mathcal{S} = \mathcal{Z}$, is a smooth map which takes two smooth vector fields $\mathcal{H}$ and $\mathcal{S}$ and returns another smooth vector field $\mathcal{Z}$, satisfying the following properties [37]:

- $\nabla_{(\mathcal{H}+\mathbf{f}\mathcal{P})}\mathcal{S} = \nabla_{\mathcal{H}}\mathcal{S} + \mathbf{f}\nabla_{\mathcal{P}}\mathcal{S}$,
- $\nabla_{\mathcal{H}}(\mathcal{S} + \mathbf{f}\mathcal{P}) = \nabla_{\mathcal{H}}\mathcal{S} + \mathbf{f}\nabla_{\mathcal{H}}\mathcal{P} + (\mathcal{H}\mathbf{f})\mathcal{P}$

for $\mathcal{H}, \mathcal{S}, \mathcal{P}$ smooth vector fields and $\mathbf{f}$ a smooth scalar function on the manifold (in particular, it can be seen by considering $\mathbf{f}$ constant that ∇ is bilinear). From the previous properties it follows that $\mathcal{Z}(X) = (\nabla_{\mathcal{H}}\mathcal{S})(X)$ depends on the value $\mathcal{H}(X)$ and on the values of $\mathcal{S}$ in a neighborhood of X.

In terms of a coordinate chart, the affine connection ∇ can be expressed by

$$\nabla_{\partial_i}\partial_j = \sum_{m=1}^{k} \Gamma_{ij}^{m}\partial_m \,,$$

where the functions $\Gamma_{i,j}^m$ are called the *Christoffel symbols* of ∇. Conversely, any set of k^3 smooth functions $\Gamma_{i,j}^m$ can be chosen in a chart to define an affine connection.

The affine connection ∇ defines univocally the *covariant derivative along a curve* γ, denoted by $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}$, that assigns to every smooth vector field $\mathcal{R}$ defined on γ another smooth vector field along γ. The covariant derivative along the curve $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}$ is compatible with the affine connection ∇ in the sense that $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}(\mathcal{S}\circ\gamma) = (\nabla_{\dot\gamma}\mathcal{S})\circ\gamma$ for any smooth vector field $\mathcal{S}$ defined on the whole manifold. A vector field $\mathcal{R}$ along a curve γ is called *parallel* if $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}(\mathcal{R}) = 0$; similarly, a vector field $\mathcal{S}$ on the whole manifold is called *parallel with respect to the curve* γ if $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}(\mathcal{S}\circ\gamma) = 0$.

In terms of an affine connection defined on the manifold, a ∇-*geodesic* is a curve γ with *acceleration* $\frac{\mathrm{D}_\gamma}{\mathrm{d}t}\dot\gamma$ equal to zero, which means that *the velocity of the curve is parallel to the curve.* In coordinates of a chart, where $\bar{q}_{(t)}$ are the coordinates of $\gamma_{(t)}$, the previous condition results in

$$\ddot{q}^m + \Gamma_{i,j}^m(\bar{q})\dot{q}^i\dot{q}^j = 0\,. \tag{2.13}$$

Given a point X in the manifold and a velocity V in the corresponding tangent space there exist an interval $(-\varepsilon,\varrho)\subset\mathbb{R}$ containing 0 and a unique ∇-geodesic curve γ defined for t in the interval that satisfies $\gamma_{(0)} = X$ and $\dot\gamma_{(0)} = V$. In the case that the interval includes 1, then the ∇-*exponential function* is defined by $\mathrm{EXP}_X(V) = \gamma_{(t=1)}$.

Relation with Riemannian geodesics. Given a Riemannian metric $(\mathbf{G}, \langle\!\langle\cdot,\cdot\rangle\!\rangle_\cdot)$ it is said that the affine connection ∇ is *compatible with the Riemannian metric* when, for any curve γ and two parallel vector fields along it $\mathcal{R}$ and $\mathcal{P}$, the inner products $\langle\!\langle \mathcal{R}(\gamma_{(t)}), \mathcal{P}(\gamma_{(t)})\rangle\!\rangle_{\gamma_{(t)}}$ remain constant along time. The *Fundamental Theorem of Riemannian Geometry* states that for a given (pseudo-)Riemannian metric $(\mathbf{G}, \langle\!\langle\cdot,\cdot\rangle\!\rangle_\cdot)$ there exists a unique affine connection $\nabla^{\langle\!\langle,\rangle\!\rangle}$ on $\mathbf{G}$ that is compatible with the metric $\langle\!\langle\cdot,\cdot\rangle\!\rangle_\cdot$ and fulfills that $\nabla_{\mathcal{H}}^{\langle\!\langle,\rangle\!\rangle}\mathcal{S} - \nabla_{\mathcal{S}}^{\langle\!\langle,\rangle\!\rangle}\mathcal{H} = \mathcal{H}\circ\mathcal{S} - \mathcal{S}\circ\mathcal{H}$. The resulting affine connection is called the *Levi-Civita connection* for the given (pseudo)metric. With respect to a coordinate chart, the Christoffel symbols Γ_{ij}^m corresponding to the Levi-Civita connection $\nabla^{\langle\!\langle,\rangle\!\rangle}$ satisfy at each point

$$2\sum_{m=1}^{k} \Gamma_{ij}^m \langle\!\langle \partial_m, \partial_c\rangle\!\rangle = \partial_i\langle\!\langle \partial_j, \partial_c\rangle\!\rangle + \partial_j\langle\!\langle \partial_i, \partial_c\rangle\!\rangle - \partial_c\langle\!\langle \partial_i, \partial_j\rangle\!\rangle\,. \tag{2.14}$$

It can be shown that $\nabla^{\langle\!\langle,\rangle\!\rangle}$-geodesics are local extremals of the corresponding Riemannian energy given in (2.5), leaving the endpoints fixed [142]; i.e., they are Riemannian geodesics, and $\mathrm{EXP}_X(V) \equiv \mathrm{Exp}_X(V)$.

Left-invariant affine connections. On a group $\mathbf{G}$, an affine connection ∇ is said to be left-invariant [143] if for any two left-invariant vector fields $\widetilde{\mathcal{H}}$ and $\widetilde{\mathcal{S}}$, the resulting $\nabla_{\widetilde{\mathcal{H}}}\widetilde{\mathcal{S}}$ is also a left-invariant vector field.[9] As any smooth vector field can be described as a linear combination of left-invariant vector fields, a left-invariant affine connection is completely defined by its action on algebra vectors. Therefore, ∇ is uniquely defined in terms of a bilinear multiplication operation known as the *left connection function* for ∇ [150]:

$$\boldsymbol{\alpha}(\cdot,\cdot)\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}\colon (U,W)\mapsto \boldsymbol{\alpha}(U,W)=\left(\nabla_{\widetilde{U}}\widetilde{W}\right)(I)\,.$$

In terms of algebra coordinates $U = \sum_i^k B_i\mu^i$ and $W = \sum_i^k B_i\omega^i$, the multiplication $\boldsymbol{\alpha}$ is determined by a $k{\times}k{\times}k$ array α_{ij}^m such that

$$(\boldsymbol{\alpha}(U,W))^m = \sum_{i,j}\alpha_{ij}^m\mu^i\omega^j,$$

and any array of k^3 numbers defines univocally a left-invariant affine connection.

For a left-invariant affine connection ∇, the corresponding ∇-geodesics are left-invariant; i.e., $Y\mathrm{EXP}_X(tV)=\mathrm{EXP}_{YX}(tYV)$. In terms of the array α_{ij}^m, a ∇-geodesic depends only on the symmetric part $(\alpha_{ij}^m+\alpha_{ji}^m)/2$.

For a left-invariant Riemannian metric $\langle\!\langle\cdot,\cdot\rangle\!\rangle_{\cdot}$ on $\mathbf{G}$, its corresponding Levi-Civita connection $\nabla^{\langle\!\langle,\rangle\!\rangle}$ is a left-invariant affine connection [13, 92] and satisfies

$$2\,\boldsymbol{\alpha}(U,W)=[U,W]-\boldsymbol{\mathcal{M}}^{-1}(\mathbf{ad}_U^*(\boldsymbol{\mathcal{M}}(W)))-\boldsymbol{\mathcal{M}}^{-1}(\mathbf{ad}_W^*(\boldsymbol{\mathcal{M}}(U)))\,,$$

where $\mathbf{ad}^*$ is the *dual* of the adjoint operator (see section 3.3).

Bi-invariant affine connections. Left-invariant connections defined by $\boldsymbol{\alpha}(U,W)=\lambda[U,W]$, where λ is any real number, are also bi-invariant connections and are called *Cartan connections* [97, 143, 135, 134]. Their symmetric parts vanish for all λ, and therefore their ∇-geodesics coincide. Moreover, ∇-geodesics departing from the identity coincide with the one-parameter subgroup curves defined in section 2.5. Therefore, under a Cartan connection, $\mathrm{EXP}_I(U)\equiv\exp(U)$ for any $U\in\mathfrak{g}$, and, more generally, $\mathrm{EXP}_X(V)=X\exp(X^{-1}V)$.

Final remarks. To summarize this background section, we now make a few remarks on the definition of geodesics. Roughly speaking, a geodesic on a manifold is a curve, i.e., a time-dependent smooth function, which may be specified by nothing more than a point passing through and its velocity in that point. A more elaborate and formal definition requires additional information about the structures defined on the given manifold. In the case of a manifold endowed with an affine connection, a geodesic is a curve whose velocity is parallel, and the curve goes forward in the "straightest" way. If a (pseudo-)Riemannian metric is defined on the manifold, then a geodesic is a curve connecting two fixed points with an extreme value of its Riemannian energy. If a Riemannian metric is defined on the manifold, a geodesic is an arc-length parametrized curve which is formed by concatenating locally shortest segments.

Three different exponential functions were defined in this section: the ∇-exponential, $\mathrm{EXP}(\cdot)$, is a function defined by an affine connection; the Riemannian exponential, $\mathrm{Exp}(\cdot)$, is

[9] Alternatively, left-invariant connections are also defined [135, 109, 150] as those satisfying $\nabla_{(T\mathbf{L}_Y\mathscr{H})}(T\mathbf{L}_Y\mathcal{S})=T\mathbf{L}_Y\,(\nabla_{\mathscr{H}}\mathcal{S})$ for any $Y\in\mathbf{G}$ and any smooth vector fields $\mathscr{H}$ and $\mathcal{S}$.

a function on a manifold equipped with a (pseudo-)Riemannian metric; and the group exponential, $\exp(\cdot)$, is defined for groups, and only group operations are involved in its definition. On a Lie group, the three curves, $\mathrm{EXP}_I(tU)$, $\mathrm{Exp}_I(tU)$, and $\exp(tU)$, start at the identity with the same velocity $U \in \mathfrak{g}$. The former is the "straightest" curve, the second is guided by a metric yielding *least action* curves, and the latter is guided by the group composition generating one-parameter subgroups.

In a spatial transformation group (or, in general, for any Lie group) Riemannian metrics can be defined taking into account the symmetries of the group. A nice property of these invariant Riemannian metrics is that $\mathrm{Exp}_X(\cdot)$ is a surjective mapping from $T_X\mathbf{G}$ to the connected component of the group. In addition, a Riemannian metric provides a metric space structure to the manifold. For many applications (for example, those involving statistical analysis on manifolds), a metric structure is a key requirement.

3. Computing left-invariant Riemannian geodesics.

Several descriptions of a geodesic were overviewed in the previous section, but from now on only Riemannian geodesics will be considered and will be called geodesics for short. In addition to the theoretical aspects of geodesics previously discussed, several approaches to computing geodesics are presented next.

3.1. Computing geodesics in coordinate charts.

Let $(\mathcal{D}, \varphi)$ be a chart of $\mathbf{G}$, and let $\Omega \subseteq \mathbb{R}^k$ be the image of φ, as was defined in section 2.2. In the following it will be assumed that all variables involved in the expressions can be represented in the same chart $(\mathcal{D}, \varphi)$.

The left-translation operator, $\mathbf{L}_Y$, was introduced in section 2.4. In chart coordinates the operation $\mathsf{L}_{\bar{y}}$ can be defined as $\mathsf{L}_{\bar{y}}\bar{x} = \varphi(\mathbf{L}_{\varphi^{-1}(\bar{y})}\varphi^{-1}(\bar{x})) = \varphi(\varphi^{-1}(\bar{y})\varphi^{-1}(\bar{x}))$. Analogously, the operation $(\bar{x})^{-1}$ can be defined as the coordinates of the element X^{-1}, i.e., $(\cdot)^{-1}\colon \Omega \to \Omega\colon \bar{x} \mapsto (\bar{x})^{-1} = (\varphi \circ \mathbf{inv} \circ \varphi^{-1})(\bar{x})$.

Any curve γ with $\gamma_{(0)} = X$ and $\dot{\gamma}_{(0)} = V$ can be described in chart coordinates as $\varphi \circ \gamma = \bar{\gamma} \in \Omega \subseteq \mathbb{R}^k$ passing through $\bar{x} = \varphi(X)$ at $t = 0$ with velocity $\bar{v} = \mathrm{d}_s\bar{\gamma}_{(s)}\big|_{(s=0)} \in T_{\bar{x}}\Omega \equiv \mathbb{R}^k$. We define the function $\mathrm{d}\varphi\colon T\mathcal{G} \to \mathbb{R}^k$ as the linear operator such that $\mathrm{d}\varphi(V) = \bar{v}$.

Finally, the tangent-lift of the operation $\mathbf{L}_Y$, $T_X\mathbf{L}_Y V$, which translates the velocity $V \in T_X\mathcal{G}$ to $T_{YX}\mathcal{G}$, has a corresponding operation in coordinates $T_{\bar{x}}\mathsf{L}_{\bar{y}}\bar{v} = (\mathrm{d}\varphi \circ T_{\varphi^{-1}(\bar{x})}\mathbf{L}_{\varphi^{-1}(\bar{y})} \circ (\mathrm{d}\varphi)^{-1})\bar{v} = \bar{w} \in \mathbb{R}^k$, where $\bar{w}$ is the velocity of the curve $\mathsf{L}_{\bar{y}}\bar{\gamma}$ at $t = 0$.

According to section 2.7, a left-invariant metric fulfills $\langle\!\langle \cdot, \cdot \rangle\!\rangle_X = \langle\!\langle T_X\mathbf{L}_{X^{-1}} \cdot\, , T_X\mathbf{L}_{X^{-1}} \cdot \rangle\!\rangle_I$, which is written in coordinates as

$$\begin{aligned}
\bar{v}^T\boldsymbol{m}_{\bar{x}}\bar{v} &= \left\langle\!\!\left\langle \mathrm{d}\varphi^{-1}(\bar{v})\, , \mathrm{d}\varphi^{-1}(\bar{v}) \right\rangle\!\!\right\rangle_{\varphi^{-1}(\bar{x})} \\
&= \langle\!\langle V, V \rangle\!\rangle_X = \langle\!\langle T_X\mathbf{L}_{X^{-1}}V\, , T_X\mathbf{L}_{X^{-1}}V \rangle\!\rangle_I \\
&= \left(T_{\bar{x}}\mathsf{L}_{(\bar{x})^{-1}}\bar{v}\right)^T \boldsymbol{m}_{\bar{e}} \left(T_{\bar{x}}\mathsf{L}_{(\bar{x})^{-1}}\bar{v}\right) \\
&= \bar{u}^T\boldsymbol{m}_{\bar{e}}\bar{u}\,,
\end{aligned} \tag{3.1}$$

where $\bar{e}$ are the coordinates of the identity element and $\bar{u} = T_{\bar{x}}\mathsf{L}_{(\bar{x})^{-1}}\bar{v}$ are the resulting coordinates of translating the velocity $\bar{v}$ to the identity. The $k{\times}k$ matrix $\boldsymbol{m}_{\bar{e}}$ is symmetric and positive-definite and defines the Riemannian energy of a differential perturbation around the identity; i.e., it is a representation of the inertia operator $\boldsymbol{\mathcal{M}}$ in the given coordinate chart.

The variational formulation of Riemannian geodesics results in

$$\begin{array}{ll} \underset{\bar{q}_{(t)}\in\Omega\subset\mathbb{R}^k}{\text{extremize}} & \frac{1}{2}\int_0^1 \dot{\bar{q}}_{(t)}^T \boldsymbol{m}_{\bar{q}} \dot{\bar{q}}_{(t)} \\ \text{subject to} & \bar{q}_{(0)} = \bar{q}_0 \quad , \quad \bar{q}_{(1)} = \bar{q}_1 \end{array}$$

as long as $\bar{q}_0$, $\bar{q}_1$, and the whole path $\bar{q}_{(t)}$ belong to Ω. This is a classical variational problem with Lagrangian function $\mathscr{L}(t,\bar{q},\dot{\bar{q}}) = \frac{1}{2}\,(\,\dot{\bar{q}}_{(t)}^T \boldsymbol{m}_{\bar{q}} \dot{\bar{q}}_{(t)}\,)$.

It is widely known that the extremals of the functional are given by the curves $\bar{q}_{(t)}$ which satisfy the endpoint constraints and the Euler–Lagrange equation [169]

$$\partial_{\bar{q}}\, \mathscr{L}(t,\bar{q},\dot{\bar{q}}) - \mathrm{d}_t\, \partial_{\dot{\bar{q}}}\, \mathscr{L}(t,\bar{q},\dot{\bar{q}}) = 0\,. \tag{E–L}$$

In matrix notation and using the matrix calculus rules described in [1, Chapter 13], [111], it can be shown that

$$\begin{aligned} \partial_{\bar{q}}\mathscr{L}(t,\bar{q},\dot{\bar{q}}) &= \frac{1}{2}\left(\dot{\bar{q}}^T \otimes \dot{\bar{q}}^T\right)(\boldsymbol{\mathcal{D}}\boldsymbol{m}_{\bar{q}}), \\ \partial_{\dot{\bar{q}}}\mathscr{L}(t,\bar{q},\dot{\bar{q}}) &= \dot{\bar{q}}^T\, \boldsymbol{m}_{\bar{q}}, \\ \mathrm{d}_t\, \partial_{\dot{\bar{q}}}\mathscr{L}(t,\bar{q},\dot{\bar{q}}) &= \ddot{\bar{q}}^T\, \boldsymbol{m}_{\bar{q}} \,+\, \dot{\bar{q}}^T\,(\boldsymbol{\mathcal{D}}\boldsymbol{m}_{\bar{q}})^T\, \left(\dot{\bar{q}} \otimes I_k\right), \end{aligned}$$

where $(\boldsymbol{\mathcal{D}}\boldsymbol{m}_{\bar{q}})$ is the Fréchet derivative of the matrix $\boldsymbol{m}_{\bar{q}}$ with respect to $\bar{q}$ as a $k^2{\times}k$ matrix.[10]

By rearranging terms, it is obtained that the extremality conditions can be written in coordinates as

$$\ddot{\bar{q}} = \boldsymbol{m}_{\bar{q}}^{-1}\left(\frac{1}{2}(\boldsymbol{\mathcal{D}}\boldsymbol{m}_{\bar{q}})^T\left(\dot{\bar{q}}\otimes\dot{\bar{q}}\right) - \left(\dot{\bar{q}}^T \otimes I_k\right)(\boldsymbol{\mathcal{D}}\boldsymbol{m}_{\bar{q}})\dot{\bar{q}}\right). \tag{odeChart}$$

The Riemannian exponential function $\mathrm{Exp}_X(V)$ is obtained by solving the ODE (odeChart) with initial conditions $\bar{q}_{(0)} = \varphi(X)$ and $\dot{\bar{q}}_{(0)} = \mathrm{d}\varphi(V)$ up to time $t=1$, $\mathrm{Exp}_X(V) = \varphi^{-1}(\bar{q}_{(1)})$. The expression given in (odeChart) is a matrix expression corresponding to the classical geodesic equation $\ddot{q}^m + \Gamma^m_{i,j}(\bar{q})\dot{q}^i\dot{q}^j = 0$ given in (2.13) using $\Gamma^m_{i,j}$ from (2.14)

$$\Gamma^m_{i,j}(\bar{q}) = \frac{1}{2}\sum_{c=1}^{k}\left(\boldsymbol{m}_{\bar{q}}^{-1}\right)^{m,c}\left(\partial_{q^i}\boldsymbol{m}_{\bar{q}}^{c,j} + \partial_{q^j}\boldsymbol{m}_{\bar{q}}^{c,i} - \partial_{q^c}\boldsymbol{m}_{\bar{q}}^{j,i}\right).$$

The left-invariant Riemannian exponential can be solved in a general way by performing the following steps: (a) choose a chart $(\mathcal{D},\varphi)$; (b) translate the initial velocity to a convenient element of the chart domain $\mathcal{D}$ by using (2.10); (c) evolve the ODE (odeChart) up to a time t ensuring that the solution remains in $\Omega \equiv \varphi^{-1}(\mathcal{D})$; (d) eventually retranslate the solution at time t, and continue the evolution up to time $t=1$.

Note that all operations written in coordinates, along with the matrices $\boldsymbol{m}_{\bar{q}}$, are uniquely determined once the function φ and $\boldsymbol{m}_{\bar{e}}$ are chosen.

[10] For a matrix function of matrix argument $\mathbf{F}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{p\times q}$, $\boldsymbol{\mathcal{D}}\mathbf{F}(M)$ is the $(pq){\times}(mn)$ linear operator fulfilling $(\boldsymbol{\mathcal{D}}\mathbf{F}(M))^{i,j} = \partial_{\overline{M}^j}\overline{\mathbf{F}(M)}^i$ such that $\overline{\mathbf{F}(M+\varepsilon P)} = \overline{\mathbf{F}(M)} + \varepsilon(\boldsymbol{\mathcal{D}}\mathbf{F}(M))\overline{P} + O(\varepsilon)$.

Appendix A describes the derivation of the differential equations which define geodesics in the scale and the translation group in one dimension, $\mathcal{ST}(1)$.

This method to compute geodesics has two limitations. First, different charts $(\mathcal{D}, \varphi)$ must be selected for each group, and each choice of the function φ involves a different ODE, making it more difficult to define a general computational method or algorithm valid for any group. Second, the selected chart is usually not valid for the whole domain, and it is required to test when the solution approaches the boundary of the domain. On the other hand, smart selections of the chart may yield particular ODEs for which closed-form solutions are known. Unfortunately, this happens in very few cases.

3.2. Solving geodesics in the group exponential chart. Some of the limitations of using the ODE (odeChart) can be circumvented by selecting a function φ valid for any group. The group exponential function $\exp(\cdot)$, defined in section 2.5, is a possible choice where group elements can be expressed as $\exp(\sum_i B_i \mu^i)$, $\{B_i\}$ being bases of the algebra $\mathfrak{g}$. Moreover, this choice allows the definition of a domain where the chart is valid. Using this chart, geodesics can be computed for any matrix group where only the algebra basis $\{B_i\}$ and the inertia operator must be specified.

The resulting framework is roughly described in Appendix B, being a generic solution valid for any matrix group and under any left-invariant metric. However, this approach requires computation of the derivatives of the matrix exponential function, $\boldsymbol{\mathcal{D}}\exp(\cdot)$ and $\boldsymbol{\mathcal{D}}^2\exp(\cdot)$, and of the matrix logarithm function, $\boldsymbol{\mathcal{D}}\log(\cdot)$, which are computationally expensive to evaluate.

3.3. Evolution of geodesics described in the algebra. The description of geodesics in local coordinates given above usually involves highly nonlinear differential equations, which require numerical solutions with high computational cost. Alternatively, the symmetries of the group and metrics allow us to describe the evolution of geodesics in any tangent space. For convenience the algebra is chosen. Accordingly, the following variational problem is proposed:

$$\text{(varQ)} \qquad \begin{aligned} &\underset{Q_{(t)}\in\mathbf{G}}{\text{extremize}} \quad \mathbf{E}(Q) = \frac{1}{2}\int_0^1 \left\langle\!\!\left\langle \dot{Q}_{(t)}, \dot{Q}_{(t)} \right\rangle\!\!\right\rangle_{Q_{(t)}} \mathrm{d}t \\ &\text{subject to} \quad Q_{(t=0)} = Q_0 \quad ; \quad Q_{(t=1)} = Q_1 \end{aligned}$$

with the corresponding Lagrangian $\mathscr{L}(t, Q, \dot{Q}) = \frac{1}{2}\left\langle\!\!\left\langle \dot{Q}_{(t)}, \dot{Q}_{(t)} \right\rangle\!\!\right\rangle_{Q_{(t)}}$. For a smooth family of curves $Q_{(\cdot,\epsilon)} \in \mathbf{G}$, such that $\partial_\epsilon Q_{(t,\epsilon)}\big|_{(\epsilon=0)} = \delta Q_{(t)} \in T_{Q_{(t)}}\mathbf{G}$ and with $\delta Q_{(0)} = \delta Q_{(1)} = 0$, the curve $Q_{(\cdot,0)}$ is a geodesic if $\mathrm{d}_\epsilon \mathbf{E}(Q_{(\cdot,\epsilon)})\big|_{(\epsilon=0)}$ vanishes.

Using the left-invariance property given in (2.8), the Lagrangian can be rewritten as $\mathscr{L}(t, Q, \dot{Q}) = \frac{1}{2}\left\langle\!\!\left\langle Q_{(t)}^{-1}\dot{Q}_{(t)}, Q_{(t)}^{-1}\dot{Q}_{(t)} \right\rangle\!\!\right\rangle_I$. Letting $U_{(t)} = Q_{(t)}^{-1}\dot{Q}_{(t)} \in \mathfrak{g}$, problem (varQ) is equivalent to

$$\text{(varU)} \qquad \begin{aligned} &\underset{U_{(t)}=Q_{(t)}^{-1}\dot{Q}_{(t)}}{\text{extremize}} \quad \frac{1}{2}\int_0^1 \left\langle\!\left\langle U_{(t)}, U_{(t)} \right\rangle\!\right\rangle_I \mathrm{d}t \\ &\text{subject to} \quad Q_{(t=0)} = Q_0 \quad ; \quad Q_{(t=1)} = Q_1, \end{aligned}$$

where the Lagrangian is *reduced* to $\boldsymbol{\ell}(U) = \frac{1}{2}\left\langle\!\left\langle U_{(t)}, U_{(t)} \right\rangle\!\right\rangle_I$, i.e., the restriction to the identity and to the algebra $\boldsymbol{\ell} = \mathscr{L}\big|_{I,\mathfrak{g}} : \mathfrak{g} \to \mathbb{R}$.

Let $U_{(\cdot,\epsilon)} = Q^{-1}_{(\cdot,\epsilon)}\dot{Q}_{(\cdot,\epsilon)}$ be the family of translated velocities from the curves $Q_{(\cdot,\epsilon)}$ to the algebra. The extremal curves solving (varU) are those fulfilling

$$\mathrm{d}_\epsilon \int_0^1 \boldsymbol{\ell}(U_{(\cdot,\epsilon)})\,\mathrm{d}t = \int \langle \partial_U \boldsymbol{\ell}, \partial_\epsilon U_{(\cdot,\epsilon)} \rangle = 0\,.$$

It is shown in[11] [84] that $\partial_\epsilon U_{(\cdot,\epsilon)} = \dot{W}_{(\cdot,\epsilon)} + \mathbf{ad}_{U_{(\cdot,\epsilon)}} W_{(\cdot,\epsilon)}$, where $W_{(\cdot,\epsilon)} = Q^{-1}_{(\cdot,\epsilon)}\delta Q_{(\cdot)}$ is a curve in $\mathfrak{g}$. Therefore,

$$\begin{aligned}
\left(\mathrm{d}_\epsilon \int_0^1 \boldsymbol{\ell}(U_{(\cdot,\epsilon)})\right)\bigg|_{(\epsilon=0)} &= \left(\int \left\langle \partial_U \boldsymbol{\ell}\,,\, \dot{W}_{(\cdot,\epsilon)} + \mathbf{ad}_{U_{(\cdot,\epsilon)}} W_{(\cdot,\epsilon)} \right\rangle\right)\bigg|_{(\epsilon=0)} \\
&\overset{(a)}{=} \int \left\langle -\frac{\mathrm{d}}{\mathrm{d}t}(\partial_U \boldsymbol{\ell}), W \right\rangle + \int \left\langle \mathbf{ad}^*_U(\partial_U \boldsymbol{\ell})\,,\, W \right\rangle \\
&= \int \left\langle -\frac{\mathrm{d}}{\mathrm{d}t}(\partial_U \boldsymbol{\ell}) + \mathbf{ad}^*_U(\partial_U \boldsymbol{\ell})\,,\, W \right\rangle .
\end{aligned}$$

The first term in (a) was obtained by integrating by parts and using that $W_{(0)} = W_{(1)} = 0$. For the second term, notice that $\mathbf{ad}_U\colon \mathfrak{g} \to \mathfrak{g}$ is a linear operator and any linear operator $\mathbf{f}\colon \mathbb{V} \to \mathbb{V}$ has an associated *dual* operator $\mathbf{f}^*\colon \mathbb{V}^* \to \mathbb{V}^*$ such that $\langle \mathfrak{w}, \mathbf{f}(u)\rangle = \langle \mathbf{f}^*(\mathfrak{w}), u\rangle$. Finally, $\partial_U \boldsymbol{\ell} = \boldsymbol{\mathcal{M}}(U) \in \mathfrak{g}^*$, and since the above variations must vanish for any W, it follows that extremal curves of (varU) satisfy the *Euler–Poincaré equation*[12,13]

$$\text{(E–P)}\qquad \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\mathcal{M}}(U_{(t)}) = \mathbf{ad}^*_{U_{(t)}}(\boldsymbol{\mathcal{M}}(U_{(t)}))\,.$$

The Euler–Poincaré equation (E–P) together with the *reconstruction equation* $\dot{Q}_{(t)} = Q_{(t)}U_{(t)}$ define the evolution of a geodesic.

The major benefit of (E–P) is that the geodesic evolution is described in a single vector space, avoiding the nonlinearities in the metric when using a coordinate chart.

Recalling section 2.4, the velocity $U_{(t)} \in \mathfrak{g}$ can be written as $\sum_i B_i\mu^i$. The corresponding $\boldsymbol{\mathcal{M}}(U) \in \mathfrak{g}^*$ can be written as $\sum_j \mathfrak{D}_j(\boldsymbol{m}\bar{\mu})^j$, where $\{\mathfrak{D}_j\}$ is a basis of $\mathfrak{g}^*$ such that $\mathrm{span}(\{\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_k\}) \equiv \mathfrak{g}^*$ and $\langle \mathfrak{D}_j, B_i\rangle = \delta_{ji}$. A schematic illustration of a curve in the group $\mathbf{G}$ and many involved concepts used for describing the evolution of a geodesic are given in Figure 1.

In terms of the bases, both the primal adjoint $\mathbf{ad}_{(\cdot)}(\cdot)$ and the dual adjoint $\mathbf{ad}^*_{(\cdot)}(\cdot)$ operators are described by the *structure constants* C_{ijm} [154, 42]

$$\mathbf{ad}_{B_i} B_j = \sum_{m=1}^{k} B_m C_{ijm} \qquad \text{and therefore} \qquad \mathbf{ad}^*_{B_i}\mathfrak{D}_j = \sum_{m=1}^{k} \mathfrak{D}_m C_{imj}\,.$$

[11]For a formal proof valid for any abstract group without the particularization to be a matrix group (see [24, Appendix]).

[12]In some textbooks it is called the *basic Euler–Poincaré equation* to emphasize that it only includes the *kinetic energy of a free-body*.

[13]In [121, 83] it is also found that $\dot{U}_{(t)} = \mathbf{ad}^{\intercal}_{U(t)}U_{(t)}$, where $\mathbf{ad}^{\intercal}_{(\cdot)} = (\boldsymbol{\mathcal{M}}^{-1} \circ \mathbf{ad}^*_{(\cdot)} \circ \boldsymbol{\mathcal{M}})$ is the *transpose* of $\mathbf{ad}$ with respect to the inner product $\langle\!\langle \cdot,\cdot \rangle\!\rangle$ satisfying $\langle\!\langle U_1, \mathbf{ad}_W U_2\rangle\!\rangle = \langle\!\langle \mathbf{ad}^{\intercal}_W U_1, U_2\rangle\!\rangle$. The distinction between $\mathbf{ad}^*_W$ and $\mathbf{ad}^{\intercal}_W$ is that the former is defined for a “raw” vector space $\mathbb{V}$, while the latter makes sense only when an inner product is assigned to $\mathbb{V}$.

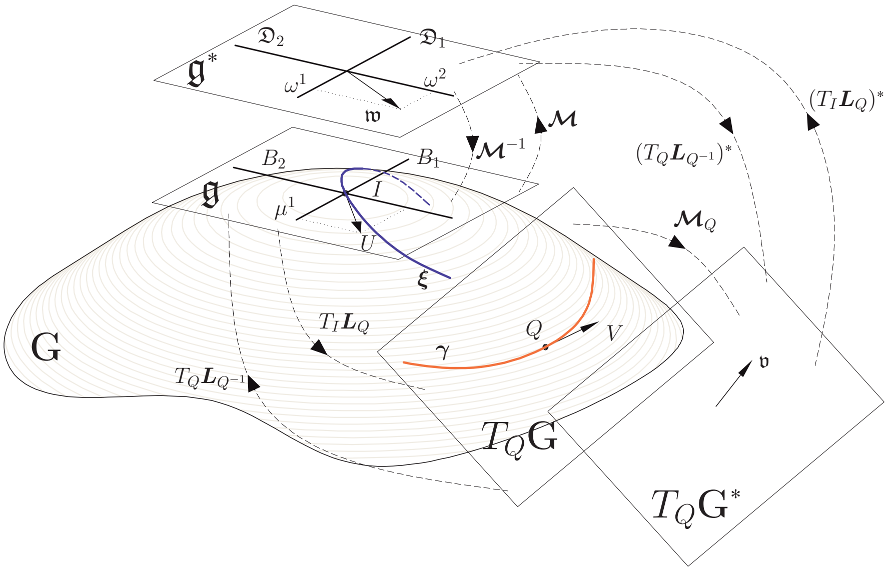


Figure 1. *Schematic illustration of a curve in the group $\boldsymbol{G}$ and many involved concepts used for describing the evolution of a geodesic. Shown are the group $\boldsymbol{G}$, the algebra $\mathfrak{g}$ as the tangent space at I, and its dual space $\mathfrak{g}^*$. The algebra $\mathfrak{g}$ has basis $\{B_1, B_2, \dots\}$, while $\mathfrak{g}^*$ has dual basis $\{\mathfrak{D}_1, \mathfrak{D}_2, \dots\}$ such that $\langle \mathfrak{D}_j, B_i\rangle = \delta_{ji}$. A curve on the group passing through $Q \in \boldsymbol{G}$ with velocity $V \in T_Q\boldsymbol{G}$ is denoted by γ. Translating the curve with $\mathbf{L}_{Q^{-1}}$ the curve ξ is obtained. It passes through I with velocity $U = T_Q\mathbf{L}_{Q^{-1}}V \in \mathfrak{g}$ having algebra coordinates $\bar{\mu} = \left(\mu^1, \mu^2, \dots\right)^T$. This vector can be identified with $\mathfrak{w} = \boldsymbol{\mathcal{M}}(U) \in \mathfrak{g}^*$ with coalgebra coordinates $\left(\omega^1, \omega^2, \dots\right)^T = \boldsymbol{m}\bar{\mu}$. Conversely $U = \boldsymbol{\mathcal{M}}^{-1}(\mathfrak{w})$. The inner product $\langle\!\langle U, U\rangle\!\rangle_I = \langle \mathfrak{w}, U\rangle$ defines a Riemannian metric which computes the* instantaneous kinetic energy *at $\mathfrak{g}$. By the left-invariance requirement of the Riemannian metric, $\langle\!\langle U, U\rangle\!\rangle_I = \langle\!\langle V, V\rangle\!\rangle_Q$, and therefore $V \in T_Q\boldsymbol{G}$ is identified with $\mathfrak{v} = \boldsymbol{\mathcal{M}}_Q(V) \in T_Q^*\boldsymbol{G}$ such that $\langle\!\langle V, V\rangle\!\rangle_Q = \langle \mathfrak{v}, V\rangle_{(T_Q\boldsymbol{G})}$ gives the same kinetic energy as $\langle\!\langle U, U\rangle\!\rangle_I$.*

These structure constants can be computed in the primal space using the matrix representations of the bases of the algebra. Recalling that $\mathbf{ad}_{B_i} B_j = [B_i, B_j]$ where $[\cdot,\cdot]$ is the matrix commutator, the elements C_{ijm} can be calculated as

$$C_{ijm} = \left(\boldsymbol{B}^\dagger(\overline{B_iB_j - B_jB_i})\right)^m .$$

Using the bilinearity property of $\mathbf{ad}^*_{(\cdot)}(\cdot)$,

$$\mathbf{ad}^*_U\mathfrak{w} = \sum_{m=1}^{k} \mathfrak{D}_m \left(\sum_{i,j}^{k} C_{imj}\, \mu^i\omega^j\right) = \sum_{m=1}^{k} \mathfrak{D}_m \nu^j; \tag{3.2}$$

$\bar{\nu} = \left(\nu^1, \nu^2, \dots, \nu^k\right)^T$ are the coalgebra coordinates of $\mathbf{ad}^*_U\mathfrak{w}$. The $\mathbf{ad}^*$ operator in terms of the algebra and coalgebra coordinates will be denoted as $\mathrm{ad}^*_{\bar{\mu}}\bar{\omega} = \bar{\nu}$.

The problem of computing the geodesic $Q_{(t)}$ passing through X with velocity V at $t = 0$ is resolved first by obtaining the curve $\bar{\mu}_{(t)} \in \mathbb{R}^k$ given by the ODE

$$\begin{aligned} & \dot{\bar{\mu}}_{(t)} = \boldsymbol{m}^{-1}(\mathrm{ad}^*_{\bar{\mu}_{(t)}}(\boldsymbol{m}\bar{\mu}_{(t)})) \\ \text{with initial condition}\quad & \bar{\mu}_{(0)} = \boldsymbol{B}^\dagger(\overline{X^{-1}V}), \end{aligned}$$

and afterward by solving the reconstruction equation given by the ODE

$$\text{(recons)} \qquad \begin{aligned} &\dot{Q}_{(t)} = Q\left(\sum_{i=1}^{k} B_i \mu^i_{(t)}\right) \\ \text{with initial condition} \qquad &Q_{(0)} = X\,. \end{aligned}$$

Additional numerical details to solve the reconstruction step will be given in section 9. Once the geodesic is solved, the corresponding Riemannian exponential is $\mathrm{Exp}_X(V) = Q_{(t=1)}$. An illustration of the geodesic evolution, its symmetries, and its relation to the Riemannian exponential function is shown in Figure 2.

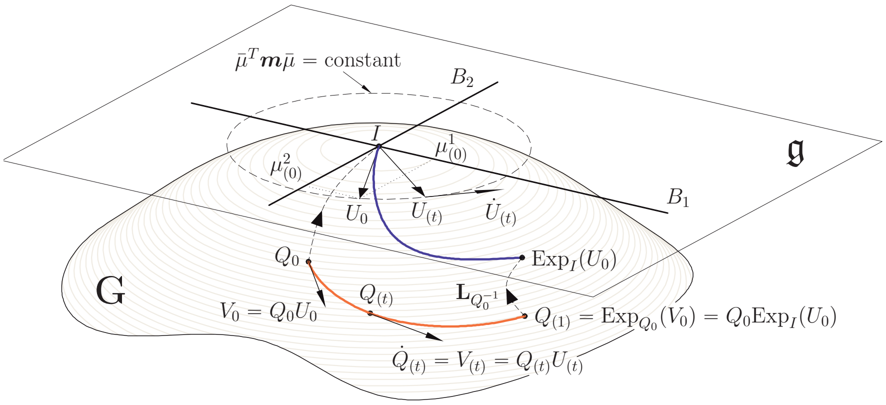


Figure 2. *Illustration of the geodesic evolution. The* red *geodesic segment $Q_{(t)}$ in the group $\boldsymbol{G}$ departs from the point Q_0 with initial velocity V_0. Every instantaneous velocity $V_{(t)} = \dot{Q}_{(t)}$ can be translated from the left by $Q^{-1}_{(t)}$, providing a curve of algebra vectors $U_{(t)} = Q^{-1}_{(t)} V_{(t)} \in \mathfrak{g}$ with coordinates $\bar{\mu}_{(t)} = \left(\mu^1_{(t)}, \mu^2_{(t)}, \ldots\right)^T$ in the algebra bases $\{B_1, B_2, \ldots\}$. The* red *geodesic segment can be left-translated by Q_0^{-1} resulting in the* blue *geodesic segment from I to $\mathrm{Exp}_I(U_0)$. Both segments have the same length.*

3.4. A coordinate-free solution based on (E–P).

While the elements of the algebra have been numerically represented by $n \times n$ matrices (embedding $\mathfrak{g}$ in $\mathbb{M}_n$), their corresponding dual elements (or covectors) living in $\mathfrak{g}^*$ were considered so far as abstract entities. In this section a numerical representation of coalgebra elements as matrices will provide a coordinate-free expression for (E–P).

In a finite dimensional vector space, it is common to interpret primal elements as *column-vectors* and the covectors as *row-vectors* such that the pairing operation results in the ordinary matrix multiplication. In the case we are concerned with, a primal element $U \in \mathfrak{g}$ is numerically represented by a matrix of $\mathbb{M}_n$ which is obtained by reshaping the vector resulting from $\boldsymbol{B}\bar{\mu}$ (see section 2.4). Elements $\mathfrak{w}$ of the coalgebra $\mathfrak{g}^*$ are naturally represented by matrices of the form $\sum_j D_j \omega^j$, where $\bar{\omega} = \left(\omega^1, \omega^2, \ldots, \omega^k\right)^T$ are the coordinates of $\mathfrak{w}$ and D_j are matrix representations of the dual bases $\mathfrak{D}_j$. The matrices D_j must satisfy that the pairing $\langle D_j, B_i \rangle = \delta_{ij}$. In general, there are many possible ways to both choose a matrix representation of D_j and make explicit the pairing operation. A convenient way is by reshaping the rows of a matrix $\boldsymbol{B}^\dagger$, any generalized left inverse of $\boldsymbol{B}$. Analogously to the primal case, $\boldsymbol{D} = \boldsymbol{B}^{\dagger T} =$

$\left(\overline{D}_1, \overline{D}_2, \ldots, \overline{D}_k\right)$ is the matrix built from the vectorization of the dual bases D_j. Then, a covector $\mathfrak{w} \in \mathfrak{g}^*$ with coordinates $\bar{\omega}$ is numerically represented by a matrix $W \in \mathbb{M}_n$ such that $\overline{W} = \boldsymbol{B}^{\dagger T}\bar{\omega}$. In addition, under this representation, the pairing operation $\langle\cdot,\cdot\rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ takes the form

$$\langle W, U\rangle = \sum_m \mu^m \omega^m = (\boldsymbol{D}^\dagger \overline{W})^T (\boldsymbol{B}^\dagger \overline{U}) = \operatorname{trace}(W^T U)\,,$$

where $\bar{\mu}$ and $\bar{\omega}$ are the coordinates of U and W, and $\boldsymbol{D}^\dagger$ and $\boldsymbol{B}^\dagger$ are any generalized left inverse of $\boldsymbol{D}$ and $\boldsymbol{B}$, respectively.

The first step toward a coordinate-free expression for (E–P) is to write the $\mathbf{ad}^*$ operator in terms of the matrix representations defined above. As $\mathbf{ad}_Z(\cdot) = [Z, \cdot]$ is a linear operator on $\mathbb{M}_n \equiv \mathbb{R}^{n^2}$, it can be represented by a matrix $\mathbf{A}_Z$ of $\mathbb{M}_{n^2}$ such that $\overline{\mathbf{ad}_Z U} = \mathbf{A}_Z \overline{U}$, with $\mathbf{A}_Z = (I_n \otimes Z) - (Z^T \otimes I_n)$ for $Z, U \in \mathfrak{g}$. The pairing $\langle W, \mathbf{ad}_Z U\rangle$ can be written using the matrix representations as $(\boldsymbol{D}^\dagger\overline{W})^T(\boldsymbol{B}^\dagger\overline{\mathbf{ad}_Z U})$. Using $\overline{U} = \boldsymbol{B}\boldsymbol{B}^\dagger\overline{U}$ (which is fulfilled only when $U \in \mathfrak{g}$) and $\boldsymbol{D}^T\boldsymbol{D}^{\dagger T} = I_k$, it follows that

$$\begin{aligned}
\langle W, \mathbf{ad}_Z U\rangle &= (\boldsymbol{D}^\dagger\overline{W})^T(\boldsymbol{B}^\dagger\overline{\mathbf{ad}_Z U})\\
&= \overline{W}^T \boldsymbol{D}^{\dagger T}\boldsymbol{B}^\dagger \mathbf{A}_Z \overline{U}\\
&= \overline{W}^T \boldsymbol{D}^{\dagger T}\boldsymbol{B}^\dagger \mathbf{A}_Z (\boldsymbol{B}\boldsymbol{B}^\dagger)\overline{U}\\
&= \overline{W}^T \boldsymbol{D}^{\dagger T}\boldsymbol{B}^\dagger \mathbf{A}_Z \boldsymbol{B}(\boldsymbol{D}^T\boldsymbol{D}^{\dagger T})\boldsymbol{B}^\dagger\overline{U}\\
&= \left(\boldsymbol{D}^\dagger(\boldsymbol{D}\boldsymbol{B}^T\mathbf{A}_Z^T\boldsymbol{B}^{\dagger T}\boldsymbol{D}^\dagger)\overline{W}\right)^T(\boldsymbol{B}^\dagger\overline{U})\\
&= (\boldsymbol{D}^\dagger\overline{\mathbf{ad}_Z^* W})^T(\boldsymbol{B}^\dagger\overline{U})\\
&= \langle \mathbf{ad}_Z^* W, U\rangle\,.
\end{aligned}$$

Therefore, $\mathbf{ad}_Z^*$ is identified with the matrix operation given by $\overline{\mathbf{ad}_Z^* W} = \boldsymbol{D}\boldsymbol{B}^T\mathbf{A}_Z^T\boldsymbol{D}\boldsymbol{D}^\dagger\overline{W}$. In order to get more compact expressions, from now on $(\cdot)^\dagger$ will be chosen as the Moore–Penrose pseudoinverse; then $\boldsymbol{B}^T = \boldsymbol{D}^\dagger$.

It is easy to show that for any $Z \in \mathbb{M}_n$, $\mathbf{A}_Z^T = \mathbf{A}_{Z^T}$ holds; i.e., $\mathbf{A}_Z^T$ is the linear operator associated to the matrix commutation $[Z^T, \cdot]$. Therefore,

$$\overline{\mathbf{ad}_Z^* W} = \boldsymbol{D}\boldsymbol{D}^\dagger\mathbf{A}_{Z^T}\overline{W} \Rightarrow \mathbf{ad}_Z^* W = \boldsymbol{\pi}[Z^T, W]\,,$$

where $\boldsymbol{\pi}\colon \mathbb{M}_n \to \mathfrak{g}^*\colon \overline{\boldsymbol{\pi}(X)} = \boldsymbol{D}\boldsymbol{D}^\dagger\overline{X} = \boldsymbol{B}^{\dagger T}\boldsymbol{B}^T\overline{X}$ is a projector and $\boldsymbol{D}\boldsymbol{D}^\dagger\overline{W} = \overline{W}$ for $W \in \mathfrak{g}^*$. In this expression, $[\cdot,\cdot]$ is the matrix commutation and should not be confused with the Lie bracket. The matrix commutation coincides with the Lie bracket only when both operands belong to the algebra. In general, the resulting matrix $[Z^T, W]$ does not represent elements belonging to either $\mathfrak{g}$ or $\mathfrak{g}^*$. Therefore, the projector $\boldsymbol{\pi}$ is mandatory in the expression of the $\mathbf{ad}^*$ operator. In [47, 27] the authors obtained the same coordinate-free expression, but a metric was used in the space $\mathbb{M}_n$.

The second step is to consider the inertia operator $\boldsymbol{\mathcal{M}}$. Using the previous numerical representations, the identification between $\mathfrak{g}$ and $\mathfrak{g}^*$ carried out by the operation $W = \boldsymbol{\mathcal{M}}(U)$ is written as $\overline{W} = (\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^\dagger)\overline{U}$. The inverse, $\boldsymbol{\mathcal{M}}^{-1}\colon \mathfrak{g}^* \to \mathfrak{g}\colon W \mapsto U = \boldsymbol{\mathcal{M}}^{-1}(W)$, can be expressed by $\overline{U} = (\boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^T)\overline{W}$. To alleviate algebraic expressions, it is useful to define the linear operators[14]

$$\begin{aligned} &\boldsymbol{\chi}\colon \mathbb{M}_n \to \mathbb{M}_n \quad \text{such that} \quad \overline{\boldsymbol{\chi}(X)} = (\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger})\overline{X}, \\ &\boldsymbol{\sigma}\colon \mathbb{M}_n \to \mathbb{M}_n \quad \text{such that} \quad \overline{\boldsymbol{\sigma}(X)} = (\boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^{T})\overline{X}\,, \end{aligned} \tag{3.3}$$

such that $\boldsymbol{\mathcal{M}}(U) = \boldsymbol{\chi}U$ and $\boldsymbol{\mathcal{M}}^{-1}(W) = \boldsymbol{\sigma}W$.

Recalling the Euler–Poincaré equation (E–P) and considering the matrix representations, we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(\overline{\boldsymbol{\mathcal{M}}(U)}) &= \overline{\mathbf{ad}^*_U\boldsymbol{\mathcal{M}}(U)} \\ &= \boldsymbol{B}^{\dagger T}\boldsymbol{B}^T\mathbf{A}_U^T\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger}\overline{U} \\ \Rightarrow \dot{\overline{U}} = \boldsymbol{\mathcal{M}}^{-1}\left(\overline{\mathbf{ad}^*_U\boldsymbol{\mathcal{M}}(U)}\right) &= \boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^T\boldsymbol{B}^{\dagger T}\boldsymbol{B}^T\mathbf{A}_U^T\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger}\overline{U} \\ &= (\boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^T)\mathbf{A}_{U^T}(\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger})\overline{U}\,. \end{aligned}$$

Then the equation of the velocity guiding a geodesic can be compactly written as

$$\dot{U}_{(t)} = \boldsymbol{\sigma}[U^T_{(t)}, \boldsymbol{\chi}U_{(t)}]\,, \tag{3.4}$$

which, together with the reconstruction step given in (recons), provides a coordinate-free free expression for left-invariant geodesics. Equivalent expressions can be found in [47, 27, 127].

Note that in the case of Rotation group $\mathbf{SO}(d)$ (where the elements of the algebra are skew-symmetric matrices (see section 8)), it is obtained that $\boldsymbol{\chi}\dot{U} = [\boldsymbol{\chi}U, U]$, which is well known as the *generalized rigid body equation* [25].

When $U_{(t)}$ remains constant, i.e., $\boldsymbol{\sigma}[U_0^T, \boldsymbol{\chi}U_0] = 0$, the corresponding geodesic is a translated one-parameter subgroup. Note that even if the metric were not invariant, $\mathrm{Exp}_I(t\,U_0)$ would coincide with $\exp(t\,U_0)$ for some particular initial velocities. These cases are described for the particular case of the Rotation group in [12, Appendix 2.E] and were called *stationary rotations*. Moreover, the reverse can be proven; if $\mathrm{Exp}_I(t\,U_0)$ coincides with $\exp(t\,U_0)$, then $\boldsymbol{\sigma}[U^T_{(t)}, \boldsymbol{\chi}U_{(t)}]$ vanishes for all t. In [145] it was shown that for a given U_0 there always exists an invariant Riemannian metric such that the group exponential $\exp(t\,U_0)$ coincides with $\mathrm{Exp}_I(t\,U_0)$.

In general, (3.4) is a second-order polynomial ODE, and no closed-form solution is known except for a few cases. An interesting particular case of (3.4) occurs when $\boldsymbol{\sigma} = \boldsymbol{\chi} = \mathrm{identity}_{\mathbb{M}_n} = I_{n^2}$, and the solution is given in closed form [10, 170] by

$$U_{(t)} = \exp\left(-t\,(U_0 - U_0^T)\right)\,U_0\,\exp\left(t\,(U_0 - U_0^T)\right)\,.$$

3.5. Left-invariant geodesics on matrix groups using optimal control. The computation of left-invariant geodesics and the Riemannian exponential function on matrix groups can be formulated as the solution of an optimal control problem. Optimal control also provides a

[14]For brevity, when there is no possible misunderstanding, $\boldsymbol{\chi}(X)$ and $\boldsymbol{\sigma}(X)$ will be written as $\boldsymbol{\chi}X$ and $\boldsymbol{\sigma}X$, respectively.

general framework valid for any matrix group and is developed for any left-invariant Riemannian metric, as in sections 3.3 and 3.4. The added value of this approach is that the costate variable will allow us to find conserved quantities (see section 4).

Optimal control is a technique used to minimize an integral cost functional that depends on the evolution of the states of a dynamical system [101, 151], [86, Chapters 11–12]. In a first stage, the dynamical system and the functional are described in terms of *control variables*. Then, the optimal time-behavior of the control variables is solved, and later the dynamical system variables are obtained.

Recalling the variational problem (varQ), the curve $Q_{(t)}$ can be considered as a dynamical system governed or guided by a *control variable* $V_{(t)} = \dot{Q}_{(t)}$, the velocity of the curve. Accordingly, (varQ) is reformulated as a control problem:

$$\text{(ocpV)} \qquad \begin{aligned} &\underset{V_{(t)} \in T_{Q(t)}\mathbf{G}}{\text{extremize}} && \frac{1}{2}\int_0^1 \left\langle\!\left\langle V_{(t)}, V_{(t)} \right\rangle\!\right\rangle_{Q_{(t)}} \mathrm{d}t \\ &\text{subject to} && \dot{Q}_{(t)} = V_{(t)}, \\ & && Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1 \,. \end{aligned}$$

For a left-invariant metric, (2.8) holds; i.e., $\langle\!\langle V_1, V_2\rangle\!\rangle_Q = \langle\!\langle U_1, U_2\rangle\!\rangle_I$, where $U_i = Q^{-1}V_i$. Then, (ocpV) can be written as

$$\text{(ocpU)} \qquad \begin{aligned} &\underset{U_{(t)} \in \mathfrak{g}}{\text{extremize}} && \frac{1}{2}\int_0^1 \left\langle\!\left\langle U_{(t)}, U_{(t)} \right\rangle\!\right\rangle_I \mathrm{d}t \\ &\text{subject to} && \dot{Q}_{(t)} = Q_{(t)}U_{(t)}, \\ & && Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1 \,. \end{aligned}$$

Three different constraints can be identified in (ocpU): the *fixed bound constraints* specify the initial and final elements Q_0 and Q_1, respectively; the *dynamic law* $\dot{Q}_{(t)} = Q_{(t)}U_{(t)}$ constrains the evolution of the dynamical system; finally, the *control variable domain constraint* $U_{(t)} \in \mathfrak{g}$ guarantees that $Q_{(t)} \in \mathbf{G}$ for all t, since $Q_{(t)}U_{(t)} \in T_{Q(t)}\mathbf{G}$.

Writing velocities in terms of the algebra basis, $\overline{U}_{(t)} = \boldsymbol{B}\bar{\mu}_{(t)}$, and using (2.9) to express the instantaneous energy, the optimal control problem can be reformulated without constraints on the control variable:

$$\text{(ocp}\bar{\mu}\text{)} \qquad \begin{aligned} &\underset{\bar{\mu}_{(t)} \in \mathbb{R}^k}{\text{extremize}} && \frac{1}{2}\int_0^1 \bar{\mu}_{(t)}^T \boldsymbol{m}\bar{\mu}_{(t)} \mathrm{d}t = \mathbf{E}(\bar{\mu}) \\ &\text{subject to} && \dot{Q}_{(t)} = Q_{(t)} \sum B_i \mu^i_{(t)}, \\ & && Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1 \,. \end{aligned}$$

Optimal control solves (ocp$\bar{\mu}$) by including the dynamic constraint in an *augmented functional* with the corresponding Lagrange multiplier. This Lagrange multiplier enforces the constraint $\dot{Q}_{(t)} - Q_{(t)}U_{(t)} = 0$, which can be equivalently written as $\left\langle \mathfrak{p}, \dot{Q} - QU \right\rangle_Q = 0$ for all $\mathfrak{p} \in T^*_Q\mathbf{G}$, where $\langle \cdot, \cdot \rangle_Q$ is the pairing in the vector space $T_Q\mathbf{G}$. Again, matrix representations are used to express the pairing $\langle \mathfrak{p}, V\rangle_Q$ as[15] $\mathrm{trace}(P^T V)$. Accordingly, (ocp$\bar{\mu}$) is equivalent to extremizing

$$\mathbf{E}'(\bar{\mu}) = \int_0^1 \left(\frac{1}{2}\bar{\mu}_{(t)}^T \boldsymbol{m}\bar{\mu}_{(t)} + \mathrm{trace}\left(P_{(t)}^T \left(\dot{Q}_{(t)} - Q_{(t)} \sum_i^k B_i \mu_{(t)}^i \right)\right)\right) \mathrm{d}t$$
$$= \int_0^1 \left(\mathscr{H}(Q, \bar{\mu}, P, t) + \mathrm{trace}(P_{(t)}^T \dot{Q}_{(t)}) \right) \mathrm{d}t\,, \tag{3.5}$$

where

$$\mathscr{H}(Q, \bar{\mu}, P, t) = \frac{1}{2}\bar{\mu}_{(t)}^T \boldsymbol{m}\bar{\mu}_{(t)} - \mathrm{trace}\left(P_{(t)}^T Q_{(t)} \sum_i^k B_i \mu_{(t)}^i \right). \tag{3.6}$$

The extremality conditions for the functional $\mathbf{E}'$, defined in (3.5), are [101]

$$\partial_{\bar{\mu}}\mathscr{H} = \boldsymbol{m}\bar{\mu}_{(t)} - \boldsymbol{B}^T(\overline{Q_{(t)}^T P_{(t)}}) = 0, \tag{3.7a}$$

$$-\partial_P \mathscr{H} = Q_{(t)} \sum B_i \mu_{(t)}^i = \dot{Q}_{(t)}, \tag{3.7b}$$

$$-\partial_Q \mathscr{H} = P_{(t)} \sum B_i^T \mu_{(t)}^i = -\dot{P}_{(t)}. \tag{3.7c}$$

Equation (3.7a) relates the *costate variable* $P_{(t)}$ and the *control variables* $\bar{\mu}_{(t)}$ via an algebraic relation with the *state variable* $Q_{(t)}$. This equation can be rewritten in terms of the matrix representation as

$$\overline{U}_{(t)} = \boldsymbol{B}\bar{\mu}_{(t)} = (\boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^T)(\overline{Q_{(t)}^T P_{(t)}}) \Rightarrow U_{(t)} = \boldsymbol{\sigma}(Q_{(t)}^T P_{(t)})\,, \tag{3.8}$$

where $\boldsymbol{\sigma}(\cdot)$ is as defined in (3.3).

The Riemannian geodesic starting at Q_0 with velocity V satisfies the IVP defined by

$$U_{(t)} = \boldsymbol{\sigma}(Q_{(t)}^T P_{(t)}), \tag{3.9a}$$

$$\dot{Q}_{(t)} = Q_{(t)} U_{(t)}, \tag{3.9b}$$

$$\dot{P}_{(t)} = -P_{(t)} U_{(t)}^T, \tag{3.9c}$$

$$Q_{(t=0)} = Q_0. \tag{3.9d}$$

While the initial condition for Q is specified, $P_{(0)}$ remains unspecified. However, $U_{(0)} = Q_0^{-1}V$ is given by the specifications of the geodesic. In addition, $P_{(0)}$ must fulfill $U_{(0)} = \boldsymbol{\sigma}(Q_0^T P_{(0)})$. It can be shown[16] that $P_{(0)} = Q_0^{-T}\boldsymbol{\chi} U_{(0)}$.

Note that this IVP involves two coupled evolutions ($Q_{(t)}$ and $P_{(t)}$), while in the methodologies described in sections 3.3 and 3.4, the evolution of $U_{(t)}$ is independent of $Q_{(t)}$.

[15] Let $\{B_1, B_2, \ldots\}$ be a set of matrix bases of $\mathfrak{g}$; then for a group element Q, the set $\{QB_1, QB2, \ldots\}$ can be used as a set of bases of $T_Q\mathbf{G}$. Moreover, for $\{D_1, D_2, \ldots\}$ a set of bases of a matrix representation of $\mathfrak{g}^*$ (see section 3.4), then $\{Q^{-T}D_1, Q^{-T}D_2, \ldots\}$ can be used as bases of a matrix representation of $T_Q^*\mathbf{G}$ which satisfies $\langle Q^{-T}D_j, QB_i\rangle_Q = \delta_{ji}$. Under these representations $P = \sum_j Q^{-T}D_j\rho^j \in T_Q^*\mathbf{G}$, $V = \sum_i QB_i\nu^i \in T_Q\mathbf{G}$, and $\langle P, V\rangle_Q = \sum \rho^m \nu^m = \mathrm{trace}(P^T V)$.

[16] For a matrix P representing an element of $T_\chi^*\mathcal{G}$, the cotangent-lift $(T_\chi \mathbf{L}_q)^* : T_{q\chi}^*\mathcal{G} \to T_\chi^*\mathcal{G}$ is represented by $Q^T P$. Therefore, $Q^T P \in \mathfrak{g}^*$, and applying from the left the operator $\boldsymbol{\chi}$ defined in (3.3) we obtain $Q^T P = \boldsymbol{\chi} U$.

3.6. Symmetric representation of the geodesic evolution. Some authors have proposed a symmetric representation of the geodesic equations for particular matrix groups and more recently for a general matrix case. A symmetric formulation is given below starting from IVP (3.9). $P_{(t)}$ can be written as $P_{(t)} = P_0 Y_{(t)}$, with $Y_{(0)} = I$; then $\dot{P}_{(t)} = P_0 \dot{Y}_{(t)} = -P_0 Y_{(t)} U_{(t)}^T$. Although P_0 is not necessarily invertible, the previous equation is fulfilled if $\dot{Y}_{(t)} = Y_{(t)}(-U_{(t)}^T)$. Therefore, the IVP (3.9) can be rewritten in terms of $Y_{(t)}$ as

$$U_{(t)} = \boldsymbol{\sigma}(Q_{(t)}^T P_0 Y_{(t)}), \tag{3.10a}$$

$$\dot{Q}_{(t)} = Q_{(t)} U_{(t)}, \tag{3.10b}$$

$$\dot{Y}_{(t)} = Y_{(t)}(-U_{(t)}^T), \tag{3.10c}$$

$$Q_{(t=0)} = Q_0, \tag{3.10d}$$

$$Y_{(t=0)} = I, \tag{3.10e}$$

$$P_0 = Q_0^{-T} \boldsymbol{\chi} U_0 \,. \tag{3.10f}$$

It can be observed that $Y_{(t)}$ starts at the identity and follows a left-invariant field lying on the transposed algebra. Therefore, $Y_{(t)}$ belongs to the *transposed group*[17] $\mathbf{G}^T = \{ X \mid X^T \in \mathbf{G} \subset \mathbb{M}_n \}$, or, equivalently, the group whose algebra is given by $\operatorname{span}(\{B_1^T, B_2^T, \ldots, B_k^T\})$. Then the IVP (3.10) defines a curve in the space $\mathbf{G} \times \mathbf{G}^T$.

In [25] a similar approach was performed on Rotation group $\mathbf{SO}(n)$; in this case the group $\mathbf{SO}(n)^T$ coincides with $\mathbf{SO}(n)$, and the authors called this framework the *symmetric representation of the rigid body*. Later the framework was extended to a similar symmetric representation of some quadratic groups, which include interesting groups like Indefinite Rotation groups and Symplectic groups [26]. A general framework has been recently proposed in [27] from the viewpoint of embedded optimal control problems in $\mathbb{M}_n$.

4. Conserved quantities along geodesics. Noether's theorem states that for every continuous symmetry of a dynamical system there is a corresponding conserved quantity. In the following, two conserved quantities are identified for a geodesic starting at Q_0 with initial velocity V_0: E corresponds to the invariance of the time origin; and S corresponds to the invariance under the group action.

- Geodesics are arc-length parametrized curves, and therefore $\|\dot{Q}_{(t)}\|_{Q_{(t)}}$ is constant, as was illustrated in Figure 2, where $\bar{\mu}_{(t)}$ is constrained to evolve along the elliptical surface defined by $\boldsymbol{m}$. This means that $\langle\!\langle U_{(t)}, U_{(t)} \rangle\!\rangle = \langle \boldsymbol{\mathcal{M}}(U_{(t)}), U_{(t)} \rangle = \langle \boldsymbol{\mathcal{M}}(U_0), U_0 \rangle$. Following section 3.4, $\boldsymbol{\mathcal{M}}(U)$ is represented by $\boldsymbol{\chi} U$, and the pairing is expressed in terms of the trace function. Then, $\langle\!\langle U_{(t)}, U_{(t)} \rangle\!\rangle$ can be compactly written as $\operatorname{trace}\left((\boldsymbol{\chi} U_{(t)})^T U_{(t)}\right)$.

  Finally, in terms of the states of a geodesic, the following quantity is conserved:

$$\begin{aligned} E :&= \langle\!\langle U_{(t)}, U_{(t)} \rangle\!\rangle \\ &= \operatorname{trace}\left(\boldsymbol{\chi} (Q_{(t)}^{-1} \dot{Q}_{(t)})^T Q_{(t)}^{-1} \dot{Q}_{(t)}\right) \\ &= \operatorname{trace}\left(\boldsymbol{\chi} (Q_0^{-1} V_0)^T Q_0^{-1} V_0\right) = \text{constant}. \end{aligned} \tag{4.1}$$

[17]This definition is valid only for matrix groups.

- Another conserved quantity is $P_{(t)}Q_{(t)}^T$, because its time-derivative (see (3.9b) and (3.9c)) vanishes:

$$\begin{aligned}(P_{(t)}Q_{(t)}^T)^{\cdot} &= \dot{P}Q^T + P\dot{Q}^T \\ &= -PU^TQ^T + PU^TQ^T = 0.\end{aligned}$$

Using (3.8) and the fact that $P_{(t)} \in T^*_{Q(t)}\mathbf{G}$, in a similar way as was done to obtain $P_{(0)}$ in section 3.5, we obtain that $P_{(t)} = Q_{(t)}^{-T}\boldsymbol{\chi}U_{(t)}$.

Finally, in terms of the states of a geodesic, the following quantity is also conserved:

$$\begin{aligned}S := {}& P_{(t)}Q_{(t)}^T \\ ={}& Q_{(t)}^{-T}\boldsymbol{\chi}(U_{(t)})Q_{(t)}^T \\ ={}& Q_{(t)}^{-T}\boldsymbol{\chi}\left(Q_{(t)}^{-1}\dot{Q}_{(t)}\right)Q_{(t)}^T \\ ={}& Q_0^{-T}\boldsymbol{\chi}\left(Q_0^{-1}V_0\right)Q_0^T = \text{constant}.\end{aligned} \tag{4.2}$$

From a *physical* viewpoint the quantity E corresponds to the *kinetic energy* which is preserved for a free-body, and the quantity S corresponds to the *spatial angular momentum* [113].

It is worth mentioning that the conserved quantity S in (4.2) can be written in terms of intrinsic operations of the group. The expression $Q^{-T}WQ^T$ can be interpreted as the dual of the Adjoint action defined in section 2.4,

$$\left\langle Q^{-T}WQ^T, U\right\rangle = \operatorname{trace}(QW^TQ^{-1}U) = \operatorname{trace}(W^TQ^{-1}UQ) = \left\langle W, \mathbf{Ad}_{Q^{-1}}U\right\rangle,$$

for $W \in \mathfrak{g}^*$ and $U \in \mathfrak{g}$; therefore[18] $\left(\mathbf{Ad}_{Q^{-1}}\right)^* : \mathfrak{g}^* \to \mathfrak{g}^* : W \mapsto Q^{-T}WQ^T$. Accordingly the conserved quantity S is written as[19] $(\mathbf{Ad}_{Q(t)^{-1}})^*\boldsymbol{\chi}U_{(t)} = $ constant. The quantity S has also been obtained using a different approach, where additional geometrical structures and concepts are used, such as *Jacobi brackets* and *symplectic forms* [2, 113, 84].

5. Reducing to a first-order equation.

The methods for geodesic computation revisited so far include the following:

- Second-order differential equation systems in local coordinates ($\mathbb{R}^k$); see sections 3.1 and 3.2.
- A first-order equation system in algebra coordinates followed by the reconstruction of the group curve ($\mathbb{R}^k \times \mathbf{G}$); see section 3.3.
- A first-order equation in the coalgebra representation followed by the reconstruction ($\mathfrak{g}^* \times \mathbf{G}$); see section 3.4.
- A first-order equation in the representation of cotangent spaces coupled with the group curve evolution ($T^*\mathbf{G} \times \mathbf{G}$); see section 3.5.
- A first-order equation system of coupled evolutions in ($\mathbf{G}^T \times \mathbf{G}$); see section 3.6.

[18] Two different notations are found in the literature: the most prevalent is $\mathbf{Ad}^*_{\chi} = (\mathbf{Ad}_{\chi^{-1}})^*$ satisfying $\mathbf{Ad}^*_{\chi} \circ \mathbf{Ad}^*_{y} = \mathbf{Ad}^*_{\chi y}$ (see [84, p. 215], [92]); the second is $\mathbf{Ad}^*_{\chi} = (\mathbf{Ad}_{\chi})^*$, which fulfills $\mathbf{Ad}^*_{\chi} \circ \mathbf{Ad}^*_{y} = \mathbf{Ad}^*_{y\chi}$ (note the change of the order in the foot) [12].

[19] In abstract notation, $(\,(\mathbf{Ad}_{q_{(t)}^{-1}})^* \circ \boldsymbol{\mathcal{M}} \circ T_{q_{(t)}}\mathbf{L}_{q_{(t)}^{-1}}\,)(\,\dot{q}_{(t)}\,) = $ constant.

All of these are either second-order or first-order systems but with double the number of degrees of freedom.

A direct consequence of the conservation of the quantity S is that geodesic evolution equations can be reduced to a first-order equation exclusively in terms of the current element of the curve. From (4.2), multiplying from the left by Q^T and from the right by $Q_{(t)}^{-T}$,

$$Q_{(t)}^T S Q_{(t)}^{-T} = \boldsymbol{\chi}(Q_{(t)}^{-1}\dot{Q}_{(t)})\,, \tag{5.1}$$

where $Q_{(t)}^{-1}\dot{Q}_{(t)} \in \mathfrak{g}$. Now, it can be noticed that $\overline{\boldsymbol{\sigma}(\boldsymbol{\chi}U)} = \boldsymbol{B}\boldsymbol{B}^{\dagger}\overline{U} = \overline{U}$ for all $U \in \mathfrak{g}$. Then, by applying $\boldsymbol{\sigma}$ to both sides of (5.1) and multiplying from the left by $Q_{(t)}$,

$$\dot{Q}_{(t)} = Q_{(t)}\boldsymbol{\sigma}\left(Q_{(t)}^T\, S\, Q_{(t)}^{-T}\right), \tag{dotQ}$$

yielding a first-order equation for the geodesic evolution without performing the trivial trick of doubling the dimension. The initial velocity of the curve V_0 is encoded in the value of $S = Q_0^{-T}\boldsymbol{\chi}\left(Q_0^{-1}V_0\right)Q_0^T$, particularly when $Q_0 = I$, $S = \boldsymbol{\chi}U_0$.

It is interesting to compare (dotQ) with the symmetric representation given by IVP (3.10). Notice that if $Q_{(t)}$ is an integral curve of the equation $\dot{X} = XU$, then $Q_{(t)}^{-T}$ is an integral curve of $\dot{X} = -XU^T$, which is precisely the equation that guides the evolution of the variable $Y_{(t)}$.

As was done in the previous section, the appearance of the transposed elements can be replaced by[20]

$$\dot{Q}_{(t)} = Q_{(t)}\,\boldsymbol{\sigma}\left((\mathbf{Ad}_{Q_{(t)}})^* S\right)\,,$$

which is an expression intrinsically described by group operations.

Equation (dotQ) is the extension of a similar equation given in [25] which was derived only for Rotation groups, $\mathbf{SO}(d)$. For this particular case the right-hand side turns out to be a cubic polynomial expression in the elements of Q.

6. Sensitivity with respect to initial velocity.

Given initial conditions $Q_0 = I \in \mathbf{G}$ and $\dot{Q}_{(0)} = U_0 \in \mathfrak{g}$, $\mathrm{Exp}_I(U)$ is obtained by integrating (dotQ) (or any of the methodologies described through sections 3.1–3.6) up to $t = 1$. In this section we are interested in computing the Gâteaux derivative of the geodesic $Q_{(t)}$ with respect to U in a direction $W \in \mathfrak{g}$. Let $Q_{(t,\epsilon)}$ be the geodesic starting at I with a perturbed initial velocity $U_0 + \epsilon W$, such that $Q_{(t,0)} = Q_{(t)} = \mathrm{Exp}_I(tU_0)$. The evolution of $Q_{(t,\epsilon)}$ is governed by (dotQ)

$$\dot{Q}_{(t,\epsilon)} = Q_{(t,\epsilon)}\,\boldsymbol{\sigma}\left(Q_{(t,\epsilon)}^T\,\boldsymbol{\chi}(U_0 + \epsilon W)\,Q_{(t,\epsilon)}^{-T}\right). \tag{6.1}$$

Taking derivatives with respect to ϵ on both sides of (6.1) and evaluating at $\epsilon = 0$, we

[20]For abstract groups it can be shown that the corresponding ODE is

$$\dot{q}_{(t)} = \left(T_I\mathbf{L}_{q(t)} \circ \boldsymbol{\mathcal{M}}^{-1} \circ (\mathbf{Ad}_{q(t)})^* \circ (\mathbf{Ad}_{q_0^{-1}})^* \circ \boldsymbol{\mathcal{M}} \circ T_{q_0}\mathbf{L}_{q_0^{-1}}\right)\dot{q}_{(0)}\,.$$

obtain

$$\dot{J}_{(t)} = J_{(t)}\,\boldsymbol{\sigma}\left(Q_{(t)}^T\,\boldsymbol{\chi}(U_0)\,Q_{(t)}^{-T}\right) + Q_{(t)}\,\boldsymbol{\sigma}\left(J_{(t)}^T\boldsymbol{\chi}(U_0)Q_{(t)}^{-T} + Q_{(t)}^T\boldsymbol{\chi}(W)Q_{(t)}^{-T} - Q_{(t)}^T\boldsymbol{\chi}(U_0)Q_{(t)}^{-T}J_{(t)}^T Q_{(t)}^{-T}\right), \tag{6.2}$$

where J is defined as $J_{(t)} = \partial_\epsilon Q_{(t,\epsilon)}\big|_{(\epsilon=0)} \in T_{Q_{(t)}}\mathbf{G}$ and the identity $\partial_\epsilon(Q^{-T}) = -Q^{-T}(\partial_\epsilon Q)^T Q^{-T}$ is used.

Integrating (6.2) with initial condition $J_{(0)} = 0$ results in a vector field along the geodesic which is known as the *Jacobi field* [50, Chapter 5], [160]. In these works the Jacobi field evolution equation is expressed by a second-order equation which takes into account the curvature of the manifold. Similarly to section 5, the second-order equation can be reduced to a first-order equation by using the symmetries of the group and metric.

Likewise the evolution of $Q_{(t)}$ was described in terms of algebra elements (see sections 3.3 and 3.4); it is convenient to describe $J_{(t)}$ as $R_{(t)} = Q_{(t)}^{-1}J_{(t)} \in \mathfrak{g}$. The time evolution of $R_{(t)}$ is

$$\begin{aligned}\dot{R}_{(t)} &= (Q_{(t)}^{-1}J_{(t)})^{\cdot} = -Q_{(t)}^{-1}\dot{Q}_{(t)}Q_{(t)}^{-1}J_{(t)} + Q_{(t)}^{-1}\dot{J}_{(t)}\\ &= -UR + R\,\boldsymbol{\sigma}\left(Q^T\boldsymbol{\chi}(U_0)Q^{-T}\right)\\ &\quad + \boldsymbol{\sigma}\left(J^T\boldsymbol{\chi}(U_0)Q^{-T} + Q^T\boldsymbol{\chi}(W)Q^{-T} - Q^T\boldsymbol{\chi}(U_0)Q^{-T}J^TQ^{-T}\right).\end{aligned} \tag{6.3}$$

From (dotQ), $\boldsymbol{\sigma}(Q^T\boldsymbol{\chi}(U_0)Q^{-T}) = U$, and in consequence $Q^T\boldsymbol{\chi}(U_0)Q^{-T} = \boldsymbol{\chi}U$. Also note that $R^T = J^TQ^{-T}$ and therefore $J^T = R^TQ^T$. Finally, after simple algebraic manipulations, the expression in (6.3) simplifies to[21]

$$\dot{R}_{(t)} = \boldsymbol{\sigma}\left(Q_{(t)}^T\boldsymbol{\chi}(W)Q_{(t)}^{-T}\right) + \boldsymbol{\sigma}[R_{(t)}^T, \boldsymbol{\chi}(U_{(t)})] + [R_{(t)}, U_{(t)}]\,. \tag{6.4}$$

7. Riemannian exponential functions under different left-invariant metrics.

It was noted in [118] that the selection of different left-invariant Riemannian metrics on the same Lie group may give rise to different curvature properties of the manifold and different geodesics. It is possible to find a relationship between a family of left-invariant Riemannian metrics and the family of their corresponding geodesics. Thus, given the Riemannian exponential function for an inertia operator $\boldsymbol{\mathcal{M}}$, it is possible to find a family of Riemannian exponential functions whose inertia operators are given by $\boldsymbol{\mathcal{M}}$ composed with the Adjoint action of the group.

Let $(\mathbf{G}, \langle\!\langle\cdot,\cdot\rangle\!\rangle.)$ be a Riemannian metric which is the left-invariant extension of $\langle\!\langle\cdot,\cdot\rangle\!\rangle_I$ and a geodesic γ with respect to this metric. The velocities of the conjugated curve, defined by $\xi_{(t)} = Y\gamma_{(t)}Y^{-1}$, are given by $\dot{\xi}_{(t)} = \xi_{(t)}\mathbf{Ad}_Y(\gamma_{(t)}^{-1}\dot{\gamma}_{(t)}) = Y\dot{\gamma}_{(t)}Y^{-1}$ (see section 2.4). If a second left-invariant Riemannian metric is chosen such that[22] $\langle\!\langle(\mathbf{Ad}_Y)^{-1}(\cdot), (\mathbf{Ad}_Y)^{-1}(\cdot)\rangle\!\rangle_I$, then it is

[21] In abstract notation,

$$\dot{\mathcal{R}}_{(t)} = \left(\boldsymbol{\mathcal{M}}^{-1}\circ(\mathbf{Ad}_{\mathcal{Q}_{(t)}})^*\circ\boldsymbol{\mathcal{M}}\right)\mathcal{W} + \left(\boldsymbol{\mathcal{M}}^{-1}\circ\mathbf{ad}^*_{\mathcal{R}_{(t)}}\circ\boldsymbol{\mathcal{M}}\right)\mathcal{U}_{(t)} + \mathbf{ad}_{\mathcal{R}_{(t)}}\mathcal{U}_{(t)}.$$

[22] This metric corresponds to the left-translation of the inertia operator given by $\widetilde{\boldsymbol{\mathcal{M}}} = (\mathbf{Ad}_{Y^{-1}})^*\circ\boldsymbol{\mathcal{M}}\circ\mathbf{Ad}_{Y^{-1}}$.

straightforward to verify that the velocities $\dot{\xi}_{(t)}$ are extremes of the Riemannian energy under the new metric, and, therefore, ξ is a geodesic for the second metric.

Given an inertia operator represented by $\boldsymbol{m}$ which defines the inner product $\langle\!\langle U_1, U_2\rangle\!\rangle_I = \overline{U}_1^T \boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger}\overline{U}_2$, and the operation $\overline{\mathbf{Ad}_{Y^{-1}}U} = \overline{Y^{-1}UY} = (Y^T \otimes Y^{-1})\overline{U}$, then the inner product $\langle\!\langle \mathbf{Ad}_{Y^{-1}}U_1, \mathbf{Ad}_{Y^{-1}}U_2\rangle\!\rangle_I = \overline{U}_1^T (Y \otimes Y^{-T})\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger}(Y^T \otimes Y^{-1})\overline{U}_2$. Therefore, if the representation $\tilde{\boldsymbol{m}}$ of an inertia operator can be written as

$$\tilde{\boldsymbol{m}} = \boldsymbol{B}^T (Y \otimes Y^{-T})\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger}(Y^T \otimes Y^{-1})\boldsymbol{B}$$

for some $Y \in \mathbf{G}$, and if there is a known solution of $\mathrm{Exp}_I(\cdot)$ for the metric defined by $\boldsymbol{m}$, then the Riemannian exponential function under the metric defined by $\tilde{\boldsymbol{m}}$ is given by

$$\mathrm{Exp}_I^{\tilde{\boldsymbol{m}}}(U) = Y\,\mathrm{Exp}_I^{\boldsymbol{m}}(Y^{-1}UY)\,Y^{-1}\,. \tag{7.1}$$

In addition, for a metric defined by $\hat{\boldsymbol{m}} = \alpha\,\boldsymbol{m}$, where $\alpha \in \mathbb{R}^+$, $\mathrm{Exp}_I^{\hat{\boldsymbol{m}}}(U) = \mathrm{Exp}_I^{\boldsymbol{m}}(U)$; i.e., the scaling of the inertia operator does not change the Riemannian exponential function, although it does change curve lengths and distances between group elements.

8. Closed-form solutions of $\mathrm{Exp}(\cdot)$ for common transformation groups.

The evolution of geodesics was described and formulated as ODEs in section 5 and throughout section 3. For a few groups, and under particular invariant metrics, geodesics, and therefore the Riemannian exponential function, can be expressed in closed form.[23] To conclude the theoretical aspects of this work we take the opportunity to gather, summarize, and present some new closed-form solutions of the Riemannian exponential function for common matrix groups used in computer vision and imaging applications.

For each of the following transformation models, transformation elements will be written in terms of some parameters $\bar{p}$ such that $\mathscr{h}(\bar{p})$ unambiguously identifies the transformation. The transformation action on a d-dimensional Euclidean point coordinate will be explicitly given. The matrix representation of the group element is denoted as $H(\bar{p})$. The group composition is in some cases expressed in terms of the parameters of its arguments. In any case, the group composition can be obtained via the multiplication of the matrix representation. A faithful matrix representation of bases of the algebra is also given. Finally, a closed-form expression for the Riemannian exponential function from the identity under a specified left-invariant metric is given. Riemannian exponential functions developed around another group element can be computed by using (2.10). For the case of right-invariant Riemannian metrics, (2.11) gives a closed-form expression to make the change from left- to right-invariance. Finally, most of the results given in this section are valid only for some particular inertia tensors, though the conjugation trick explained in section 7 can be used to generalize the closed-form expression to other metrics.

[23]For some groups, the closed-form expression is written in terms of the matrix exponential function. Although there is no global consensus about what a "closed-form" expression means [31], the matrix exponential function, $\exp(\cdot)$, can be computed accurately and efficiently by most scientific software packages like Mathematica and MATLAB [122, 7]. Moreover, pure algebraic and explicit formulas (although complicated) in terms of *elementary operations* can be given for matrices of size smaller than 5 and for some larger structured matrices.

8.1. Translation group, $\mathcal{T}(d)$. The group elements can be parametrized with $\bar{t} = \left(t^1, \ldots, t^d\right)^T$, and the action on a point coordinate $\mathsf{y} \equiv \left(y^1, \ldots, y^d\right)^T \in \mathbb{R}^d$ is $\hbar(\bar{t}) \cdot \mathsf{y} = \mathsf{y} + \bar{t}$. The group composition rule is $\hbar(\bar{t}_2)\hbar(\bar{t}_1) = \hbar(\bar{t}_2 + \bar{t}_1)$.

A matrix representation of group elements and algebra vectors is given, respectively, by $(d+1)\times(d+1)$ matrices of the form

$$H(\bar{t}) = \left(\begin{array}{c|c} I_d & \bar{t} \\ \hline 0_d^T & 1 \end{array}\right) \qquad \text{and} \qquad U(\bar{\tau}) = \left(\begin{array}{c|c} 0_{d\times d} & \bar{\tau} \\ \hline 0_d^T & 0 \end{array}\right),$$

where $\bar{\tau} \in \mathbb{R}^d$.

Finally, the Riemannian exponential under any left-invariant metric is

$$\mathrm{Exp}_I^{\mathcal{T}(d)}\left(U(\bar{\tau})\right) = H(\bar{\tau}).$$

8.2. Isotropic Scale group, $\mathcal{S}^+(1)$. This group is a resizing around the origin of the coordinate system. The superindex plus sign means that inversions are not allowed, preserving the orientation.[24] The group elements are parametrized by a positive scalar number $s \in \mathbb{R}^+$, and its action on coordinate points is $\hbar(s) \cdot \mathsf{y} = s\mathsf{y}$.

The group composition law is $\hbar(s_2)\hbar(s_1) = \hbar(s_2 s_1)$. A matrix representation is given by 1×1 matrices $H(s) = s$, and the algebra vectors are of the form $U(\sigma) = \sigma$, with $\sigma \in \mathbb{R}$.

The $\mathrm{Exp}(\cdot)$ function under any left-invariant metric is

$$\mathrm{Exp}_I^{\mathcal{S}^+(1)}\left(U(\sigma)\right) = H(e^{\sigma}).$$

8.3. Non-Isotropic Scale group, $\mathcal{S}^+(d)$. Asymmetrically resized objects can be obtained by using this transformation model, where a scale factor is used for each dimensional coordinate.

Transformations $\hbar(\bar{s})$ are parametrized with $\bar{s} = \left(s^1, s^2, \ldots, s^d\right)^T$, where $s^i \in \mathbb{R}^+$. The action is $\hbar(\bar{s}) \cdot \mathsf{y} = \left(s^1 y^1, s^2 y^2, \ldots, s^d y^d\right)^T$.

The group composition is $\hbar(\bar{r})\hbar(\bar{s}) = \hbar(\left(r^1 s^1, r^2 s^2, \ldots, r^d s^d\right)^T)$. Matrix representations for group elements and algebra vectors are

$$H(\bar{s}) = \begin{pmatrix} s^1 & 0 & \ldots & 0 \\ 0 & s^2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & s^d \end{pmatrix} \qquad \text{and} \qquad U(\bar{\sigma}) = \begin{pmatrix} \sigma^1 & 0 & \ldots & 0 \\ 0 & \sigma^2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma^d \end{pmatrix},$$

where $\sigma^i \in \mathbb{R}$.

The $\mathrm{Exp}_I^{\mathcal{S}^+(d)}$ function, for any left-invariant metric, is

$$\mathrm{Exp}_I^{\mathcal{S}^+(d)}\left(U\left(\left(\sigma^1, \sigma^2, \ldots, \sigma^d\right)^T\right)\right) = H\left(\left(e^{\sigma^1}, e^{\sigma^2}, \ldots, e^{\sigma^d}\right)^T\right).$$

[24] "Orientation" refers to the handedness of a frame.

Remark. The three previous groups are commutative, and therefore any left-invariant metric is also right-invariant, and the corresponding Riemannian exponential function $\mathrm{Exp}_I(\cdot)$ coincides with the group exponential function $\exp(\cdot)$ (see section 2.8).

8.4. Isotropic Scale + Translation group, $\mathcal{ST}(d)$. Under this group action the objects are isotropically resized around the origin of the coordinate system and then translated. The $\mathcal{ST}(1)$ group is a simple case of a noncompact and noncommutative Lie group where bi-invariant metrics cannot be defined [14].

A parametrization of the group elements is $\hbar(s,\bar{t})$, where $s \in \mathbb{R}^+$ and $\bar{t} \in \mathbb{R}^d$. The action on a coordinate point is $\hbar(s,\bar{t}) \cdot \mathsf{y} = s\mathsf{y} + \bar{t}$.

The composition law is given by $\hbar(r,\bar{w})\hbar(s,\bar{t}) = \hbar(rs, \bar{w} + r\bar{t})$.

Matrix representations are

$$H(s,\bar{t}) = \left(\begin{array}{c|c} sI_d & \bar{t} \\ \hline 0_d^T & 1 \end{array}\right) \qquad \text{and} \qquad U(\sigma,\bar{\tau}) = \left(\begin{array}{c|c} \sigma I_d & \bar{\tau} \\ \hline 0_d^T & 0 \end{array}\right)$$

with $\sigma \in \mathbb{R}$ and $\bar{\tau} \in \mathbb{R}^d$.

The following set of $(d+1)\times(d+1)$ matrices is used as bases of the algebra:

$$\left\{ \left(\begin{array}{c|c} I_d & 0_d \\ \hline 0_d^T & 0 \end{array}\right), \left(\begin{array}{c|c} 0_{d\times d} & \mathrm{e}_1 \\ \hline 0_d^T & 0 \end{array}\right), \dots, \left(\begin{array}{c|c} 0_{d\times d} & \mathrm{e}_d \\ \hline 0_d^T & 0 \end{array}\right) \right\},$$

where e_j is the jth canonical basis of $\mathbb{R}^d$.

With these bases and for $\boldsymbol{m} = I_{d+1}$, the left-invariant Riemannian exponential has the following closed-form expression:

$$\mathrm{Exp}_I^{\mathcal{ST}(d)}\left(U(\sigma,\bar{\tau})\right) = \begin{cases} H\left(\frac{(b^2+1)e^w}{b^2e^{2w}+1}, \frac{b(e^{2w}-1)}{b^2e^{2w}+1}\frac{\bar{\tau}}{\|\bar{\tau}\|}\right) & \text{if } \|\bar{\tau}\| \neq 0, \\ H\left(e^\sigma, 0\right) & \text{else,} \end{cases} \tag{8.1}$$

where $\|\cdot\|$ is the d-Euclidean norm, $w = \sqrt{\sigma^2 + \|\bar{\tau}\|^2}$, and $b = (w-\sigma)/\|\bar{\tau}\|$.

The expressions in (8.1) can be derived by noticing that the $\mathcal{ST}(1)$ group with a left-invariant metric is isomorphic to the Poincaré half-plane [14, 168, 161]. These can be derived from the IVP obtained in Appendix A after involved algebraic work.

Note that the closed-form in (8.1) is valid for any spatial dimension d, while in most of the literature only the one-dimensional case is found.

8.5. Rotation group, $\mathcal{SO}(d)$. Group elements are linear isometries that are orientation preserving and keep the point at the origin fixed. Many parametrizations have been proposed to describe rotations, especially in $\mathbb{R}^3$, such as quaternions [157, 75, 171], Euler angle conventions [45], axis-angle representation [124, 68], and also for higher dimensions [63]. Let $\hbar(\bar{\theta})$ be a parametrization; then the action is given by $\hbar(\bar{\theta}) \cdot \mathsf{y} = R_\theta \mathsf{y}$, where R_θ is a $d\times d$ matrix fulfilling $R_\theta R_\theta^T = R_\theta^T R_\theta = I_d$ and $\det(R_\theta) = 1$.

The matrix representation of an element of the algebra $\mathfrak{so}(d)$ can be obtained by noticing that $\dot{R}_\theta R_\theta^T + R_\theta \dot{R}_\theta^T = 0$, and particularizing to the identity, it is obtained that $\dot{R} = -\dot{R}^T$.

Therefore, the algebra is composed by the set of skew-symmetric $d{\times}d$ matrices which is a $d(d-1)/2$ dimensional space [119, 63, 76]

$$U(\bar{\omega}) = \Sigma(\bar{\omega}) = \Sigma\left(\left(\omega^1, \omega^2, \dots, \omega^{d(d-1)/2}\right)^T\right),$$

where $\Sigma(\bar{\omega})$ is a $d{\times}d$ skew-symmetric matrix built by rearranging the entries of the vector $\bar{\omega}$.

It is well known that the $\mathcal{SO}(d)$ group admits a bi-invariant metric [14, 132, 76]. It corresponds to the inertia operator represented by the identity matrix I_k when the algebra bases are given by $\{\Sigma(\mathrm{e}_i)\}_{i=1}^{d(d-1)/2}$. Therefore, the inner product in the algebra results in $\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \operatorname{trace}(U_1^T U_2)/2$, and the corresponding Riemannian exponential is given by the group exponential,

$$\operatorname{Exp}_I^{\mathcal{SO}(d)}(\,\Sigma(\bar{\omega})\,) = \exp(\,\Sigma(\bar{\omega})\,)\,.$$

8.6. Special Euclidean group (a.k.a. Rigid Body Transformation group), $\mathcal{SE}(d)$. This group is composed of rotations and translations. It is the most general transformation model which preserves the Euclidean distance between all pairs of points.

A parametrization of a group element is $\hbar(\bar{\theta}, \bar{t})$, where $\bar{\theta}$ is a parametrization of rotation and $\bar{t} \in \mathbb{R}^d$ is the translation parameter. The action on an $\mathbb{R}^d$ point is $\hbar(\bar{\theta}, \bar{t}) \cdot \mathsf{y} = R_\theta \mathsf{y} + \bar{t}$.

Matrix representations of elements and algebra vectors are

$$H(\bar{\theta}, \bar{t}) = \left(\begin{array}{c|c} R_\theta & \bar{t} \\ \hline 0_d^T & 1 \end{array}\right) \qquad \text{and} \qquad U(\bar{\omega}, \bar{\tau}) = \left(\begin{array}{c|c} \Sigma(\bar{\omega}) & \bar{\tau} \\ \hline 0_d^T & 0 \end{array}\right),$$

where $\Sigma(\bar{\omega})$ is a skew-symmetric $d{\times}d$ matrix with entries ω^i, analogously to the case of $\mathcal{SO}(d)$ group.

For a left-invariant metric where $\boldsymbol{m}$ is the identity, the Riemannian exponential function results in [129]

$$\operatorname{Exp}_I^{\mathcal{SE}(d)}(\,U(\bar{\omega}, \bar{\tau})\,) = \left(\begin{array}{c|c} \exp(\,\Sigma(\bar{\omega})\,) & \bar{\tau} \\ \hline 0_d^T & 1 \end{array}\right).$$

8.7. Similarity group, $\mathcal{SIM}(d)$. This group is a combination of rotations, translations, and isotropic scalings.[25]

Group elements are parametrized by $\hbar(s, \bar{\theta}, \bar{t})$, with $s \in \mathbb{R}^+$ a scaling parameter, $\bar{\theta}$ a parametrization of a $d{\times}d$ rotation, and $\bar{t} \in \mathbb{R}^d$ a translation parameter.

The action is computed by $\hbar(s, \bar{\theta}, \bar{t}) \cdot \mathsf{y} = sR_\theta \mathsf{y} + \bar{t}$.

A matrix representation is

$$H(s, \bar{\theta}, \bar{t}) = \left(\begin{array}{c|c} sR_\theta & \bar{t} \\ \hline 0 & 1 \end{array}\right) \qquad \text{and} \qquad U(\sigma, \bar{\omega}, \bar{\tau}) = \left(\begin{array}{c|c} \sigma I_d + \Sigma(\bar{\omega}) & \bar{\tau} \\ \hline 0 & 0 \end{array}\right).$$

A set of bases of the algebra is given by

$$\left(\begin{array}{c|c} I_d & 0_d \\ \hline 0_d^T & 0 \end{array}\right) \bigcup \left\{ \left(\begin{array}{c|c} \Sigma(\mathrm{e}_i) & 0_d \\ \hline 0_d^T & 0 \end{array}\right) \right\}_{i=1}^{d(d-1)/2} \bigcup \left\{ \left(\begin{array}{c|c} 0_{d\times d} & \mathrm{e}_j \\ \hline 0_d^T & 0 \end{array}\right) \right\}_{j=1}^{d}.$$

[25] In other works this group is called conformal [53, 173] or similitude [70]. Sometimes it is defined allowing orientation changes, but here it will be defined only for positive scales.

Let $\phi\colon \mathcal{SIM}(d) \to \mathcal{SO}(d) \times \mathcal{ST}(d)$ be a smooth map, where $\times$ denotes the direct product of the group sets [34]

$$\phi\left(\hbar^{\mathcal{SIM}}(s,\bar{\theta},\bar{t})\right) = \left(\hbar^{\mathcal{SO}}(\bar{\theta})\,,\, \hbar^{\mathcal{ST}}(s,\bar{t})\right)\,.$$

Under a block-scalar metric, i.e., a metric under which the contributions from both spaces do not interact, the mapping ϕ results in an isometry [64]; thus under this metric, geodesics on $\mathcal{SIM}(d)$ are obtained by lifting geodesics on $\mathcal{SO}(d)$ and $\mathcal{ST}(d)$, respectively:

$$\begin{aligned}\mathrm{Exp}_I^{\mathcal{SIM}(d)}&\left(\left(\begin{array}{c|c}\sigma I_d + \Sigma(\bar{\omega}) & \bar{\tau}\\ \hline 0_d^T & 0\end{array}\right)\right)\\ &= \phi^{-1}\left(\left(\mathrm{Exp}_I^{\mathcal{SO}(d)}\left(\Sigma(\bar{w})\right)\,,\, \mathrm{Exp}_I^{\mathcal{ST}(d)}\left(\left(\begin{array}{c|c}\sigma I_d & \bar{\tau}\\ \hline 0_d^T & 0\end{array}\right)\right)\right)\right)\,.\end{aligned}$$

8.8. Orientation Preserving General Linear group, $\mathcal{GL}^+(d)$. These transformations preserve the parallelism between lines and the point at the origin. A transformation can be represented by a matrix $M_p \in \mathbb{M}_d$ with positive determinant. Its action on a coordinate point $\mathsf{y} \in \mathbb{R}^d$ is given by $\hbar(\bar{p}) \cdot \mathsf{y} = M_p\mathsf{y}$, and a parametrization is given by the matrix elements themselves ($\bar{p} = \overline{M}_p \in \mathbb{R}^{d^2}$).

The algebra $\mathfrak{gl}(d)$ is the set of all $d{\times}d$ real matrices $\mathbb{M}_d$. For the canonical bases of $\mathbb{M}_d$ and under a left-invariant metric defined by $\boldsymbol{m} = I_{d^2}$, a closed-form solution of the Riemannian exponential function was derived in [10, 170]:

$$\mathrm{Exp}_I^{\mathcal{GL}^+(d)}(\,U\,) = \exp(U^T)\exp(U - U^T)\,. \tag{8.2}$$

The corresponding inner product is $\langle\!\langle U_1, U_2\rangle\!\rangle_I = \mathrm{trace}(U_1^T U_2)$.

Note that the group exponential $\exp(\cdot)$ is not surjective (section 2.5), while the Riemannian exponential $\mathrm{Exp}(\cdot)$ is (section 2.7). There is no contradiction between these statements and (8.2), where a composition of two different group exponentials appears at the right-hand side. This is in agreement with the results of [175], where it is proved that the composition of two group exponentials is enough to fully cover the identity component of any group.

8.9. Special Linear group, $\mathcal{SL}(d)$. This group consists of the subset of transformations in $\mathcal{GL}^+(d)$ which preserve the volume of the objects.

A matrix representation of an element is given by a matrix M_p with determinant equal to 1. An algebra representation consists of the set of $d{\times}d$ traceless matrices [164, 104] which is a $d^2 - 1$ dimensional space.

Although in [104] a closed-form expression was derived from scratch, the same result can be obtained using the concept of a *totally geodesic subgroup*. Given a group equipped with a Riemannian metric $(\mathbf{G}, \langle\!\langle\cdot,\cdot\rangle\!\rangle.)$ and a proper subgroup $\mathbf{N}$ with the induced metric, $\mathbf{N}$ is called a *totally geodesic subgroup* of $\mathbf{G}$ with respect to $\langle\!\langle\cdot,\cdot\rangle\!\rangle.$ if each geodesic of $\mathbf{N}$ is also a geodesic of $\mathbf{G}$. Sufficient and necessary conditions to be a totally geodesic subgroup with respect to a left-invariant Riemannian metric are given in [121]. It can be shown that $\mathcal{SL}(d)$ is a totally geodesic subgroup of $\mathcal{GL}^+(d)$ under the metric defined by the inner product $\langle\!\langle U_1, U_2\rangle\!\rangle_I = \mathrm{trace}(U_1^T U_2)$. Therefore, (8.2) is also a closed-form solution valid for $\mathcal{SL}(d)$.

Although it is beyond the scope of this paper, it is interesting to note that $\mathcal{SL}(d)$ is a *semisimple group*, and therefore its algebra admits a *Cartan decomposition* into skew-symmetric matrices and traceless-symmetric matrices [80, Chapter IV]. It is shown in [80, p. 277], [172] that, under the particular left-invariant metric given by minus the product of the Killing form and the Cartan involution, the Riemannian exponential function has the same closed-form solution. Moreover, expressing $\mathcal{GL}^+(d)$ as the direct product $\mathcal{SL}(d) \times \mathcal{S}^+(1)$, the closed-form of $\mathrm{Exp}_I^{\mathcal{GL}^+(d)}$ in (8.2) can be easily obtained.

8.10. Centered Transformation group, $(\mathcal{G}(d) \times \mathcal{T}(d))$. In some of the previous transformation models, the action includes translations. There is an alternative characterization of these transformations where the objects are first transformed with respect to a center defined on the object (instead of about the origin of the coordinate system), and then the object and its center are translated together. With this alternative formulation, the transformations act on a pair (y, c) composed of a point with a defined center

$$\hbar^{(\mathcal{G}\times\mathcal{T})}(\bar{p},\bar{t}) \cdot (\mathsf{y},\mathsf{c}) = \left(\hbar^{\mathcal{G}}(\bar{p}) \cdot (\mathsf{y}-\mathsf{c}) + (\mathsf{c}+\bar{t})\,,\, \mathsf{c}+\bar{t} \right) = \left(\mathsf{y}', \mathsf{c}'\right) ,$$

where y' and c' are the transformed positions of the point and the center, respectively. The group composition is $\hbar(\bar{p}_2,\bar{t}_2)\hbar(\bar{p}_1,\bar{t}_1) = \hbar(g(\bar{p}_2,\bar{p}_1), \bar{t}_2+\bar{t}_1)$, where the function $g(\cdot,\cdot)$ is the corresponding composition law of the $\mathcal{G}$ group, and therefore the proposed group is a direct product of $\mathcal{G}$ and the Translation group $\mathcal{T}(d)$.

Possible matrix representations of these group elements and algebra vectors are

$$H(\bar{p},\bar{t}) = \left(\begin{array}{c|c|c} H^{\mathcal{G}}(\bar{p}) & 0_{n\times d} & 0_n \\ \hline 0_{d\times n} & I_d & \bar{t} \\ \hline 0_n^T & 0_d^T & 1 \end{array}\right) \quad \text{and} \quad U(\bar{\rho},\bar{\tau}) = \left(\begin{array}{c|c|c} U^{\mathcal{G}}(\bar{\rho}) & 0_{n\times d} & 0_n \\ \hline 0_{d\times n} & 0_{d\times d} & \bar{\tau} \\ \hline 0_n^T & 0_d^T & 0 \end{array}\right) .$$

The algebra bases are the direct sum $\mathfrak{g}(d) \oplus \mathfrak{t}(d)$.

For a left-invariant metric defined by a block structured matrix

$$\boldsymbol{m} = \left(\begin{array}{c|c} \boldsymbol{m}^{\mathcal{G}} & 0_{k\times d} \\ \hline 0_{d\times k} & \boldsymbol{m}^{\mathcal{T}} \end{array}\right) ,$$

the Riemannian exponential results in

$$\mathrm{Exp}_I^{(\mathcal{G}\times\mathcal{T})}\left(U(\bar{\rho},\bar{\tau})\right) = \left(\begin{array}{c|c|c} \mathrm{Exp}_I^{\mathcal{G}}(U^{\mathcal{G}}(\bar{\rho})) & 0_{n\times d} & 0_n \\ \hline 0_{d\times n} & I_d & \bar{\tau} \\ \hline 0_n^T & 0_d^T & 1 \end{array}\right) .$$

8.11. Projective group (a.k.a. Homography group), $\mathcal{PGL}(d)$. These transformations map straight lines into straight lines but without preserving either angles or parallelism.

A matrix representation of the group elements is attained considering the action on a *projective space*. In a projective space every point of $\mathbb{R}^d$ is identified with a real line passing through the origin of a $(d+1)$-dimensional space. The mappings ψ and ψ^{-1} allow us to identify $\mathsf{y} \in \mathbb{R}^d$ with its corresponding line in $\mathbb{R}^{d+1}$ and vice-versa:

$$\psi(\mathsf{y}) = \left(\mathsf{y}^T, 1\right)^T ,$$
$$\psi^{-1}(\bar{z}) = \frac{\left(z^1, \ldots, z^d\right)^T}{z^{d+1}} .$$

Projective transformations act linearly on $\mathbb{R}^{d+1}$. If the denominator in the last expression vanishes, the corresponding $\mathbb{R}^d$ points are not well defined, and these points are known as *ideals* [77]. Technically, this group does not generate bijections for points of $\mathbb{R}^d$, but it is well defined for lines through the origin in $\mathbb{R}^{d+1}$. Due to the equivalence relation $\psi^{-1}(\alpha z) = \psi^{-1}(z)$ for any nonzero scalar α, linear transformations related by a scalar factor are identified. Therefore, the transformation group $\mathcal{PGL}(d+1)$ is defined as the quotient $\mathcal{GL}(d+1)/(\mathbb{R}-\{0\})$. Every transformation in a neighborhood of the identity can be identified with a matrix of $\mathbf{GL}(d+1)$ with determinant equal to 1, and therefore the group becomes, at least locally, isomorphic to the $\mathcal{SL}(d+1)$ group.

Matrix representations of the group elements are the $(d+1)\times(d+1)$ matrices of determinant 1, and matrix representations of the algebra vectors are $(d+1)\times(d+1)$ matrices with trace equal to 0, resulting in a $(d+1)^2 - 1$ dimensional space.

The action on $\mathbb{R}^d$ points, when it is defined, can be written as

$$\hbar(\bar{p}) \cdot \mathsf{y} = \psi^{-1}\left(H(\bar{p})\, \psi(\mathsf{y}) \right) ,$$

where $\bar{p}$ are the parameters of a matrix with determinant equal to 1.

Finally, under the left-invariant metric generated by $\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \mathrm{trace}(U_1^T U_2)$, the Riemannian exponential function can be computed via (8.2).

8.12. Möbius Transformation group, $\mathcal{M}(d)$.

This group includes transformations of the Euclidean space that locally preserve angles. Möbius transformations are obtained by composition of translations, rotations, dilations, and inversions with respect to the unit sphere ($\mathbf{i}\colon \mathsf{y} \mapsto \mathsf{y}/\|\mathsf{y}\|^2$). Note that an even number of inversions must be used in order to preserve the orientation. A parametrization and its action are given by $\hbar(s, \bar{\theta}, \bar{t}, \bar{w}) \cdot \mathsf{y} = s R_{\bar{\theta}}\, \mathbf{i}(\mathbf{i}(\mathsf{y}) + \bar{w}) + \bar{t}$.

In order to obtain a matrix representation of the transformation element, a mapping to a hyperbolic space can be done: let $\bar{z} \in \mathbb{R}^{d+1,1}$ with a norm defined by $\|\bar{z}\|^2_{hyp} = \sum_i^{d+1} (z^i)^2 - (z^{d+2})^2$, and let $\psi\colon \mathbb{R}^d \to \mathbb{R}^{d+1,1}$ be the mapping given by

$$\psi(\mathsf{y}) = \left(\mathsf{y}^T, \frac{\|\mathsf{y}\|^2 - 1}{2}, -\frac{\|\mathsf{y}\|^2 + 1}{2} \right)^T ,$$
$$\psi^{-1}(\bar{z}) = \frac{\left(z^1, \ldots, z^d\right)^T}{-z^{d+1} - z^{d+2}} .$$

Note that $\|\psi(\mathsf{y})\|^2_{hyp} = 0$ and $\psi^{-1}(\alpha\bar{z}) = \psi^{-1}(\bar{z})$ for any nonzero scalar α. With the mapping ψ, conformal transformations in $\mathbb{R}^d$ consist of isometries in $\mathbb{R}^{d+1,1}$ modulo $\{\pm\mathcal{ID}(d+2)\}$, which is the projective orthogonal group $\mathcal{PO}(d+1,1)$ [64]. This group is locally isomorphic to the indefinite special orthogonal group $\mathcal{SO}(d+1,1)$. Specifically, matrix representations of the elementary transformations that make up the Möbius group are H_D, H_R, H_T, and $H_{\mathbf{i}}$, corresponding to dilations, rotations, translations, and the inversion, respectively:

$$H_D(s) = \left(\begin{array}{c|cc} I_d & 0_d & 0_d \\ \hline 0_d^T & \frac{1}{2}(\frac{1}{s}+s) & \frac{1}{2}(\frac{1}{s}-s) \\ 0_d^T & \frac{1}{2}(\frac{1}{s}-s) & \frac{1}{2}(\frac{1}{s}+s) \end{array}\right), \quad H_R(\bar{\theta}) = \left(\begin{array}{c|cc} R_\theta & 0_d & 0_d \\ \hline 0_d^T & 1 & 0 \\ 0_d^T & 0 & 1 \end{array}\right),$$

$$H_T(\bar{t}) = \left(\begin{array}{c|cc} I_d & -\bar{t} & -\bar{t} \\ \hline \bar{t}^T & 1-\frac{\|\bar{t}\|^2}{2} & -\frac{\|\bar{t}\|^2}{2} \\ -\bar{t}^T & \frac{\|\bar{t}\|^2}{2} & 1+\frac{\|\bar{t}\|^2}{2} \end{array}\right), \qquad H_{\mathbf{i}} = \left(\begin{array}{c|cc} I_d & 0_d & 0_d \\ \hline 0_d^T & -1 & 0 \\ 0_d^T & 0 & 1 \end{array}\right),$$

where $s \in \mathbb{R}^+$, $\bar{\theta}$ is a parametrization of a rotation matrix, and $\bar{t} \in \mathbb{R}^d$.

In terms of the previous elementary transformations, a matrix representation of $\hbar(s, \bar{\theta}, \bar{t}, \bar{w})$ is given by

$$H(s, \bar{\theta}, \bar{t}, \bar{w}) = H_T(\bar{t})\, H_D(s)\, H_R(\bar{\theta})\, H_{\mathbf{i}}\, H_T(\bar{w})\, H_{\mathbf{i}}\,.$$

As in the case of the projective transformations, the action of a Möbius transformation on a coordinate point of $\mathbb{R}^d$ is $\hbar(s, \bar{\theta}, \bar{t}, \bar{w}) \cdot \mathsf{y} = \psi^{-1}(H\,\psi(\mathsf{y}))$ when it is well defined and no vanishing denominator appears.

The following set the of $(d+2)\times(d+2)$ matrices can be used as a representation of the bases of the algebra:

$$\left(\begin{array}{c|cc} 0_{d\times d} & 0_d & 0_d \\ \hline 0_d^T & 0 & -1 \\ 0_d^T & -1 & 0 \end{array}\right) \bigcup \left\{ \left(\begin{array}{c|cc} \Sigma(\mathrm{e}_i) & 0_d & 0_d \\ \hline 0_d^T & 0 & 0 \\ 0_d^T & 0 & 0 \end{array}\right) \right\}_{i=1}^{d(d-1)/2} \bigcup \cdots$$

$$\cdots \left\{ \left(\begin{array}{c|cc} 0_{d\times d} & -\mathrm{e}_j & -\mathrm{e}_j \\ \hline \mathrm{e}_j^T & 0 & 0 \\ -\mathrm{e}_j^T & 0 & 0 \end{array}\right) \right\}_{j=1}^{d} \bigcup \left\{ \left(\begin{array}{c|cc} 0_{d\times d} & \mathrm{e}_j & -\mathrm{e}_j \\ \hline -\mathrm{e}_j^T & 0 & 0 \\ -\mathrm{e}_j^T & 0 & 0 \end{array}\right) \right\}_{j=1}^{d}.$$

It can be proved that with the above representation, the $\mathcal{M}(d)$ group is a totally geodesic subgroup of the $\mathcal{GL}^+(d+2)$ group under the left-invariant metric given by $\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \operatorname{trace}(U_1^T U_2)$. Therefore, for this metric, the Riemannian exponential function can be computed explicitly using (8.2).

Commonly used transformation groups without known closed form solution for $\mathrm{Exp}(\cdot)$. The following well-known transformation models are commonly used in computer vision and image analysis applications (for example, they are very useful for medical image registration). To the best of our knowledge, there is no known closed-form solution for the $\mathrm{Exp}(\cdot)$ function under any left-invariant Riemannian metric for these groups (except for the trivial case $d = 1$). However, their action and matrix representations are included in the present list for completeness.

8.13. Orientation Preserving General Affine group, $\mathcal{GA}^+(d)$. This group is composed of transformations in the $\mathcal{GL}^+(d)$ group and translations. It is the most general transformation model preserving collineation, parallelism, and orientation.

The transformations in this group are the composition of a general linear transformation followed by a translation. The action is $\hbar(\bar{p},\bar{t})\cdot \mathsf{y} = M_p\mathsf{y} + \bar{t}$, where $\bar{p}$ is a parametrization of the general linear matrix M_p (for example, $\bar{p} = \overline{M}_p$) and $\bar{t} \in \mathbb{R}^d$. The composition law is $\hbar(\bar{p}_2,\bar{t}_2)\hbar(\bar{p}_1,\bar{t}_1) = \hbar(\overline{M_{p_2}M_{p_1}}, M_{p_2}\bar{t}_1 + \bar{t}_2)$.

A matrix representation of the group elements is

$$H(\bar{p},\bar{t}) = \left(\begin{array}{c|c} M_p & \bar{t} \\ \hline 0_d^T & 1 \end{array}\right) .$$

A matrix representation of the algebra vectors is obtained by the $(d+1)\times(d+1)$ matrices with the last row equal to zero.

8.14. Special Affine group, $\mathcal{SA}(d)$. This group is the subset of transformations in the $\mathcal{GA}^+(d)$ group that also preserve the volume of the objects.

The parametrization must be done with a parametrization M_p of a matrix of determinant equal to 1.

A matrix representation of the algebra is given by the set of $(d+1)\times(d+1)$ matrices with the last row equal to zero and trace equal to zero.

9. Application examples. A few examples are given in this section to illustrate some applications of invariant geodesics. Before that, some details are given about simple numerical procedures for distance computation and for the integration of the differential equations involved in the geodesic evolution.

Explicit forward Lie–Euler method. There is a large number of numerical integration methods for ODEs and IVPs [40, 39, 116, 73, 85, 89]. Each numerical procedure has different levels of accuracy and complexity. Furthermore, some of them are specifically designed to preserve some properties of dynamical systems such as the conserved quantities, the group structure, and the symplectic characteristics of the Hamiltonian.

This paper does not aim to explore large families of numerical integrators; only the explicit forward Lie–Euler method is considered as an illustrative case to integrate (dotQ).

While the forward Euler integration method in a vector space involves additive updates, multiplicative updates should be used when integrating on groups to preserve the group structure. The evolution of the geodesic is given by $\dot{Q}_{(t)} = Q_{(t)}U_{(t)}$, where $U_{(t)}$ is given by $\boldsymbol{\sigma}(Q_{(t)}^T S Q_{(t)}^{-T})$ in (dotQ). If it is assumed that U remains constant along each update step $\Delta = t_{j+1} - t_j$, then $\dot{Q} = QU$ can be solved exactly at t_{j+1} by

$$\begin{aligned} Q_{(t_{j+1})} \approx Q_{j+1} &= Q_j\, e^{\Delta\, U_j} \\ &= Q_j\, e^{\Delta\, \boldsymbol{\sigma}\left(\, Q_j^T\, S\, Q_j^{-T}\,\right)} . \end{aligned} \tag{9.1}$$

With this scheme, the group element is updated by the group composition, and therefore Q_{j+1} remains within the group set.

Summarizing, let $\boldsymbol{B} = \left(\overline{B}_1, \ldots, \overline{B}_k\right)$ be a matrix containing the vectorization of the k bases of the algebra $\mathfrak{g}$, and let $\boldsymbol{m}$ be a $k{\times}k$ matrix representing the inertia tensor at the identity. Let $\boldsymbol{\sigma}$ be the linear operator represented by the matrix $(\boldsymbol{B}\boldsymbol{m}^{-1}\boldsymbol{B}^T)$ and $\boldsymbol{\chi}$ its inverse represented by the matrix $(\boldsymbol{B}^{\dagger T}\boldsymbol{m}\boldsymbol{B}^{\dagger})$. Given $Q_0 \in \mathbf{G}$ and $V_0 \in T_{Q_0}\mathbf{G}$, arguments of the desired Riemannian exponential function, compute $S = Q_0^{-T}\boldsymbol{\chi}(Q_0^{-1}V_0)Q_0^T$. Finally, $\mathrm{Exp}_{Q_0}(V_0)$ can be computed by iterating (9.1) with a (sufficiently small) step Δ up to time $t = 1$.

Computing the Riemannian logarithm function. The Riemannian exponential function defines geodesics in terms of their departing elements and velocities. However, in several applications it is desired to specify geodesic segments in terms of their departing and arrival elements (see section 2.6). For this purpose the computation of the Riemannian logarithm is a key tool. The Riemannian logarithm is defined as an inverse function. This problem can be considered as a *two-point boundary value problem* (TPBVP) [144, 88] instead of an IVP. There are different approaches to solving this kind of problem, such as *shooting* or *relaxation* methods.

The shooting method starts with an initial guess of the initial velocity, integrates the geodesic up to time 1, and quantifies the discrepancy or error between the actual arrival point and the desired target. The current discrepancy is used to smartly update the initial velocity, and the process continues until the error vanishes. The process can also be described as a root-finding problem.

Given A and B, the initial and final elements to be joined by a geodesic segment, $\mathrm{Log}_A(B)$ is formulated as the inverse problem of finding the initial velocity V such that $\mathrm{Exp}_A(V) = B$. The following optimization problem is proposed to solve the inverse problem:

$$V^* = \operatorname*{argmin}_V \left\|\mathrm{Exp}_A(V) - B\right\|_F^2 ,$$

where the *Frobenius matrix norm* $\|\cdot\|_F$ was selected for simplicity.

Due to the metric invariance and using (2.10), the problem can be equivalently formulated as

$$V^* = A\left(\operatorname*{argmin}_U \left\|\mathrm{Exp}_I(U) - A^{-1}B\right\|_F^2\right) . \tag{9.2}$$

For most of the cases enumerated in section 8 the $\mathrm{Exp}(\cdot)$ function can be algebraically inverted, and there exist closed-form expressions for $\mathrm{Log}(\cdot)$. Nevertheless, the algebraic expressions corresponding to groups such as $\mathcal{GL}^+(d)$, $\mathcal{SL}^+(d)$ cannot be algebraically inverted [178]. For these cases, and those problems where $\mathrm{Exp}(\cdot)$ have to be computed numerically, a numerical procedure must be employed to compute $\mathrm{Log}(\cdot)$.

A gradient descent strategy is proposed in order to solve (9.2). The corresponding descent direction can be computed using $\partial_X \|X - A^{-1}B\|_F^2|_{(X=Q)} = 2(Q - A^{-1}B)$, and the derivatives of the discrepancy functional with respect to the components μ^i of U are

$$\partial_{\mu^i}\left\|\mathrm{Exp}_I(U) - A^{-1}B\right\|_F^2 = 2\,\mathrm{trace}\left(\left(\mathrm{Exp}_I(U) - A^{-1}B\right)^T \mathrm{Exp}_I(U) R^{B_i}_{(1)}\right) ,$$

where $R^{B_i}_{(1)}$ is the solution of (6.4) at $t = 1$ with $W = B_i$ and initial condition $R_{(0)} = 0$. Alternatively, a finite differences scheme on $\left\|\mathrm{Exp}_I(U) - A^{-1}B\right\|^2_F$ can also be used to compute the descent direction [144, 48]. Note that the derivatives have to be computed along perturbations of the components of U in the algebra $\mathfrak{g}$:

$$\partial_{\mu^i}\left\|\mathrm{Exp}_I(U) - A^{-1}B\right\|^2_F \approx \frac{\left\|\mathrm{Exp}_I(U+\varepsilon B_i) - A^{-1}B\right\|^2_F - \left\|\mathrm{Exp}_I(U-\varepsilon B_i) - A^{-1}B\right\|^2_F}{2\varepsilon}$$

for an appropriate small ε. Once a descent direction is computed, a simple line search procedure is performed to update U.

A limitation of the descent strategy is that for some groups the stationary point reached after convergence may be not a global optimum, i.e., $\mathrm{Exp}_I(U) \neq A^{-1}B$, or even being a global optimum, it may be not the smallest initial velocity which satisfies $U = \mathrm{Log}_I(A^{-1}B)$ [178].

Whenever a right-invariant metric is used, the Riemannian logarithm can be computed from its left-invariant version as

$$\mathrm{Log}^{right}_X(Y) = -\,\mathrm{Log}^{left}_I(XY^{-1})\,X\,.$$

9.1. Interpolating spatial transformations. In this example the geodesic framework is used to interpolate between two transformations or between two transformed versions of an object. The methodology provides a time-parametrized path between two transformations which can be used, for example, in computer graphics for key-framing animation [157, 9, 106, 21].

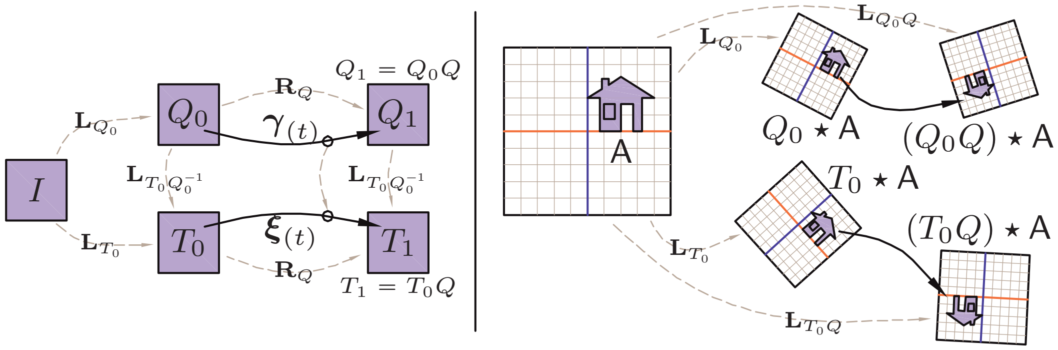
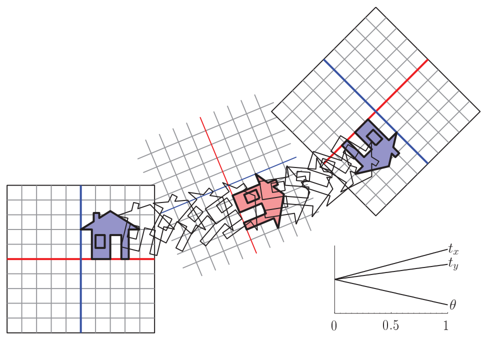


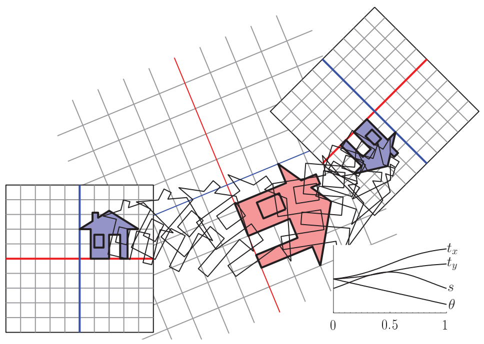
Figure 3. *Illustration of the framework to interpolate spatial transformations.*

The framework is depicted in Figure 3. In the left panel we illustrate the problem of finding the geodesic γ that interpolates two spatial transformations Q_0 and $Q_1 = \mathbf{R}_Q(Q_0) = Q_0Q$. The geodesic interpolating path is given by $\gamma_{(t)} = \mathrm{Exp}_{Q_0}(t\,\mathrm{Log}_{Q_0}(Q_1))$. Due to the left-invariance, the geodesic between any pair of left-translated instances $T_0 = \mathbf{L}_{T_0Q_0^{-1}}(Q_0)$ and $T_1 = \mathbf{L}_{T_0Q_0^{-1}}(Q_1)$ will be $\xi_{(t)} = \mathbf{L}_{T_0Q_0^{-1}}(\gamma_{(t)})$. In this framework the following distances are also preserved: $\mathbf{distance}(Q_0, Q_1) = \mathbf{distance}(I, Q) = \mathbf{distance}(T_0, T_1)$. The right panel shows the transformation actions on an object A defined in the Euclidean space $\mathbb{R}^2$. To emphasize that the spatial transformations act over the whole ambient space, a grid is also shown in the figure. Note that the interpolation between transformations applies to the ambient space independently of the object. However, different objects may define different inertia tensors to take their geometries into account. In addition, using the presented framework, a distance between the objects $Q_0 \cdot \mathsf{A}$ and $Q_1 \cdot \mathsf{A}$ may be inherited from the distance between the transformation mapping them.

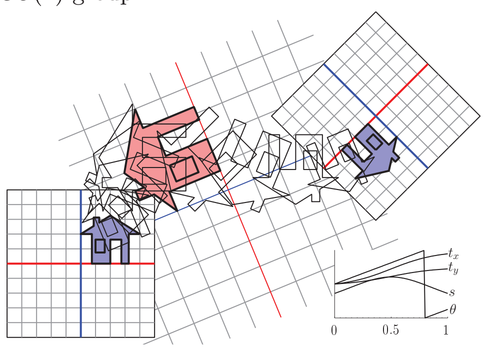
A first experiment was designed to illustrate the results of interpolation of spatial transformation when using different groups. Two instances to be interpolated were arbitrarily chosen: Q_0 as the identity transformation; and Q_1 as a clockwise rotation of 135 degrees followed by a translation of four units in the horizontal axis and two units along the vertical axis. Both transformations are elements from the $\mathcal{SE}(2)$ group. The metric used in all interpolations was selected as the one given by $\langle\!\langle U, U \rangle\!\rangle_I = \mathrm{trace}(U^T U)$, where U is the matrix representation of the velocity located at the identity.



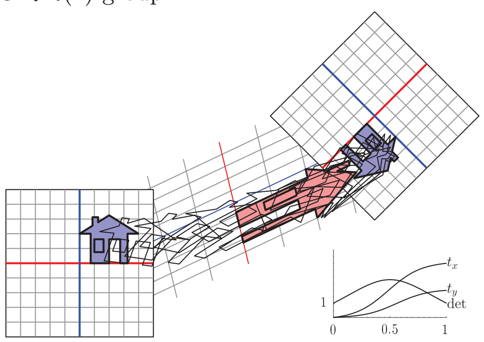
(a) $\mathcal{SE}(2)$ group



(b) $\mathcal{SIM}(2)$ group



(c) $\mathcal{SIM}(2)$ group, but not the shortest geodesic

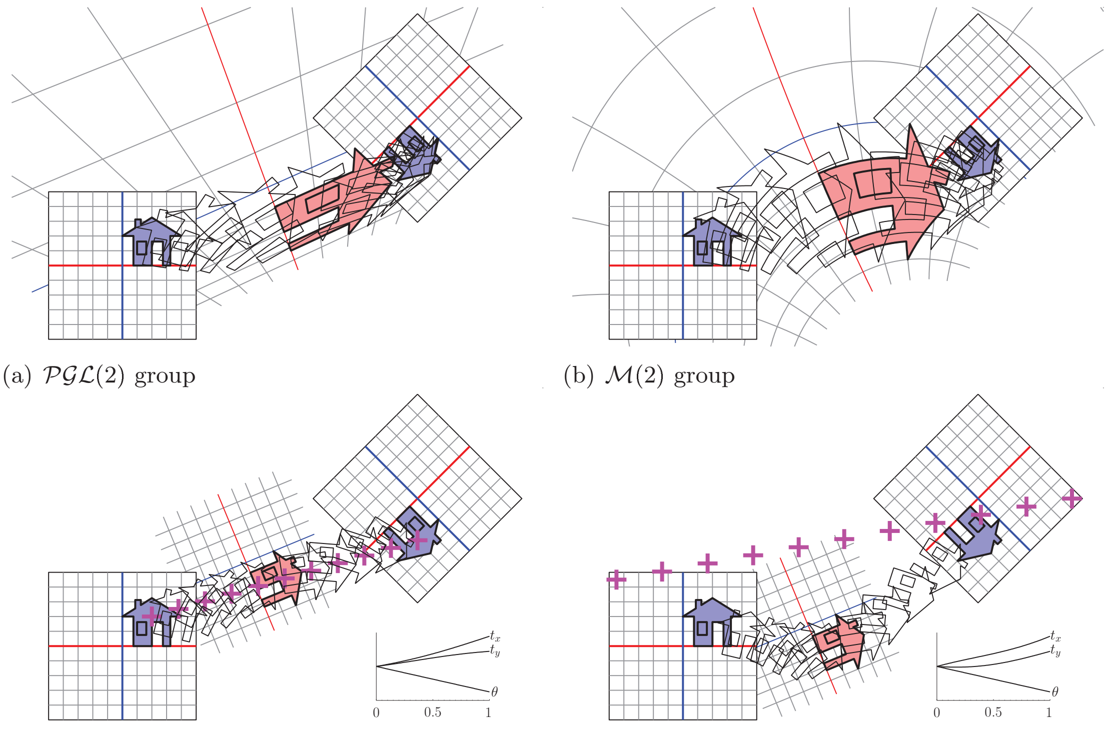


(d) $\mathcal{GA}(2)$ group

Figure 4. *Geodesic interpolation between two instances of $\mathcal{SE}(2)$ on different groups. The time course of some parameters of the transformations is also shown in arbitrary units.*

The geodesic interpolation on $\mathcal{SE}(2)$ is shown in Figure 4(a). In this case the interpolating path corresponds to a linear interpolation of the transformation parameters t_x, t_y, θ matching both instances (see the attached panel with the time-course of the parameters describing the transformation: θ denotes the rotation angle; t_x and t_y denote translations; s and det denote size and volume change, respectively; and vertical axes of these panels are in arbitrary units).

Figures 4(b)–(c) show two possible geodesics on $\mathcal{SIM}(2)$. Both curves are solutions of (9.2), both being global minima. This illustrates the nonuniqueness issue mentioned before. The shortest curve, shown in Figure 4(b), defines the distance between the transformations. It is interesting to see that, although in both the initial and final transformations the scale factor equals 1, the geodesics (b)–(c) change the object scale in order to minimize the curve lengths (this effect will also be discussed in section 9.3). A similar behavior can be seen in Figure 4(d) for the $\mathcal{GA}(2)$ group. It can be noted that the interpolating paths, and also the distances, between both transformations depend on the acting group. The resulting distances are 5.57 for $\mathcal{SE}(2)$; 4.83 (case (b)) and 6.57 (case (c)) for $\mathcal{SIM}(2)$; 4.66 for $\mathcal{SA}(2)$ (not shown in the figure); and 4.42 for $\mathcal{GA}(2)$. It can be said that for self-contained subgroups, the larger the dimension of the subgroup, the smaller the distance between a pair of transformations.



(a) $\mathcal{PGL}(2)$ group

(b) $\mathcal{M}(2)$ group

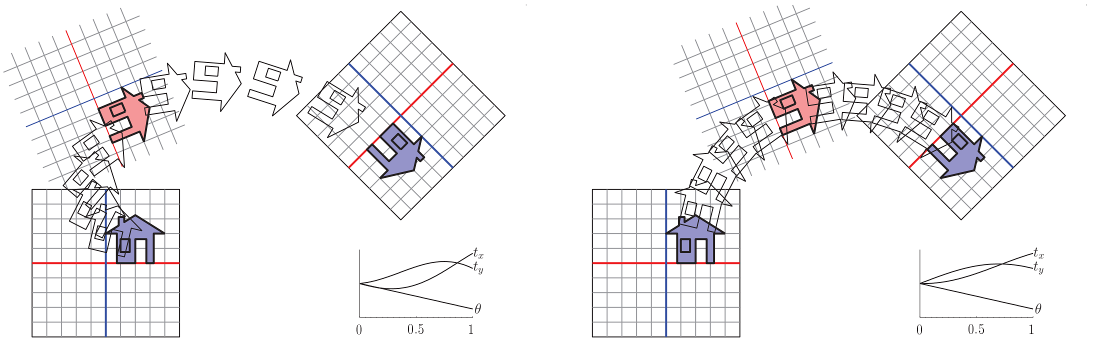
(c) $\mathcal{SO}(2) \times \mathcal{T}(2)$ group

(d) $\mathcal{SO}(2) \times \mathcal{T}(2)$ group

Figure 5. *Interpolation curves using $\mathcal{PGL}(2)$, $\mathcal{M}(2)$, and centered transformation $(\mathcal{SO}(2) \times \mathcal{T}(2))$ groups. In the cases* (c) *and* (d) *"+" markers indicate the center of rotation.*

Figure 5 shows interpolating paths similar to those in the previous figure but now obtained when using the following transformation groups: Projective group $\mathcal{PGL}(2)$; Möbius group $\mathcal{M}(2)$; and two examples of centered transformations $(\mathcal{SO}(2) \times \mathcal{T}(2))$. In the cases of centered transformations, the interpolation paths depend on the choice of the rotation center (indicated by "+" markers). This transformation model should be used only when a well-defined center can be determined; otherwise the result will depend on the arbitrary choice of the center.

Transformations can also be interpolated without using the left-invariant geodesic framework shown in Figure 3. For example, Figure 6(a) shows the interpolating geodesic on the $\mathcal{SE}(2)$ group but this time using a right-invariant metric. The one-parameter subgroup (or its translations) might also be used for interpolation [9, 79] (see Figure 6(b)), although it does not always exist (see section 2.5). This curve is computed by using the group exponential function $\gamma_{(t)} = Q_0 \exp(t\, Q_0^{-1} Q_1)$, where no metric is specified. This curve is computationally simpler [122, 7] than Riemannian geodesics. The elements along the curve belong to the smallest group containing both transformations, which in this example is the $\mathcal{SE}(2)$ group.



(a) $\mathcal{SE}(2)$ (right-invariant) (b) Group exponential

Figure 6. *Alternative interpolation paths to the framework illustrated in Figure 3.*

A second experiment was designed to illustrate the geodesic interpolation between two transformations which cannot be joined by a one-parameter curve. The instances to be interpolated are $Q_0 = I$ and Q_1 defined as the composition of a 180 degree rotation followed by a scale factor of 2 along the horizontal axis. Figure 7 shows the action of a left-invariant geodesic on the $\mathcal{GL}^+(2)$ group interpolating Q_0 and Q_1. In this particular example there are two possible geodesics with minimal length joining Q_0 and Q_1. In addition to the path shown in Figure 7, the other geodesic starts with a rotation in the opposite direction.

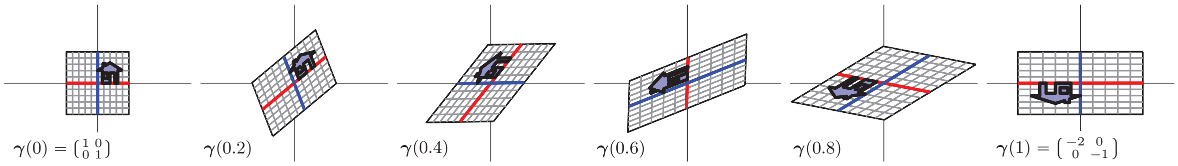


Figure 7. *Action of the left-invariant geodesic on $\mathcal{GL}(2)$ interpolating Q_0 and Q_1 which cannot be joined by a translated one-parameter subgroup curve.*

9.2. Intrinsic filtering. In this application, a spatial filtering with a Gaussian kernel is performed on spatial transformations distributed on a spatial domain $\Omega \subset \mathbb{R}^2$. The data to be filtered in this example is a field $J(x) \colon \Omega \to \mathbf{GL}^+(2)$ of spatial derivatives of an invertible and differentiable mapping $\Phi \colon \Omega \to \Omega$. Figure 8(a) shows a deformation field Φ where the glyphs represent the local deformation defined by a 2×2 matrix with positive determinant

$$J(x) = \begin{pmatrix} \partial_1\phi^1\big|_x & \partial_2\phi^1\big|_x \\ \partial_1\phi^2\big|_x & \partial_2\phi^2\big|_x \end{pmatrix} \in \mathbf{GL}^+(2)\,.$$

Regarding the practical details, the deformation mapping Φ was simulated from a continuous and piecewise cubic spline velocity field with vanishing boundary conditions. $\Phi(x)$ was obtained as the time integration up to time 1 of a particle on an initial point x [33]. Moreover, $J(x)$ was also computed integrating the local deformation guided by the Jacobian of the velocity field [16]. This parametrization guarantees (with an appropriate integration scheme) that all Jacobian matrices $J(x)$ belong to $\mathbf{GL}^+(2)$.

Figure 8(b) shows a glyph representation of a matrix-valued discrete grid (image) which corresponds to the field $J(x)$ in *Lagrangian coordinates* [84, p. 335]. The color of each cell represents the local volume change in logarithm units, $\log_{10}(\det(J(x)))$.

In order to simulate a noisy observation of the matrix-valued image $J(x)$, the deformation mapping Φ (Figure 8(a)) was composed with a random deformation with circular boundary conditions. The "noisy" Jacobian matrices from the perturbed deformation are shown in Figure 8(c).

The matrix-valued image of Figure 8(c) was filtered with a normalized 2D Gaussian kernel with σ equal to 1. If the filtering is performed independently on each component of the Jacobian matrices, the data elements are considered to belong to the vector space $\mathbb{M}_2$. Figure 8(d) shows the result of the componentwise filtering. It can be seen that the filtering yielded two locations with negative values of the Jacobian determinant (shown as green pixels).

In order to preserve the group structure an intrinsic filtering was proposed. Filtering of manifold-valued data can be interpreted as a *weighted mean* in a similar manner as in the case of vectorial data [137].

Given a data field $F\colon \Omega \to \mathbb{R}$ and a normalized spatial kernel $K(r)$, the discrete convolution is $(K * F)(x) = \sum_r K(r)F(x-r)$. The filtered field satisfies

$$\begin{aligned}\hat{F}(x) = (K * F)(x) &= \operatorname*{argmin}_M \sum_r K(r) \left\| F(x-r) - M \right\|_F^2 \\ &= \operatorname*{argmin}_M \sum_r K(r)\, \mathbf{distance}(F(x-r), M)^2 ,\end{aligned}$$

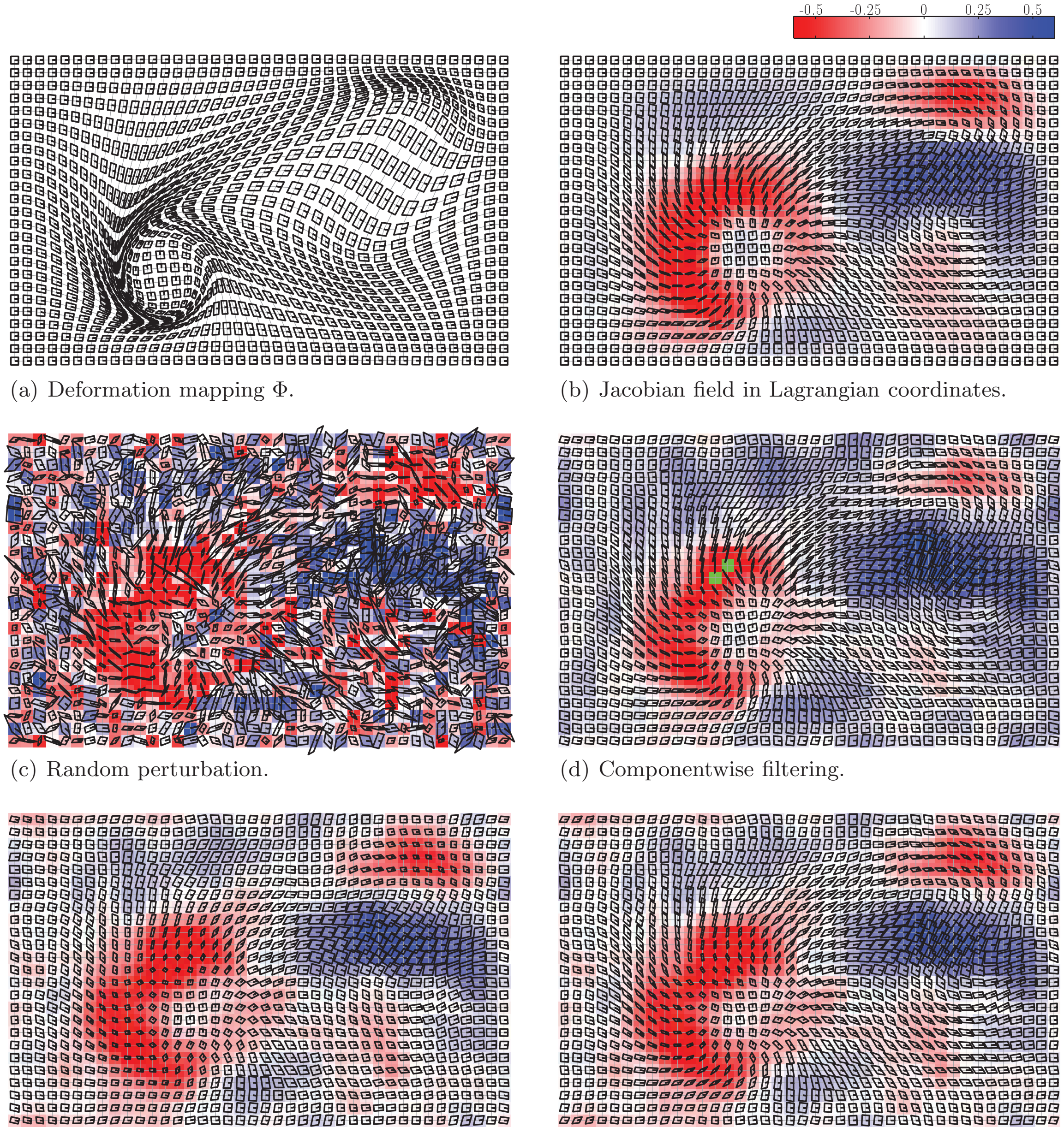
where $\mathbf{distance}(F(x), M)$ is the Euclidean distance between $F(x)$ and M. This interpretation of the convolution with a normalized kernel can be straightforwardly extended to Riemannian manifold-valued data by selecting an appropriate distance function. Note that two metrics are involved in this framework—the metric on the data, and the Euclidean metric for the spatial domain defining the relative weights between the neighboring elements via the kernel K.

For elements in a group $\mathcal{G}$ with a Riemannian distance, a necessary condition for the weighted mean $\hat{F}$ is

$$\sum_r K(r) \operatorname{Log}_{\hat{F}} (F(x-r)) = 0 ,$$

where all vectors $\operatorname{Log}_{\hat{F}} (F(x+r))$ belong to the same tangent space $T_{\hat{F}}\mathcal{G}$ and the usual addition can be performed. The following gradient descent procedure on $\mathcal{G}$ can be performed to compute the weighted mean:

$$\hat{F}_{k+1} = \operatorname{Exp}_{\hat{F}_k} \left(\varepsilon \sum_r K(r) \operatorname{Log}_{\hat{F}_k} (F(x-r)) \right) ,$$



(a) Deformation mapping Φ.

(b) Jacobian field in Lagrangian coordinates.

(c) Random perturbation.

(d) Componentwise filtering.

(e) Intrinsic filtering with left-invariant metric.

(f) Intrinsic filtering with right-invariant metric.

Figure 8. *Intrinsic Riemannian filtering of a field of general linear transformations. In all images, glyphs represent the local deformation. In panels* (b)–(f) *glyphs are located in Lagrangian coordinates, while in panel* (a) *they are located in deformed coordinates. Color represents local volume change in logarithm scale.*

where ε is a small enough step such that

$$\sum_r K(r)\,\mathbf{distance}(F(x-r),\hat{F}^{\varepsilon}_{k+1})^2 < \sum_r K(r)\,\mathbf{distance}(F(x-r),\hat{F}_k)^2,$$

which can be determined by a backtracking procedure.

While the convergence to a stationary point is guaranteed, necessary conditions ensuring uniqueness of minima are difficult to obtain, chiefly for compact groups (see, for example, [38, 76]). The proposed descent procedure converges to the weighted mean for a sufficiently small dispersion of the instances.

Intrinsic filtering on $\mathbf{GL}^+(2)$ was applied to the Jacobian matrices of Figure 8(c) using invariant metrics defined by $\langle\!\langle U, U \rangle\!\rangle_I = \mathrm{trace}(U^T U)$. The results are shown in Figures 8(e)–(f) for left- and right-invariant metrics, respectively. Note that both results differ, and the appropriate selection of the invariance depends on the application.

In this example, where the elements belong to $\mathcal{GL}^+(n) \equiv \mathcal{SL}(n) \times \mathcal{S}^+(1)$, it can be checked that intrinsic filtering satisfies

$$\log(\det(K * F)) = K * \log(\det(F))$$

for both left- and right-invariance cases. The logarithm of the local volume change descriptor of the intrinsically filtered transformations coincides with the scalar filtering of the logarithm of the volume descriptor from the original data.

9.3. Extension to any left-invariant metric of the closed-form expression for $\mathcal{ST}(1)$ group.

This section makes use of the conjugation of geodesics explained in section 7 to provide a closed-form expression for the $\mathcal{ST}(1)$ group under any invariant metric.

An element of $\mathcal{ST}(1)$ can be represented by the 2×2 matrix $Y = \left(\begin{smallmatrix} s & t \\ 0 & 1 \end{smallmatrix}\right)$, where $s \in \mathbb{R}^+$ is the scale factor and $t \in \mathbb{R}$ is the displacement parameter. A set of bases for its algebra is given by $\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \}$ (see section 8.4). Let $\boldsymbol{m} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ be the representation of the inertia operator for which a closed-form solution for $\mathrm{Exp}(\cdot)$ is given in (8.1). For this particular group the transformed and scaled inertia operator $\tilde{\boldsymbol{m}}$ given in section 7 takes the form

$$\tilde{\boldsymbol{m}} = \alpha\, \boldsymbol{B}^T (Y \otimes Y^{-T}) \boldsymbol{B}^{\dagger T} \boldsymbol{m} \boldsymbol{B}^{\dagger} (Y^T \otimes Y^{-1}) \boldsymbol{B} = \frac{\alpha}{s^2} \begin{pmatrix} s^2 + t^2 & t \\ t & 1 \end{pmatrix}. \tag{9.3}$$

Therefore, any possible left-invariant metric for $\mathcal{ST}(1)$ can be written as (9.3), where

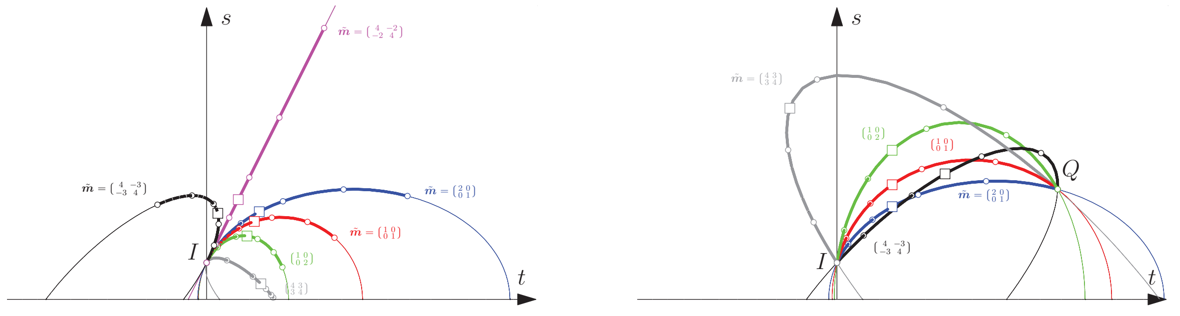
$$\tilde{\boldsymbol{m}} = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \Rightarrow \begin{cases} s = \frac{\sqrt{\det(\tilde{\boldsymbol{m}})}}{B}, \\ t = \frac{C}{B}, \\ \alpha = \frac{\det(\tilde{\boldsymbol{m}})}{B}. \end{cases} \tag{9.4}$$

Then, for any inertia operator represented by $\tilde{\boldsymbol{m}}$, the element $Y = \left(\begin{smallmatrix} s & t \\ 0 & 1 \end{smallmatrix}\right)$ that satisfies (7.1) can be computed from (9.4). Finally, the general closed-form expression for the Riemannian exponential is

$$\mathrm{Exp}_I^{\tilde{\boldsymbol{m}}}(U) = Y\, \mathrm{Exp}_I(Y^{-1} U Y)\, Y^{-1}, \qquad \text{where} \qquad Y = \begin{pmatrix} \sqrt{\det(\tilde{\boldsymbol{m}})}/B & C/B \\ 0 & 1 \end{pmatrix}$$

and $\mathrm{Exp}_I(Y^{-1} U Y)$ can be computed using (8.1).

Figure 9(a) shows several geodesics, all starting from the identity with a given initial velocity $U = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, but under different inertia operators $\tilde{\boldsymbol{m}}$. The corresponding inertia operator for each curve is given explicitly in the figure. The white circles along each curve show equidistant points along the segment from $t = 0$ to 1, and the white square is the element at the midpoint of the trajectory ($t = 0.5$). Geodesics under the inertia operator $\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ are known to be semicircles in the (t, s) plane with their centers on the $s = 0$ axis [161, 14] (see the red curves in Figure 9). Under a general $\tilde{\boldsymbol{m}}$ the geodesics become elliptical trajectories (considering a straight line as a degenerated ellipse).



(a) Geodesics from the identity with the same initial velocity.

(b) Geodesics from the identity to a given element.

Figure 9. *Geodesics for $\mathcal{ST}(1)$ using different inertia operators $\tilde{\boldsymbol{m}}$.*

Figure 9(b) shows several geodesics between the identity and the element $Q = \left(\begin{smallmatrix}3&6\\0&1\end{smallmatrix}\right)$. The experiments performed on this simple group illustrated the strong dependence of the metric on the geodesic trajectories [118]. This behavior is expected to hold also on other spatial transformation groups. Accordingly, the choice of the metric tensor is a crucial and application-dependent issue.

For a general k-dimensional group, the metric tensor belongs to a $k(k+1)/2$-dimensional family. In the best case, the conjugacy action can help to extend the closed-form solution for a $(k+1)$-dimensional family of metric tensors.

Appendix A. Example: Scale and translation group in one dimension in local coordinates. In order to illustrate the method to compute geodesics for a given chart a very simple group of spatial transformations will be used: scaling and translation in one-dimensional space, also known as 1D-affine transformations. The elements in this group can be represented by matrices of the form $\left(\begin{smallmatrix}s&t\\0&1\end{smallmatrix}\right)$, where $s \in \mathbb{R}^+$ is a scaling factor and $t \in \mathbb{R}$ represents a translation.

A possible chart for this group is $\varphi\left(\left(\begin{smallmatrix}s&t\\0&1\end{smallmatrix}\right)\right) = \left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)$. The domain $\mathbf{D}$ of this particular chart is the whole group $\mathbf{G}$, and the image of the function φ is $\Omega \equiv \mathbb{R}^+ \times \mathbb{R}$. The coordinates of the identity element are $\varphi(I) = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. Left-translations are written in coordinates as $\mathsf{L}_{\left(\begin{smallmatrix}r\\d\end{smallmatrix}\right)}\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right) = \left(\begin{smallmatrix}rs\\rt+d\end{smallmatrix}\right)$ and the inverse as $\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)^{-1} = \begin{pmatrix}1/s\\-t/s\end{pmatrix}$. Regarding velocities, they are represented by matrices of the form $\left(\begin{smallmatrix}\dot{s}&\dot{t}\\0&0\end{smallmatrix}\right)$. The function $\mathrm{d}\varphi$ is given by $\mathrm{d}\varphi\left(\left(\begin{smallmatrix}\dot{s}&\dot{t}\\0&0\end{smallmatrix}\right)\right) = \left(\begin{smallmatrix}\dot{s}\\\dot{t}\end{smallmatrix}\right)$ and $T_{\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)}\mathsf{L}_{\left(\begin{smallmatrix}r\\d\end{smallmatrix}\right)}\left(\begin{smallmatrix}\dot{s}\\\dot{t}\end{smallmatrix}\right) = \left(\begin{smallmatrix}r\dot{s}\\r\dot{t}\end{smallmatrix}\right)$.

To compute the 2×2 matrix $\boldsymbol{m}_{\bar{q}}$ for $\bar{q} = \left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)$ note that

$$\left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix}\right)^T \boldsymbol{m}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix}\right) = \left(T_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)} \mathsf{L}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)^{-1}} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix}\right)\right)^T \boldsymbol{m}_{\bar{e}} \left(T_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)} \mathsf{L}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)^{-1}} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix}\right)\right)$$
$$= \begin{pmatrix} \dot{s}/s \\ \dot{t}/s \end{pmatrix}^T \boldsymbol{m}_{\bar{e}} \begin{pmatrix} \dot{s}/s \\ \dot{t}/s \end{pmatrix},$$

and therefore $\boldsymbol{m}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)} = \frac{1}{s^2} \boldsymbol{m}_{\bar{e}}$.

Choosing $\boldsymbol{m}_{\bar{e}} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, its derivative is given by $(\boldsymbol{\mathcal{D}} \boldsymbol{m}_{\bar{q}})_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)} = -\frac{2}{s^3} \begin{pmatrix} \begin{smallmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{smallmatrix} \end{pmatrix}$. Substituting it into (odeChart), the following IVP system is obtained for a geodesic:

$$\ddot{s}_{(t)} = \frac{\dot{s}_{(t)}^2 - \dot{t}_{(t)}^2}{s_{(t)}},$$
$$\ddot{t}_{(t)} = 2\frac{\dot{s}_{(t)}\,\dot{t}_{(t)}}{s_{(t)}},$$
$$\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right)_{(0)} = \varphi(X),$$
$$\left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix}\right)_{(0)} = \mathrm{d}\varphi(V).$$

It is worth mentioning that computing geodesics in chart coordinates for a very simple transformation group yielded an IVP for which it is not straightforward to find a solution.

Appendix B. Euler–Lagrange equation using the group exponential chart. For an $n \times n$ matrix group $\mathbf{G}$ of dimension k, an open set of the algebra coordinates can be mapped to an open set of the group via the group exponential function $\varphi^{-1}\colon \Omega \subset \mathbb{R}^k \to \mathbf{D} \subset \mathbf{G}\colon \bar{x} \mapsto X = \exp(\sum_i B_i x^i)$, where Ω is an open set around 0 and $\mathbf{D}$ is an open set of the group around the identity. $\bar{x}$ are usually called canonical coordinates of the first kind. The inverse of this mapping defines the group exponential chart denoted by $(\mathbf{D}, \varphi)$, where $\varphi\colon \mathbf{D} \subset \mathbf{G} \to \Omega\colon X \mapsto \bar{x} = \boldsymbol{B}^\dagger \overline{\log(X)}$. It was shown in [149] that an admissible set Ω such that $(\varphi \circ \varphi^{-1})(\bar{x}) = \bar{x}$ is given by $\left\| \sum_i B_i x^i \right\|_F < \log(2)$.

Using the group exponential chart, the group identity $I \in \mathbf{G}$ is mapped to $0 \in \Omega$. The left-translation operation in chart coordinates is given by

$$\mathsf{L}_{\bar{y}}(\bar{x}) = \varphi\left(\varphi^{-1}(\bar{y})\varphi^{-1}(\bar{x})\right) = \boldsymbol{B}^\dagger \overline{\log\left(\exp(B_j y^j)\exp(B_i x^i)\right)},$$

where $B_i x^i$ is used as shorthand for $\sum_i B_i x^i$. In this chart, the inverse operation results in $(\bar{x})^{-1} = -\bar{x}$.

Let $\bar{x}_{(t)}$ be the coordinates of a curve in Ω, and let $\dot{\bar{x}}_{(t)} = \mathrm{d}\varphi(\dot{X}_{(t)}) \in \mathbb{R}^k$ be the coordinates of the curve velocity; then the chart coordinates of the velocity of the left-translated curve $Z_{(t)} = YX_{(t)}$ are

$$\begin{aligned} \mathrm{d}\varphi(Z)|_t &= \boldsymbol{B}^\dagger\, \boldsymbol{\mathcal{D}}\mathrm{log}\left(\exp(B_j y^j)\exp(B_i x^i_{(t)})\right)\left(I_n \otimes \exp(B_j y^j)\right) \boldsymbol{\mathcal{D}}\mathrm{exp}(B_i x^i_{(t)}) \boldsymbol{B}\, \dot{\bar{x}}_{(t)} \\ &= T_{\bar{x}} \mathsf{L}_{\bar{y}}\left(\dot{\bar{x}}_{(t)}\right), \end{aligned} \tag{B.1}$$

where $\boldsymbol{\mathcal{D}}$ is the Fréchet derivative operator.

Given a left-invariant inner product defined by the matrix $\boldsymbol{m}$ in the algebra, the inertia operator in coordinates of the chart $\boldsymbol{m}_{\bar{x}}$ is defined by (see (3.1))

$$\dot{\bar{x}}_{(t)}^T \boldsymbol{m}_{\bar{x}} \dot{\bar{x}}_{(t)} = \left(T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \dot{\bar{x}}_{(t)}\right)^T \boldsymbol{m} \left(T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \dot{\bar{x}}_{(t)}\right) .$$

Note that in this particular chart $\boldsymbol{m}_{\bar{e}} = \boldsymbol{m}$, as can be seen from the definition of m in (2.9). By using $\bar{y} = (\bar{x})^{-1}$, (B.1) simplifies to

$$T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \dot{\bar{x}}_{(t)} = \underbrace{\boldsymbol{B}^\dagger \left(I_n \otimes \exp(-B_i x^i)\right) \boldsymbol{\mathcal{D}}\mathrm{exp}(B_i x^i_{(t)}) \boldsymbol{B}}_{\mathsf{t}_{\bar{x}}} \ \dot{\bar{x}}_{(t)} = \mathsf{t}_{\bar{x}}\, \dot{\bar{x}}_{(t)} \,,$$

where $\mathsf{t}_{\bar{x}}$ is a $k{\times}k$ matrix. Therefore, the matrix $\boldsymbol{m}_{\bar{x}} = \mathsf{t}_{\bar{x}}^T\, \boldsymbol{m}\, \mathsf{t}_{\bar{x}}$.

The geodesic problem written in chart coordinates is solved by a curve $\bar{q}_{(t)}$ satisfying (odeChart) given in section 3.1. The Fréchet derivative of the metric tensor, $(\boldsymbol{\mathcal{D}} \boldsymbol{m}_{\bar{q}})$, must be computed, which is given by

$$\boldsymbol{\mathcal{D}} \boldsymbol{m}_{\bar{q}} = (K_{kk} + I_{k^2}) \left(I_k \otimes \mathsf{t}_{\bar{q}}^T\, \boldsymbol{m}\right) \boldsymbol{\mathcal{D}} \mathsf{t}_{\bar{q}} \,, \tag{B.2}$$

where K_{mm} is the *commutation matrix* [1] defined by $K_{mm}\overline{M} = \overline{M^T}$ for an $m{\times}m$ matrix M, and

$$\begin{aligned}\boldsymbol{\mathcal{D}} \mathsf{t}_{\bar{q}} = \left(\boldsymbol{B}^T \otimes \boldsymbol{B}^\dagger\right) \Big(& \left(I_{n^3} \otimes \exp\left(-B_i q^i\right)\right) \boldsymbol{\mathcal{D}}^2\mathrm{exp}\left(B_i q^i\right) - \\ & \left(\boldsymbol{\mathcal{D}}\mathrm{exp}\left(B_i q^i\right)^T \otimes I_{n^2}\right) \left(K_{nn} \otimes I_{n^2}\right) \left(I_n \otimes \overline{I_n} \otimes I_n\right) \boldsymbol{\mathcal{D}}\mathrm{exp}\left(-B_i q^i\right)\Big) \boldsymbol{B} \,.\end{aligned} \tag{B.3}$$

To evaluate the expressions in (B.2) and (B.3) it is also necessary to compute $\boldsymbol{\mathcal{D}}\mathrm{exp}(M)$ and $\boldsymbol{\mathcal{D}}^2\mathrm{exp}(M)$. There are different attempts to compute $\boldsymbol{\mathcal{D}}\mathrm{exp}(M)$ [8, 7, 125, 102, 114]. In our computations we have used the approach given in [115], $\mathbf{F}$ being an analytic matrix function; then

$$\mathbf{F}\left(\left(\begin{array}{c|c} M & P \\ \hline 0 & M \end{array}\right) \right) = \left(\begin{array}{c|c} \mathbf{F}(M) & \mathrm{d}_\varepsilon \mathbf{F}(M + \varepsilon P) \\ \hline 0 & \mathbf{F}(M) \end{array}\right) ,$$

and, if P is the jth canonical perturbation, then the vectorization of the upper-right submatrix $\overline{\mathrm{d}_\varepsilon \mathbf{F}(M + \varepsilon P)}$ is the jth column of the matrix $\boldsymbol{\mathcal{D}}\mathbf{F}(M)$. Additionally [82],

$$\mathbf{F}\left(\left(\begin{array}{c|c|c|c} M & P_1 & P_2 & 0 \\ \hline 0 & M & 0 & P_2 \\ \hline 0 & 0 & M & P_1 \\ \hline 0 & 0 & 0 & M \end{array}\right) \right) = \left(\begin{array}{c|c|c|c} \mathbf{F}(M) & \mathrm{d}_\varepsilon \mathbf{F}(M + \varepsilon P_1) & \mathrm{d}_\varrho \mathbf{F}(M + \varrho P_2) & \mathrm{d}_\varrho \mathrm{d}_\varepsilon \mathbf{F}(M + \varepsilon P_1 + \varrho P_2) \\ \hline 0 & \mathbf{F}(M) & 0 & \mathrm{d}_\varrho \mathbf{F}(M + \varrho P_2) \\ \hline 0 & 0 & \mathbf{F}(M) & \mathrm{d}_\varepsilon \mathbf{F}(M + \varepsilon P_1) \\ \hline 0 & 0 & 0 & \mathbf{F}(M) \end{array}\right) .$$

Rearranging the elements of the upper-right block in an $n^4{\times}n^2$ matrix for the n^2 possible canonical perturbations P_1 and P_2, it is possible to build the matrix $\boldsymbol{\mathcal{D}}^2\mathbf{F}(M)$ such that $\overline{\boldsymbol{\mathcal{D}}\mathbf{F}(M + \varepsilon P)} = \overline{\boldsymbol{\mathcal{D}}\mathbf{F}(M)} + \varepsilon(\boldsymbol{\mathcal{D}}^2\mathbf{F}(M))\overline{P} + O(\varepsilon)$.

Finally, regarding $\boldsymbol{\mathcal{D}}\log(M)$, using the inverse function theorem, we can obtain

$$\boldsymbol{\mathcal{D}}\log(M) = \left(\boldsymbol{\mathcal{D}}\mathrm{exp}(\log(M))\right)^{-1} .$$

REFERENCES

[1] K. M. Abadir and J. R. Magnus, *Matrix Algebra*, Cambridge University Press, Cambridge, UK, 2005.

[2] R. H. Abraham and J. E. Marsden, *Foundations of Mechanics*, AMS, Providence, RI, 1978.

[3] P.-A. Absil, C. G. Baker, and K. A. Gallivan, *Trust-region methods on Riemannian manifolds*, Found. Comput. Math., 7 (2007), pp. 303–330.

[4] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2009.

[5] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization on manifolds: Methods and applications*, in Recent Advances in Optimization and Its Applications in Engineering, M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, eds., Springer, Berlin, Heidelberg, 2010, pp. 125–144.

[6] R. L. Adler, J. P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal., 22 (2002), pp. 359–390.

[7] A. H. Al-Mohy, *Algorithms for the Matrix Exponential and Its Fréchet Derivative*, Ph.D. thesis, University of Manchester, Manchester, UK, 2010.

[8] A. H. Al-Mohy and N. J. Higham, *Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation*, SIAM J. Matrix Anal. Appl., 30 (2009), pp. 1639–1657.

[9] M. Alexa, *Linear combination of transformations*, ACM Trans. Graphics, 21 (2002), pp. 380–387.

[10] E. Andruchow, G. Larotonda, L. Recht, and A. Varela, *The Left Invariant Metric in the General Linear Group*, preprint, arXiv, http://arxiv.org/abs/1109.0520 (2011).

[11] A. Angelidis, *Shape Modeling by Swept Space Deformation*, Ph.D. thesis, University of Otago, Dunedin, New Zealand, 2005.

[12] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, New York, 1989.

[13] V. I. Arnold and B. A. Khesin, *Topological Methods in Hydrodynamics*, Springer, New York, 1998.

[14] V. Arsigny, *Processing Data in Lie Groups: An Algebraic Approach. Application to Non-linear Registration and Diffusion Tensor MRI*, Ph.D. thesis, École Polytechnique, Paris, France, 2006.

[15] V. Arsigny, X. Pennec, and N. Ayache, *Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations—application to the registration of histological slices*, Med. Image Anal., 9 (2005), pp. 507–523.

[16] J. Ashburner, *A fast diffeomorphic image registration algorithm*, NeuroImage, 38 (2007), pp. 95–113.

[17] M. Balasubramanian, J. R. Polimeni, and E. L. Schwartz, *Exact geodesics and shortest paths on polyhedral surfaces*, IEEE Trans. Pattern Anal. Machine Intell., 31 (2009), pp. 1006–1016.

[18] I. N. Bankman, *Handbook of Medical Imaging: Processing and Analysis*, Academic Press, New York, 2000.

[19] E. Bayro-Corrochano and J. Ortegón-Aguilar, *Lie algebra approach for tracking and 3D motion estimation using monocular vision*, Image Vis. Comput., 25 (2007), pp. 907–921.

[20] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, *Computing large deformation metric mappings via geodesic flows of diffeomorphisms*, Internat. J. Comput. Vis., 61 (2005), pp. 139–157.

[21] E. Begelfor and M. Werman, *Interpolating Transformations: Riemannian Geometry, Invariant Metrics and Geodesics*, unpublished, accessed from ftp://ftp.cs.huji.ac.il/users/aristo/papers/SYGRAPH2005/sig05.pdf (2005).

[22] B. Berkels, G. Linkmann, and M. Rumpf, *An SL(2) invariant shape median*, J. Math. Imag. Vis., 37 (2010), pp. 85–97.

[23] P. J. Besl, *Geometric modeling and computer vision*, Proc. IEEE, 76 (1988), pp. 936–958.

[24] A. Bloch, P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu, *The Euler–Poincaré equations and double bracket dissipation*, Comm. Math. Phys., 175 (1996), pp. 1–42.

[25] A. M. Bloch, P. E. Crouch, J. E. Marsden, and T. S. Ratiu, *The symmetric representation of the rigid body equations and their discretization*, Nonlinearity, 15 (2002), p. 1309.

[26] A. M. Bloch, P. E. Crouch, J. E. Marsden, and A. K. Sanyal, *Optimal control and geodesics on quadratic matrix Lie groups*, Found. Comput. Math., 8 (2008), pp. 469–500.

[27] A. M. Bloch, P. E. Crouch, N. Nordkvist, and A. K. Sanyal, *Embedded geodesic problems and optimal control for matrix Lie groups*, J. Geom. Mech., 3 (2011), pp. 197–223.

[28] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache, *Geometric variability of the scoliotic spine using statistics on articulated shape models*, IEEE Trans. Medical Imag., 27 (2008), pp. 557–568.

[29] D. Bommes and L. Kobbelt, *Accurate computation of geodesic distance fields for polygonal curves on triangle meshes*, in Proceedings of the Conference on Vision, Modelling, and Visualization, VMV 2007, Vol. 7, Saarbrücken, Germany, 2007, pp. 151–160.

[30] F. L. Bookstein, *Principal warps: Thin-plate splines and the decomposition of deformations*, IEEE Trans. Pattern Anal. Machine Intell., 11 (1989), pp. 567–585.

[31] J. M. Borwein and R. E. Crandall, *Closed forms: What they are and why we care*, Notices Amer. Math. Soc., 60 (2013), pp. 50–65.

[32] M. Bossa and S. Olmos, *Statistical model of similarity transformations: Building a multi-object pose model of brain structures*, in IEEE Computer Vision and Pattern Recognition Workshops, IEEE, Washington, DC, 2006, pp. 59–66.

[33] M. Bossa, E. Zacur, and S. Olmos, *Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation*, in IEEE Computer Vision and Pattern Recognition Workshops, IEEE, Washington, DC, 2008, pp. 23–28.

[34] M. Bossa, E. Zacur, and S. Olmos, *Statistical analysis of relative pose information of subcortical nuclei: Application on ADNI data*, NeuroImage, 55 (2011), pp. 999–1008.

[35] N. Boumal, *Interpolation and regression of rotation matrices*, in Geometric Science of Information, Lecture Notes in Comput. Sci. 8085, F. Nielsen and F. Barbaresco, eds., Springer, Berlin, Heidelberg, 2013, pp. 345–352.

[36] Y. Boykov and V. Kolmogorov, *Computing geodesics and minimal surfaces via graph cuts*, in Proceedings of the IEEE International Conference on Computer Vision, Vol. 1, IEEE, Washington, DC, 2003, pp. 26–33.

[37] K. Burns and M. Gidea, *Differential Geometry and Topology: With a View to Dynamical Systems*, CRC, Boca Raton, FL, 2005.

[38] S. R. Buss and J. P. Fillmore, *Spherical averages and applications to spherical splines and interpolation*, ACM Trans. Graphics, 20 (2001), pp. 95–126.

[39] J. C. Butcher, *Numerical methods for ordinary differential equations in the* 20*th century*, J. Comput. Appl. Math., 125 (2000), pp. 1–29.

[40] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*, Wiley, New York, 2008.

[41] V. Caselles, R. Kimmel, and G. Sapiro, *Geodesic active contours*, Internat. J. Comput. Vis., 22 (1997), pp. 61–79.

[42] H. Cendra, D. D. Holm, J. E. Marsden, and T. S. Ratiu, *Lagrangian reduction, the Euler–Poincaré equations, and semidirect products*, in Geometry of Differential Equations, A. Khovanskiĭ, ed., Amer. Math. Soc. Transl. Ser. 2 186, AMS, Providence, RI, 1998, pp. 1–26.

[43] H. Cendra, J. E. Marsden, S. Pekarsky, and T. S. Ratiu, *Variational principles for Lie–Poisson and Hamilton–Poincaré equations*, Moscow Math. J., 3 (2003), pp. 833–867.

[44] C. T. Chang, B. Gorissen, and S. Melchior, *Fast oriented bounding box optimization on the rotation group SO(*3*,R)*, ACM Trans. Graphics, 30 (2011), pp. 122:1–122:16.

[45] J. J. Craig, *Introduction to Robotics: Mechanics and Control*, Prentice–Hall, Englewood Cliffs, NJ, 2004.

[46] A. Criminisi, T. Sharp, and A. Blake, *GeoS: Geodesic image segmentation*, in Computer Vision–ECCV 2008, Lecture Notes in Comput. Sci. 5302, D. Forsyth, P. Torr, and A. Zisserman, eds., Springer, Berlin, Heidelberg, 2008, pp. 99–112.

[47] P. E. Crouch, N. Nordkvist, and A. K. Sanyal, *Optimal control and geodesics on matrix Lie groups*, in Portuguese Conference on Automatic Control–CONTROLO 2010, 2010.

[48] R. de Levie, *An improved numerical approximation for the first derivative*, J. Chem. Sci., 121 (2009), pp. 935–950.

[49] J. P. Dedieu and D. Nowicki, *Symplectic methods for the approximation of the exponential map and the Newton iteration on Riemannian submanifolds*, J. Complexity, 21 (2005), pp. 487–501.

[50] M. P. Do Carmo, *Riemannian Geometry*, Birkhäuser, Basel, 1992.

[51] T. Drummond and R. Cipolla, *Application of Lie algebras to visual servoing*, Internat. J. Comput. Vis., 37 (2000), pp. 21–41.

[52] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*, Wiley, New York, 1998.
[53] B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov, *Modern Geometry—Methods and Applications: Part* I: *The Geometry of Surfaces, Transformation Groups, and Fields*, Grad. Texts in Math., Springer-Verlag, Berlin, 1984.
[54] P. Fillard, V. Arsigny, N. Ayache, and X. Pennec, *A Riemannian framework for the processing of tensor-valued images*, in Deep Structure, Singularities, and Computer Vision, Lecture Notes in Comput. Sci. 3753, O. F. Olsen, L. Florack, and A. Kuijper, eds., Springer, Berlin, Heidelberg, 2005, pp. 112–123.
[55] P. T. Fletcher, *Geodesic regression and the theory of least squares on Riemannian manifolds*, Internat. J. Comput. Vis., 105 (2013), pp. 171–185.
[56] P. T. Fletcher and S. Joshi, *Riemannian geometry for the statistical analysis of diffusion tensor data*, Signal Process., 87 (2007), pp. 250–262.
[57] P. T. Fletcher, S. Venkatasubramanian, and S. Joshi, *The geometric median on Riemannian manifolds with application to robust atlas estimation*, NeuroImage, 45 (2009), p. S143.
[58] J. Flusser and B. Zitová, *A comment on "A novel approach for the registration of weak affine images,"* Pattern Recog. Lett., 34 (2013), pp. 1381–1385.
[59] O. Freifeld and M. J. Black, *Lie bodies: A manifold representation of* 3*D human shape*, in European Conference on Computer Vision, Vol. 1 of ECCV'12, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, eds., Springer-Verlag, New York, 2012, pp. 1–14.
[60] D. Gabay, *Minimizing a differentiable function over a differential manifold*, J. Optim. Theory Appl., 37 (1982), pp. 177–219.
[61] J. Gallier, *Logarithms and Square Roots of Real Matrices. Existence, Uniqueness and Applications in Medical Imaging*, unpublished, accessed from http://www.cis.upenn.edu/~cis610/matlog.pdf (2011).
[62] J. Gallier, *Notes on Differential Geometry and Lie Groups*, unpublished, accessed from http://www.cis.upenn.edu/~jean/gbooks/manif.html (2013).
[63] J. Gallier and D. Xu, *Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices*, Internat. J. Robotics Automation, 17 (2002).
[64] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian Geometry*, Springer-Verlag, New York, 2004.
[65] K. Gorczowski, M. Styner, J. Y. Jeong, J. S. Marron, J. Piven, H. C. Hazlett, S. M. Pizer, and G. Gerig, *Multi-object analysis of volume, pose, and shape using statistical discrimination*, IEEE Trans. Pattern Anal. Machine Intell., 32 (2010), pp. 652–661.
[66] V. Govindu, *Robustness in motion averaging*, in Asian Conference on Computer Vision–ACCV 2006, Lecture Notes in Comput. Sci. 3852, P. J. Narayanan, S. K. Nayar, and H.-Y. Shum, eds., Springer, Berlin, Heidelberg, 2006, pp. 457–466.
[67] C. Gramkow, *On averaging rotations*, J. Math. Imaging Vision, 15 (2001), pp. 7–16.
[68] F. S. Grassia, *Practical parameterization of rotations using the exponential map*, J. Graphics Tools, 3 (1998), pp. 1–13.
[69] U. Grenander and M. I. Miller, *Computational anatomy: An emerging discipline*, Quart. Appl. Math., 56 (1998), pp. 617–694.
[70] H. W. Guggenheimer, *Plane Geometry and Its Groups*, Holden–Day, San Francisco, CA, 1967.
[71] Y. Gur and N. Sochen, *Regularizing flows over Lie groups*, J. Math. Imaging Vis., 33 (2009), pp. 195–208.
[72] S. Gwak, J. Kim, and F. C. Park, *Numerical optimization on the Euclidean group with applications to camera calibration*, IEEE Trans. Robotics Automation, 19 (2003), pp. 65–74.
[73] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I: *Nonstiff Problems*, Springer, New York, 2011.
[74] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer, New York, 2003.
[75] A. J. Hanson, *Visualizing Quaternions*, Morgan Kaufmann, San Francisco, CA, 2006.
[76] R. Hartley, J. Trumpf, Y. Dai, and H. Li, *Rotation averaging*, Internat. J. Comput. Vis., 103 (2013), pp. 267–305.
[77] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2004.

[78] R. I. Hartley and F. Kahl, *Global optimization through rotation space search*, Internat. J. Comput. Vis., 82 (2009), pp. 64–79.
[79] A. Hawkins and C. M. Grimm, *Keyframing using linear interpolation of matrices*, in ACM SIGGRAPH 2005 Posters, SIGGRAPH '05, ACM, New York, 2005.
[80] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1979.
[81] N. J. Higham and A. H. Al-Mohy, *Computing matrix functions*, Acta Numerica, 19 (2010), pp. 159–208.
[82] N. J. Higham and S. D. Relton, *Higher order Fréchet derivatives of matrix functions and the level-2 condition number*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1019–1037.
[83] J. Hinkle, P. T. Fletcher, and S. Joshi, *Intrinsic polynomials for regression on Riemannian manifolds*, J. Math. Imaging Vis., doi:10.1007/s10851-013-0489-5 (2013).
[84] D. D. Holm, T. Schmah, C. Stoica, and D. C. P. Ellis, *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*, Oxford University Press, Oxford, UK, 2009.
[85] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta Numer., 9 (2000), pp. 215–365.
[86] V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, Cambridge,UK, 1996.
[87] I. G. Kang and F. C. Park, *Cubic spline algorithms for orientation interpolation*, Internat. J. Numer. Methods Engrg., 46 (1999), pp. 45–64.
[88] H. B. Keller, *Numerical Solution of Two Point Boundary Value Problems*, SIAM, Philadelphia, 1976.
[89] M. Kobilarov, K. Crane, and M. Desbrun, *Lie group integrators for animation and control of vehicles*, ACM Trans. Graphics, 28 (2009), pp. 16:1–16:14.
[90] J. Kohlrausch, K. Rohr, and H. S. Stiehl, *A new class of elastic body splines for nonrigid registration of medical images*, J. Math. Imaging Vis., 23 (2005), pp. 253–280.
[91] B. Kolev, *Lie groups and mechanics: An introduction*, J. Nonlinear Math. Phys., 11 (2004), pp. 480–498.
[92] B. Kolev, *Groupes de Lie et mécanique*, unpublished, accessed from http://www.cmi.univ-mrs.fr/~kolev/pdf/GLM.pdf (2007).
[93] J. J. Kuffner, *Effective sampling and distance metrics for 3D rigid body path planning*, in IEEE International Conference on Robotics and Automation, Vol. 4, IEEE, Washington, DC, 2004, pp. 3993–3998.
[94] J. Kwon, H. Lee, F. Park, and K. Lee, *A geometric particle filter for template-based visual tracking*, IEEE Trans. Pattern Anal. Machine Intell., 36 (2014), pp. 625–643.
[95] J. Kwon and F. C. Park, *Visual tracking via particle filtering on the affine group*, Internat. J. Robotics Res., 29 (2010), pp. 198–217.
[96] J. Kybic and M. Unser, *Fast parametric elastic image registration*, IEEE Trans. Image Process., 12 (2003), pp. 1427–1442.
[97] H. T. Laquer, *Invariant affine connections on Lie groups*, Trans. Amer. Math. Soc., 331 (1992), pp. 541–551.
[98] S. Lee, M. Choi, H. Kim, and F. C. Park, *Geometric direct search algorithms for image registration*, IEEE Trans. Image Process., 16 (2007), pp. 2215–2224.
[99] S. Lee, G. Wolberg, K. Y. Chwa, and S. Y. Shin, *Image metamorphosis with scattered feature constraints*, IEEE Trans. Visualization Comput. Graphics, 2 (1996), pp. 337–354.
[100] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, *Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing*, J. Math. Imaging Vis., 25 (2006), pp. 423–444.
[101] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*, Wiley, New York, 2012.
[102] C. Li, Y. Sheng, and M. Wang, *An effective method to compute Fréchet derivative of matrix exponential and its error analysis*, J. Inform. Comput. Sci., 7 (2010), pp. 1854–1859.
[103] D. Li and Y. Zhang, *A novel approach for the registration of weak affine images*, Pattern Recog. Lett., 33 (2012), pp. 1647–1655.
[104] G. Li, Y. Liu, J. Yin, and Z. Shi, *Newton geodesic optimization on special linear group*, in Proceedings of the IEEE Conference on Decision and Control, IEEE, Washington, DC, 2009, pp. 4346–4351.

[105] H. Li and R. Hartley, *The 3D-3D registration problem revisited*, in Proceedings of the IEEE International Conference on Computer Vision, IEEE, Washington, DC, 2007, pp. 1–8.

[106] J. Li and P. Hao, *Smooth interpolation on homogeneous matrix groups for computer animation*, J. Zhejiang University SCIENCE A, 7 (2006), pp. 1168–1177.

[107] M. Li, T. Tan, W. Chen, and K. Huang, *Efficient object tracking by incremental self-tuning particle filtering on the affine group*, IEEE Trans. Image Process., 21 (2012), pp. 1298–1313.

[108] D. Lin, E. Grimson, and J. Fisher, *Learning visual flows: A Lie algebraic approach*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Washington, DC, 2009, pp. 747–754.

[109] M. Lorenzi and X. Pennec, *Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration*, Internat. J. Comput. Vis., 105 (2013), pp. 111–127.

[110] L. Machado, F. Silva Leite, and K. Krakowski, *Higher-order smoothing splines versus least squares problems on Riemannian manifolds*, J. Dynam. Control Syst., 16 (2010), pp. 121–148.

[111] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, New York, 1988.

[112] A. Makadia, A. I. Patterson, and K. Daniilidis, *Fully automatic registration of 3D point clouds*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '06, IEEE Computer Society, Washington, DC, 2006, pp. 1297–1304.

[113] J. E. Marsden, *Lectures on Mechanics*, Cambridge University Press, Cambridge, UK, 1992.

[114] R. Mathias, *Evaluating the Fréchet derivative of the matrix exponential*, Numer. Math., 63 (1992), pp. 213–226.

[115] R. Mathias, *A chain rule for matrix functions and applications*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 610–620.

[116] R. I. McLachlan and G. R. W. Quispel, *Geometric integrators for ODEs*, J. Phys. A, 39 (2006), p. 5251.

[117] M. I. Miller, A. Trouvé, and L. Younes, *On the metrics and Euler–Lagrange equations of computational anatomy*, Ann. Rev. Biomedical Engrg., 4 (2002), pp. 375–405.

[118] J. Milnor, *Curvatures of left invariant metrics on Lie groups*, Adv. Math., 21 (1976), pp. 293–329.

[119] M. Moakher, *Means and averaging in the group of rotations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 1–16.

[120] J. Modersitzki, *Numerical Methods for Image Registration*, Oxford University Press, Oxford, UK, 2004.

[121] K. Modin, M. Perlmutter, S. Marsland, and R. McLachlan, *On Euler–Arnold equations and totally geodesic subgroups*, J. Geom. Phys., 61 (2011), pp. 1446–1461.

[122] C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Rev., 45 (2003), pp. 3–49.

[123] M. E. Mortenson, *Geometric Modeling*, Wiley, New York, 1997.

[124] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC, Boca Raton, FL, 1994.

[125] I. Najfeld and T. F. Havel, *Derivatives of the matrix exponential and their computation*, Adv. Appl. Math., 16 (1995), pp. 321–375.

[126] L. Noakes and T. Popiel, *Geometry for robot path planning*, Robotica, 25 (2007), pp. 691–702.

[127] N. Nordkvist, P. E. Crouch, A. M. Bloch, and A. K. Sanyal, *Embedded optimal control problems*, in Proceedings of the IEEE Conference on Decision and Control and European Control Conference, IEEE, Washington, DC, 2011, pp. 7311–7316.

[128] C. Olsson, F. Kahl, and M. Oskarsson, *The registration problem revisited: Optimal solutions from points, lines and planes*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, IEEE, Washington, DC, 2006, pp. 1206–1213.

[129] F. C. Park, *Distance metrics on the rigid-body motions with applications to mechanism design*, J. Mech. Design, 117 (1995), pp. 48–54.

[130] F. C. Park, J. E. Bobrow, and S. R. Ploen, *A Lie group formulation of robot dynamics*, Internat. J. Robotics Res., 14 (1995), pp. 609–618.

[131] F. C. Park and M. W. Kim, *Lie theory, Riemannian geometry, and the dynamics of coupled rigid bodies*, ZAMM Z. Angew. Math. Phys., 51 (2000), pp. 820–834.

[132] F. C. Park and B. Ravani, *Smooth invariant interpolation of rotations*, ACM Trans. Graphics, 16 (1997), pp. 277–295.

[133] X. Pennec, *Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements*, J. Math. Imaging Vis., 25 (2006), pp. 127–154.

[134] X. Pennec, *Bi-invariant means on Lie groups with Cartan–Schouten connections*, in Geometric Science of Information, Lecture Notes in Computer Science 8085, F. Nielsen and F. Barbaresco, eds., Springer, Berlin, Heidelberg, 2013, pp. 59–67.

[135] X. Pennec and V. Arsigny, *Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups*, in Matrix Information Geometry, F. Barbaresco, A. Mishra, and F. Nielsen, eds., Springer, New York, 2012, pp. 123–166.

[136] X. Pennec and N. Ayache, *Uniform distribution, distance and expectation problems for geometric features processing*, J. Math. Imaging Vis., 9 (1998), pp. 49–67.

[137] X. Pennec, P. Fillard, and N. Ayache, *A Riemannian framework for tensor computing*, Internat. J. Comput. Vis., 66 (2006), pp. 41–66.

[138] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, *A comparison of similarity measures for use in 2D-3D medical image registration*, IEEE Trans. Medical Imaging, 17 (1998), pp. 586–595.

[139] G. Peyré and L. D. Cohen, *Geodesic remeshing using front propagation*, Internat. J. Comput. Vis., 69 (2006), pp. 145–156.

[140] M. T. Pham, O. J. Woodford, F. Perbet, A. Maki, B. Stenger, and R. Cipolla, *A new distance for scale-invariant 3D shape recognition and registration*, in Proceedings of the IEEE International Conference on Computer Vision, IEEE, Washington, DC, 2011, pp. 145–152.

[141] F. Porikli, O. Tuzel, and P. Meer, *Covariance tracking using model update based on Lie algebra*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, IEEE, Washington, DC, 2006, pp. 728–735.

[142] M. M. Postnikov, *The Variational Theory of Geodesics*, Dover, New York, 1967.

[143] M. M. Postnikov, *Geometry* VI: *Riemannian Geometry*, Springer, New York, 2001.

[144] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 2007.

[145] G. T. Pripoae, *Vector fields dynamics as geodesic motion on Lie groups*, Comptes Rendus Mathematique, 342 (2006), pp. 865–868.

[146] I. U. Rahman, I. Drori, V. C. Stodden, D. L. Donoho, and P. Schröder, *Multiscale representations for manifold-valued data*, Multiscale Model. Simul., 4 (2005), pp. 1201–1232.

[147] A. Roche, X. Pennec, G. Malandain, and N. Ayache, *Rigid registration of 3-D ultrasound with MR images: A new approach combining intensity and gradient information*, IEEE Trans. Medical Imag., 20 (2001), pp. 1038–1049.

[148] K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, *Landmark-based elastic registration using approximating thin-plate splines*, IEEE Trans. Medical Imag., 20 (2001), pp. 526–534.

[149] W. Rossmann, *Lie Groups: An Introduction through Linear Groups*, Oxford University Press, Oxford, UK, 2002.

[150] A. Saccon, J. Hauser, and A. P. Aguiar, *Optimal control on Lie groups: The projection operator approach*, IEEE Trans. Automat. Control, 58 (2013), pp. 2230–2245.

[151] Y. L. Sachkov, *Control theory on Lie groups*, J. Math. Sci., 156 (2009), pp. 381–439.

[152] S. Said, N. Courty, N. Le Bihan, and S. J. Sangwine, *Exact principal geodesic analysis for data on SO(3)*, in Proceedings of the European Signal Processing Conference, EURASIP, 2007, pp. 1700–1705.

[153] D. Salomon, *Transformations and Projections in Computer Graphics*, Springer, New York, 2006.

[154] D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, Springer-Verlag, New York, 1986.

[155] T. W. Sederberg and S. R. Parry, *Free-form deformation of solid geometric models*, SIGGRAPH Comput. Graphics, 20 (1986), pp. 151–160.

[156] J. A. Sethian, *Fast marching methods*, SIAM Rev., 41 (1999), pp. 199–235.

[157] K. Shoemake, *Animating rotation with quaternion curves*, SIGGRAPH Comput. Graphics, 19 (1985), pp. 245–254.

[158] D. Skerl, B. Likar, and F. Pernus, *A protocol for evaluation of similarity measures for rigid registration*, IEEE Trans. Medical Imag., 25 (2006), pp. 779–791.

[159] S. T. Smith, *Optimization techniques on Riemannian manifolds*, in Hamiltonian and Gradient Flows, Algorithms and Control, Fields Inst. Commun. 3, AMS, Providence, RI, 1994, pp. 113–136.

[160] S. Sommer, F. Lauze, and M. Nielsen, *Optimization over geodesics for exact principal geodesic analysis*, Adv. Comput. Math., 40 (2014), pp. 283–313.

[161] S. Stahl, *A Gateway to Modern Geometry: The Poincaré Half-Plane*, Jones and Bartlett, Burlington, MA, 2007.

[162] R. Subbarao and P. Meer, *Nonlinear mean shift over Riemannian manifolds*, Internat. J. Comput. Vis., 84 (2009), pp. 1–20.

[163] V. Surazhsky, T. Surazhsky, D. Kirsanov, S. J. Gortler, and H. Hoppe, *Fast exact and approximate geodesics on meshes*, ACM Trans. Graphics, 24 (2005), pp. 553–560.

[164] K. Tapp, *Matrix Groups for Undergraduates*, AMS, Providence, RI, 2005.

[165] D'A. W. Thompson, *On Growth and Form: The Complete Revised Edition*, Dover, New York, 1992.

[166] O. Tuzel, F. Porikli, and P. Meer, *Learning on Lie groups for invariant detection and tracking*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, Washington, DC, 2008, pp. 1–8.

[167] O. Tuzel, R. Subbarao, and P. Meer, *Simultaneous multiple 3D motion estimation via mode finding on Lie groups*, in Proceedings of the IEEE International Conference on Computer Vision, Vol. 1, IEEE, Washington, DC, 2005, pp. 18–25.

[168] C. Vaccaro, *Heat Kernel Methods in Finance: The SABR Model*, preprint, arXiv, http://arxiv.org/abs/1201.1437 (2012).

[169] B. Van Brunt, *The Calculus of Variations*, Springer, New York, 2004.

[170] 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 J. Numer. Anal., 33 (2013), pp. 481–514.

[171] J. Vince, *Quaternions for Computer Graphics*, Springer, New York, 2011.

[172] H.-C. Wang, *Discrete nilpotent subgroups of Lie groups*, J. Differential Geom., 3 (1969), pp. 481–492.

[173] R. P. Woods, *Characterizing volume and surface deformations in an atlas framework: Theory, applications, and implementation*, NeuroImage, 18 (2003), pp. 769–788.

[174] R. P. Woods, S. T. Grafton, J. D. G. Watson, N. L. Sicotte, and J. C. Mazziotta, *Automated image registration:* II. *Intersubject validation of linear and nonlinear models*, J. Comput. Assisted Tomography, 22 (1998), pp. 153–165.

[175] M. Wüstner, *A connected Lie group equals the square of the exponential image*, J. Lie Theory, 13 (2003), pp. 307–309.

[176] L. Younes, *Computable elastic distances between shapes*, SIAM J. Appl. Math., 58 (1998), pp. 565–586.

[177] L. Younes, *Shapes and Diffeomorphisms*, Springer-Verlag, New York, 2010.

[178] E. Zacur, M. Bossa, and S. Olmos, *Multivariate tensor-based morphometry with a right-invariant Riemannian distance on GL+(n)*, J. Math. Imaging Vis., doi:10.1007/s10851-013-0479-7 (2013).

[179] M. Žefran and V. Kumar, *Interpolation schemes for rigid body motions*, Comput. Aided Design, 30 (1998), pp. 179–189.

[180] M. Žefran, V. Kumar, and C. Croke, *Metrics and connections for rigid-body kinematics*, Internat. J. Robotics Res., 18 (1999), p. 242.

[181] M. Žefran, V. Kumar, and C. B. Croke, *On the generation of smooth three-dimensional rigid body motions*, IEEE Trans. Robotics Automation, 14 (1998), pp. 576–589.