INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Bi-invariant Means in Lie Groups. Application to Left-invariant Polyaffine Transformations.

Vincent Arsigny — Xavier Pennec — Nicholas Ayache

N° 5885

April, 2006.

Thème BIO

Rapport de recherche

ISSN 0249-6399

Bi-invariant Means in Lie Groups. Application to Left-invariant Polyaffine Transformations.

Vincent Arsigny, Xavier Pennec , Nicholas Ayache

Thème BIO — Systèmes biologiques
Projets Epidaure

Rapport de recherche n° 5885 — April, 2006. — 49 pages

Abstract: In this work, we present a general framework to define rigorously a novel type of mean in Lie groups, called the *bi-invariant mean*. This mean enjoys many desirable *invariance properties*, which generalize to the non-linear case the properties of the arithmetic mean: it is invariant with respect to left- and right-multiplication, as well as inversion. Previously, this type of mean was only defined in Lie groups endowed with a bi-invariant Riemannian metric, like compact Lie groups such as the group of rotations. But Riemannian bi-invariant metrics do not always exist. In particular, we prove in this work that such metrics do not exist in any dimension for *rigid transformations*, which form but the most simple Lie group involved in bio-medical image registration.

To overcome the lack of existence of bi-invariant Riemannian metrics for many Lie groups, we propose in this article to define bi-invariant means in *any finite-dimensional real Lie group* via a general *barycentric equation*, whose solution is by definition the bi-invariant mean. We show the existence and uniqueness of this novel type of mean, provided the dispersion of the data is small enough, and the convergence of an efficient iterative algorithm for computing this mean has also been shown. The intuition of the existence of such a mean was first given by R.P.Woods (without any precise definition), along with an efficient algorithm for computing it (without proof of convergence), in the case of *matrix* groups.

In the case of rigid transformations, we give a simple criterion for the general existence and uniqueness of the bi-invariant mean, which happens to be the same as for rotations. We also give closed forms for the bi-invariant mean in a number of simple but instructive cases, including 2D rigid transformations. Interestingly, for general linear transformations, we show that similarly to the Log-Euclidean mean, which we proposed in recent work, the bi-invariant mean is a generalization of the (scalar) geometric mean, since the determinant of the bi-invariant mean is exactly equal to the geometric mean of the determinants of the data.

Last but not least, we use this new type of mean to define a novel class of polyaffine transformations, called *left-invariant polyaffine*, which allows to fuse local rigid or affine

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

components *arbitrarily far away from the identity*, contrary to Log-Euclidean polyaffine fusion, which we have recently introduced.

Key-words: Bi-invariant means, Fréchet means, Statistics, Lie groups, Riemannian geometry, Polyaffine transformations

Moyennes bi-invariantes dans les groupes de Lie. Application aux transformations polyaffines invariantes à gauche.

Résumé : Dans ce travail, nous présentons un cadre général pour définir rigoureusement un nouveau type de moyenne dans des groupes de Lie, appelé *moyenne bi-invariante*. Cette moyenne possède de nombreuses *propriétés d'invariance*, qui généralisent au cas non-linéaire les propriétés de la moyenne arithmétique: cette moyenne est invariante par multiplication à gauche, à droite, ainsi que par inversion. Précédemment, ce type de moyenne avait seulement été défini dans des groupes de Lie dotés d'une métrique riemannienne bi-invariante, comme les groupes de Lie compacts tels que le groupe de rotations. Mais une métrique riemannienne bi-invariante n'existe pas toujours. En particulier, nous prouvons dans cet article qu'une telle métrique n'existe jamais pour *pour les transformations rigides*, qui forment pourtant dans le groupe de Lie le plus simple impliqué dans le recalage d'images bio-médicales.

Pour surmonter l'absence d'existence d'une métrique riemannienne bi-invariante pour de nombreux de groupes de Lie, nous proposons dans cet article de définir une moyenne bi-invariante dans *tout groupe de Lie réel de dimension finie* par l'intermédiaire d'une *équation barycentrique* très générale, dont la solution est par définition la moyenne bi-invariante. Nous montrons dans cet article l'existence et l'unicité de cette nouvelle moyenne, dans le cas où la dispersion des données est assez petite ; nous montrons également la convergence d'un algorithme itératif efficace permettant de calculer cette moyenne en pratique. L'intuition de l'existence d'un tel type de moyenne a été fournie originellement par R.P.Woods (sans toutefois de définition précise de cette moyenne), ainsi que l'algorithme permettant de la calculer efficacement (sans preuve de convergence), dans le cas particulier *des groupes de matrices.*

Dans le cas des transformations rigides, nous donnons un critère simple et général pour l'existence et l'unicité de la moyenne bi-invariante, qui s'avère être le même que pour les rotations. Dans un certain nombre de cas simples mais instructifs, nous donnons également la forme analytique prise par la moyenne bi-invariante, en particulier pour les transformations rigides 2D. Pour les transformations linéaires générales, nous montrons que de manière similaire aux moyennes log-euclidiennes, que nous avons proposées dans des travaux récents, la moyenne bi-invariante est une généralisation aux transformations linéaires inversibles de la moyenne géométrique (scalaire), puisque le déterminant de la moyenne bi-invariante est exactement égale à la moyenne géométrique des déterminants des données.

Enfin, nous utilisons ce nouveau type de moyenne afin de définir une classe nouvelle de transformations de polyaffines, appelée *polyaffines invariantes à gauche*, qui permet de fusionner des composantes locales rigides ou affines *arbitrairement lointaines de l'identité*, contrairement à la fusion log-euclidienne polyaffine, que nous avons récemment proposée.

Mots-clés : Moyennes bi-invariantes, moyennes de Fréchet, statistiques, groupes de Lie, géometrie riemannienne, transformations polyaffines

Contents

1 Introduction

In recent years, the need for rigorous frameworks to compute statistics in non-linear spaces has grown considerably in the bio-medical imaging community. First, a number imaging modalities, like diffusion MRI (or dMRI) [1, 2, 3], provide researchers with data which do not live in a linear space, and nonetheless require post-processing (re-sampling, regularization, statistics, etc.). See for instance [4] for examples of Riemannian statistics on diffusion tensors and [5] for statistics on rigid transformations, in the context of the analysis of the statistical properties of the human scoliosis. Second, the one-to-one registration of bio-medical images naturally deals with data living in non-linear spaces, since many types of invertible geometrical deformations belong to *groups* of transformations, which are not vector spaces. These groups can be finite-dimensional, as in the case of rigid or affine transformations, or infinite-dimensional as in the case of groups of diffeomorphisms parameterized with time-varying speed vector fields [6].

Among statistics, the most fundamental is certainly the *mean*, which extracts from the data a central point, minimizing in some sense the dispersion of the data around it. In this paper, we focus on the generalization of the *Euclidean mean* to Lie Groups, which are a large class of non-linear spaces with relatively nice properties. Classically, in a Lie group endowed with a Riemannian metric, the natural choice of mean is called the *Fréchet mean* [7]. But this Riemannian approach is completely satisfactory *only* if a bi-invariant metric exists, which is for example the case for compact groups such as rotations [7, 8]. The bi-invariant Fréchet mean enjoys many desirable *invariance properties*, which generalize to the non-linear case the properties of the arithmetic mean: it is invariant with respect to left- and right-multiplication, as well as inversion. Unfortunately, bi-invariant Riemannian metrics do not always exist. In particular, in this work, we prove the novel result that such metrics do not exist in any dimension for *rigid transformations*, which form but the most simple Lie group involved in bio-medical image registration.

To overcome the lack of existence of bi-invariant Riemannian metrics for many Lie groups, we propose in this article to define a *bi-invariant mean* generalizing the Fréchet mean induced by bi-invariant metrics, even in cases when such metrics *do not* exist. The intuition of the existence of such a mean was actually first given in [9] (without any precise definition), along with an efficient algorithm for computing it (without proof of convergence), in the case of *matrix* groups.

In this work, we present a general framework to define rigorously bi-invariant means, this time in *any* finite dimensional real Lie group. To do this, we rely on a general *barycentric equation*, whose solution is by definition the bi-invariant mean. We show the existence and uniqueness of this novel type of mean, provided the dispersion of the data is small enough, and the convergence of the classical iterative algorithm of [9] is also shown.

In the case of rigid transformations, we have been able to determine a simple criterion for the general existence and uniqueness of the bi-invariant mean, which happens to be the same as for rotations. We also give closed forms for the bi-invariant mean in a number of simple but instructive cases, including 2D rigid transformations. Interestingly, for general linear transformations, we show that similarly to the Log-Euclidean mean, that we recently proposed in [10], the bi-invariant mean is a generalization of the (scalar) geometric mean, since the determinant of the bi-invariant mean is exactly equal to the geometric mean of the determinants of the data.

Last but not least, this new type of mean is used to define a novel class of polyaffine transformations, called *left-invariant polyaffine*, which allows to fuse local rigid or affine components *arbitrarily far away from the identity*, contrary to Log-Euclidean polyaffine fusion, which we recently introduced in [11].

The sequel of this article is organized as follows. First, we recall the fundamental properties of Lie groups and of invariant Riemannian metrics in these spaces, and prove that bi-invariant Riemannian metrics do not exist for rigid transformations. Then, we detail the properties of the group exponential and logarithm in Lie groups. In the next Section, we rely on these properties to obtain a novel definition of bi-invariant means in any finite-dimensional real Lie groups, along with a proof of its existence and uniqueness and we also prove the convergence of the efficient iterative scheme proposed in [9] to compute this mean in practice. Then, we explicit the form taken by the bi-invariant mean in a number of simple cases where a closed form exists for this mean, e.g. the Heisenberg group. Afterwards, we focus on linear transformations, and in particular on rigid transformations and tensors. Last but not least, we rely on bi-invariant means to define a novel class of polyaffine transformations, called *left-invariant polyaffine*, which allows to fuse local rigid or affine components *arbitrarily far away from the identity.*

2 Means in Lie Groups

Notations. In the sequel of this article, we will use a number of notations, which are listed below. We begin with notations for usual matrix groups and submanifolds:

- $GL(n)$ is the group of real invertible $n \times n$ matrices, and more generally, for any (finite dimensional) vector space E, $GL(E)$ will be the group of invertible linear operations acting on E.
- $SL(n)$ is the *special linear group*, i.e. the subgroup of matrices of $GL(n)$ whose determinant is equal to 1.
- $O(n)$ is the group of *orthogonal* transformations, i.e. square matrices satisfying $R.R^T = Id$, where Id is the identity matrix and R^T is the transposed matrix of R.
- $SO(n)$ is the *special orthogonal group*, better known has the group of *rotations*. It is the subgroup of $O(n)$ whose elements satisfy $det(R) = 1$.
- $SE(n)$ is the group of *special Euclidean* transformations, i.e. the group of rigid displacements.
- $M(n)$ is the space of real $n \times n$ square matrices.
- $Sym_\star^+(n)$ is the space of symmetric positive-definite real $n \times n$ matrices.

- $Sym(n)$ is the vector space of real $n \times n$ symmetric matrices.

General (abstract) Lie group notations:

- when $\mathcal{G}$ is a Lie group, its neutral element will be written e, and a typical element of $\mathcal{G}$ will be m. The Lie algebra of $\mathcal{G}$ will be written $\mathfrak{g}$.
- the tangent space of $\mathcal{G}$ at point m will be referred to as $T_m\mathcal{G}$, which can be intuitively thought of as the linear space 'best approximating' $\mathcal{G}$ around m.
- we denote L_m (resp. R_m) the left- (resp. right-)multiplication by an element $m \in \mathcal{G}$. Furthermore, we will let $Inv : \mathcal{G} \rightarrow \mathcal{G}$ be the inversion operator.
- if a mapping $\Phi : \mathcal{G} \rightarrow \mathcal{G}$ is differentiable, we write $D_m\Phi$ its *tangent map* (or differential map) at m. $D_m\Phi$ is a mapping from $T_m\mathcal{G}$ to $T_{\Phi(m)}\mathcal{G}$, which means that to a tangent vector located at m (which is basically an infinitesimal displacement) it associates a tangent vector at $\Phi(m)$ (another infinitesimal displacement, 'living' in a different vector space).

2.1 Lie Groups

Definition of Lie groups. We start by recalling the basic properties of Lie groups, along with the convenient notions which are classically used to describe these properties. Typical examples of such groups are groups of geometrical transformations (e.g., rigid or affine transformations), where the multiplication is the composition of mappings.

In simple terms, a Lie group is first a group in the *algebraic* sense, i.e. a set of elements in which a *multiplication* between elements is defined. This multiplication is assumed to have neat and intuitive properties: it is associative ($(a.b).c = a.(b.c)$), it has a neutral element e and each element a has a unique inverse a^{-1}.

Second, a Lie group has a structure of (smooth, i.e. $\mathcal{C}^\infty$) *differential manifold*. This means that it is locally similar to a vector space, but can be quite 'curved' globally.

Third, the algebraic and differential are structures *compatible*: inversion and (left- and right-) multiplications are *smooth* mappings. This means that it is possible to indefinitely differentiate them. For more formal definitions and more details, please refer to classical differential geometry books like See [12] or [13].

Examples. Many usual sets can be viewed as Lie Groups. Namely:

- vector spaces
- *multiplicative matrix groups*: $GL(n)$, $O(n)$, $SO(n)$, etc.
- *Geometric transformation groups* such as rigid transformations, similarities, affine transformations... which can anyway also be looked upon as matrix groups via their 'faithful' representation based on homogeneous coordinates.

- some infinite-dimensional *Lie groups of diffeomorphisms* have also been recently gaining much importance in computational anatomy [6]. Due to the high level of technicality needed to deal with such infinite-dimensional spaces, we will not consider this type of Lie groups in this work.

Lie Algebra and Adjoint Representation. We will need a number of notions classically used to describe the properties of Lie groups. They are the following:

- To any vector $v \in T_e\mathcal{G}$, i.e. to any tangent vector to $\mathcal{G}$ at the identity can be associated in a one-to-one manner a *left-invariant* vector field X_v, defined by $X_v(m) = D_eL_m.v$, i.e. simply by left-multiplying v.

- we can give $(T_e\mathcal{G}, +, .)$, which is by construction a *vector space*, a structure of *Lie algebra*, i.e. give it an extra algebraic operation which is an associative and bi-linear inner product called the *Lie bracket*, denoted here $[\,,]$. This operation closely reflects the multiplicative properties of the group $\mathcal{G}$. In particular, for commutative Lie groups, the Lie bracket is always null.

  This inner product is actually derived from *general* Lie bracket on smooth vector fields, since $(T_e\mathcal{G}, +, ., [\,,])$ can be considered as a *Lie subalgebra* of vector fields on $\mathcal{G}$, since we can identify $T_e\mathcal{G}$ with the set of left-invariant vector fields.

  As mentioned previously, the notation for the Lie algebra of $\mathcal{G}$ will be in this article $\mathfrak{g}$. It has a number of remarkable algebraic properties (in addition to its associativity and bi-linearity) which are the following:

  i) $\forall a, b \in \mathfrak{g},\ [a, b] = -[b, a]$ ('anti-commutativity'), which implies $[a, a] = 0$
  ii) $\forall a, b, c \in \mathfrak{g},\ [a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$ (Jacobi identity).

  Simple examples of Lie brackets is are given by $GL(n)$ and its multiplicative subgroups, like $SL(n)$ or $SO(n)$. In these cases, the Lie algebra is a *vector space of square matrices*, and the Lie bracket between two elements M and N of this algebra is the *commutator* of these two matrices, i.e. $[M, N] = M.N - N.M$. In particular, the Lie algebra of $GL(n)$ is $M(n)$, that of $SL(n)$ is the subvector space of $M(n)$ of matrices with a trace equal to zero, and the Lie algebra of $SO(n)$ is the vector space of skew symmetric matrices. For a complete account on Lie Algebras, see [14].

- $\mathcal{G}$ can be 'represented' by a group of matrices acting on $\mathfrak{g}$, via what is called its *adjoint representation*, $Ad(\mathcal{G})$. We will see that the properties of this representation and the existence of bi-invariant metrics for the group $\mathcal{G}$ are highly linked.

  This means that one can map each element of the group into a *linear operator* (i.e., a matrix) which acts on the Lie algebra. More precisely, an element m of $\mathcal{G}$ acts on an element v of $\mathfrak{g}$ by $Ad(m).v =$ '$m.v.m^{-1}$' [1]. This operation is called a representation in the sense of *representation theory* (see [15] for a complete treatment), which means that this mapping is *compatible* with the Lie group structure of $\mathcal{G}$. This compatibility consists of the following properties:

i) $Ad(e) = Id$

ii) $\forall m \in \mathcal{G},\ Ad(m^{-1}) = Ad(m)^{-1}$

iii) $\forall m, n \in \mathcal{G},\ Ad(m.n) = Ad(m).Ad(n)$

iv) $Ad : \mathcal{G} \to GL(\mathfrak{g})$ is smooth.

This amounts to saying that Ad is a *smooth group homomorphism* (or Lie group homomorphism).

2.2 Means and Algebraic Invariance

Lie groups are *not* vector spaces in general but have a more complicated structure: instead of a (commutative) addition '+' and a scalar multiplication '.', they only have a (non-commutative in general) multiplication '$\times$' and an inversion operator (which corresponds to the scalar multiplication by -1 for vector spaces).

How should one generalize the notion of *mean* to this type of non-linear space? To do so, one can rely on the *invariance properties* that the mean should *a priori* satisfy, in order to generalize the invariance properties of the arithmetic mean in vector spaces.

Indeed, in the case of vector spaces, the arithmetic means presents strong invariance properties: invariance with respect to any translation and with respect to any multiplication by a scalar. This means that the arithmetic mean is invariant with respect to all the algebraic operations induced by the vector space structure. It makes good sense that the notion of mean and the algebraic structure should be *compatible*.

In the case of groups, the invariance with respect to left- and right-multiplications (the group can be non-commutative) and the inversion operator are the equivalent of the invariance properties associated to the mean in vector spaces. When we translate a given set of samples or a probability measure, it is reasonable to wish that their mean be translated exactly in the same way, and the same property is desirable when we take the inverses of the samples.

Example 1. The Geometric Mean of positive numbers. *We can give to the set of positive numbers a structure of commutative group with the usual multiplication in $\mathbb{R}$. In this context, let (x_i) be N positive numbers and (w_i) be N non-negative normalized weights*

[1] To be completely rigorous, one has to resort to the (more complicated) differentials of left- and right-multiplication. This yields: $Ad(m).v =$ ' $m.v.m^{-1}$ ' $= D_{m^{-1}}L_m.D_eR_{m^{-1}}.v = D_mR_{m^{-1}}.D_eL_m.v$ by associativity of the group multiplication. In the matrix case, we *do* have the (simple this time) formula: $Ad(R).M = R.M.R^{-1}$, which only uses two matrix multiplications and one matrix inversion.

($\sum_i w_i = 1$). The arithmetic mean of the data is invariant with respect to multiplication, but not with respect to inversion:

$$\Sigma_i w_i \frac{1}{x_i} \neq \frac{1}{\Sigma_i w_i x_i}, \text{ in general.} \tag{1}$$

Thus the arithmetic mean is not fully adapted to this multiplicative structure. On the contrary, the geometric mean, written here $\mathbb{E}((x_i),(w_i))$, is fully adapted. It is given by:

$$\mathbb{E}((x_i),(w_i)) = \exp(\Sigma_i w_i \log(x_i)). \tag{2}$$

We recall the classical convexity inequality between the two means:

$$\exp(\Sigma_i w_i \log(x_i)) < \Sigma_i w_i x_i, \tag{3}$$

whenever the data is not reduced to a single point.

How Can We Define Invariant Means? A classical approach to define a notion of mean compatible with algebraic operations is to define first a *distance* (or metric) compatible with these operations and then to rely on this distance to define the mean.

Means and Distances. For a short time, let us now consider the problem of defining means in the general setting of *metric spaces*, i.e. the sets on which a *distance* is defined. To make the notion of mean *compatible* with the metric, one can rely on the intuitive idea of *minimal variance or dispersion* to define the mean, because a metric provides a way of quantifying how close (or far away) two elements are from each other in a metric space $(E, dist)$. More precisely, the mean can be defined as the point $\mathbb{E}(X)$ which minimizes some kind of dispersion dispersion of the data $(X_i)_{i=1}^N$ around itself (with respect to some non-negative nomalized weights (w_i)), for example:

$$\mathbb{E}(X_i) = \underset{Y \in E}{arg\ min} \sum_i w_i.dist(X_i, Y)^\alpha. \tag{4}$$

The case $\alpha = 2$ corresponds in vector spaces to the arithmetic mean, and in other spaces to their generalization, called the *Fréchet* mean [7]. For $\alpha = 1$, one obtains the generalization of the *median*. One should note that the dispersion may have several minimizers. For instance, this is classically the case in vector spaces for $\alpha = 1$, essentially because in this case the dispersion it is not strictly convex, contrary to the case $\alpha > 1$. Even when $\alpha > 1$, the dispersion of the data X_i should not be too high in order to guarantee that the dispersion has a unique global minima [7].

2.3 Bi-invariant Fréchet Means via Invariant Metrics in Lie Groups

In the case of Lie groups, we will see here how one can (or cannot) define a distance compatible with algebraic operations *and* the differentiable structure of these groups.

Riemannian Metrics and Geodesics. The distances (or metrics) *compatible* with the differentiable structure of differentiable manifolds are called *Riemannian metrics.*

Basically, the idea is to define *smoothly* in each tangent space $T_m\mathcal{G}$ a *scalar product* $< .,. >_{T_m\mathcal{G}}$. The distance between two points m and n is then obtained by computing the minimal length of a smooth curve $c(t)$ joining them in one unit of time. We recall that the length $l(c)$ of c is classically given by:

$$l(c) = \int_0^1 \left\| \frac{dc}{dt}(t) \right\|_{T_{c(t)}\mathcal{G}} dt.$$

A smooth curve of minimal length between two points is called a *geodesic*, and if there is a unique geodesic between two points, this curve is called a *minimizing geodesic.* Interestingly, for any given point in a smooth manifold endowed with a Riemannian metric, there exists an *open neighborhood* of this point which is *geodesically convex*, i.e. where any couple of points can be joined by a minimizing geodesic. When two points are 'far' apart, there can fail to be any geodesic between them (think of a set with several connected components) or on the contrary several geodesics (possibly an infinity) can join these points (think of antipodal points on a sphere). For details on the existence and possible uniqueness of geodesics, see for example [13].

Invariance Properties of Riemannian Metrics. Let us now detail the different types of invariance (or compatibility) that can exist between a Riemannian metric on a Lie group and its algebraic properties. They are the following:

- 'left-invariance': the metric is invariant with respect to any multiplication on the left. Another useful way of phrasing this is to say that left-multiplications are *isometries* of $\mathcal{G}$, i.e. do not change distances between elements of $\mathcal{G}$.

  In terms of scalar products and differentials, this means precisely that for any two points m and h of $\mathcal{G}$ and any vectors v and w of $T_m\mathcal{G}$, we have: $< D_m L_h.v, D_m L_h.w >_{T_{h.m}\mathcal{G}} = < v, w >_{T_m\mathcal{G}}$.

- 'right-invariance': invariance with respect to any multiplication on the right.

- 'inversion-invariance': invariance with respect to inversion. The inversion operator is then an isometry of $\mathcal{G}$.

These properties are *not* independent. This simply comes from the fact that for any two elements m, n of $\mathcal{G}$, we have $(m.n)^{-1} = n^{-1}.m^{-1}$. This implies for example that the left-multiplication can be obtained smoothly from one right-multiplication and two inversions in the following way: $L_m = Inv \circ R_{m^{-1}} \circ Inv$.

A simple (but rarely mentioned in classical references on Lie groups) consequence of this is that *all* right-invariant metrics can be obtained from left-invariant metrics by 'inversion', and *vice versa.* Indeed, we have:

Proposition 1. *Let $<,>$ be a left-invariant Riemannian metric defined on $\mathcal{G}$. Then the 'inverted' metric $\ll,\gg$, defined below, is right-invariant, and moreover we have $<,>_e = \ll,\gg_e$.*

For any two points m and h of $\mathcal{G}$ and any vectors v and w of $T_m\mathcal{G}$, we define the inverted metric $\ll,\gg$ as follows:

$$\ll v, w \gg_m \stackrel{def}{=} < D_m Inv.v, D_m Inv.w >_{T_{m^{-1}}\mathcal{G}} .$$

Proof. Actually, the proof relies only on differentiating the equality $(h.m)^{-1} = m^{-1}.h^{-1}$. This yields:

$$D_{h.m}Inv \circ D_m L_h = D_{m^{-1}} R_{h^{-1}} \circ D_m Inv.$$

This allows to show directly that:

$$\ll D_m R_h.v, D_m R_h.w \gg_{T_{h.m}\mathcal{G}} = \ll v, w \gg_{T_m\mathcal{G}},$$

which means that $\ll,\gg$ is right-invariant.

Last but not least, the equality $<,>_e = \ll,\gg_e$ comes from the fact that quite intuitively $D_e Inv = -Id$, where Id is the identity operator in $T_e\mathcal{G}$. This can be easily seen from the classical result $D_e \exp = D_0 \log = Id$ and the equality (valid in an open neighborhood of e) $m^{-1} = \exp(-\log(m))$, where exp and log are the group exponential and logarithm, presented in detail in the sequel of this article. □

Corollary 1. *'Left-invariance' (resp. 'right-invariance') and 'inversion-invariance' imply 'right-invariance' (resp. 'left-invariance').*

Proof. We have just seen in Proposition 1 that right-invariant metrics can be obtained by composition between left-invariant metrics and the inversion operator, and vice versa for left-invariant metrics. If left-multiplications and inversion are isometries, so are right-multiplications by composition. □

Riemannian metrics which are simultaneously left- and right- invariant are called *bi-invariant*. On these special metrics, we have the very interesting result:

Theorem 1. *Bi-invariant metrics have the following properties:*

1. *A bi-invariant metric is also invariant w.r.t. inversion*
2. *It is bi-invariant if and only if $\forall m \in \mathcal{G}, Ad(m)$ is an isometry of the Lie algebra $\mathfrak{g}$*
3. *One-parameter subgroups of $\mathcal{G}$ are geodesics for the bi-invariant metric*

Proof. See [12], chapter V. □

From this result and Proposition 1, we see that *any two invariance properties imply the third.*

Bi-invariant Means. We have seen that a metric structure induces a notion of mean called the Fréchet mean. The Fréchet mean associated to a bi-invariant metric is called the *bi-invariant mean*. Actually, it does not depend on the particular choice of bi-invariant metric, since whenever the bi-invariant mean is uniquely defined, it is given as the solution of a *barycentric equation* [7] which is independent from the arbitrary choice of bi-invariant metric.

Since the metric inducing the notion of mean is bi-invariant, so is the mean, which is then fully compatible with the algebraic properties of the Lie group. As a consequence, this notion of mean is particularly well-adapted to Lie groups [7]. However, contrary to left- or right- invariant metrics, which always exist [2] , bi-invariant metrics may fail to exist, and we will now see under which conditions bi-invariant metrics exist for a given Lie group.

Compactness of the Adjoint Representation From Theorem 1, we see that if a bi-invariant metric $< .,. >$ exists for the Lie group $\mathcal{G}$, then $\forall m \in \mathcal{G}$, $Ad(m)$ is an isometry of $\mathfrak{g}$ and can thus be looked upon as an element of the orthogonal group $O(n)$ where $n = dim(\mathcal{G})$. Then, note that $O(n)$ is a *compact* group, and that therefore $Ad(\mathcal{G})$ is necessarily included in a compact set, a situation called *relative compactness*. This notion provides indeed an excellent criterion, since we have:

Theorem 2. *The Lie group $\mathcal{G}$ admits a bi-invariant metric if and only if its adjoint representation $Ad(\mathcal{G})$ is relatively compact.*

Proof. We have already seen the first implication. For the converse part, the theory of differential forms and their integration can be used to explicitly construct a bi-invariant metric. This is done in [12]B, Theorem V.5.3. □

Compactness, Commutativity and Bi-invariant Metrics. In the case of *compact* Lie groups, we have the property that their adjoint representation is the image of a compact set by a continuous mapping and is thus also compact. Then, Theorem 2 implies that bi-invariant metrics exist in such a case. In particular, this is the case of *rotations*, for which bi-invariant means have been extensively studied and used [7]. This is also trivially the case of *commutative* Lie groups, where only one type of multiplication exists, which reduces the adjoint representation to $\{Id\}$. An illustration of this situation is given by the Lie group structure on symmetric positive-definite matrices we recently proposed in [16, 17, 18].

As shown by Theorem 2, the general non-compact and non-commutative case is not so nice, and one has to carefully check the properties of the adjoint representation of the Lie group to see whether a bi-invariant metric exists or not. This verification has to be done all the more carefully that non-commutativity and non-compactness are necessary but *not* sufficient to prevent the existence of bi-invariant metrics, as shown in the following paragraph.

[2] It suffices to propagate an arbitrary scalar product defined on $T_e\mathcal{G}$ to all tangent spaces by left- or right-multiplication to generate all left- or right-invariant Riemannian metrics.

An Example in the Non-Compact and Non-Commutative Case. We have already seen that any compact or commutative Lie group has at least one bi-invariant metric. From this remark, one can easily construct an example of non-compact and non-commutative group having a bi-invariant metric: let $\mathcal{G}_1$ be a commutative non-compact group and $\mathcal{G}_2$ be a compact non-commutative group. They both have a bi-invariant metric. Let $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$ be their *direct product*, i.e. the group obtained with the multiplication $(g_1', g_2').(g_1, g_2) = (g_1'.g_1, g_2'.g_2)$. Then $\mathcal{G}$ is neither commutative nor compact, but *has* a bi-invariant metric! In fact, let $< .,. >_1$ and $< .,. >_2$ be respectively a bi-invariant metric of $\mathcal{G}_1$ and $\mathcal{G}_2$. Then $< P_{\mathcal{G}_1}(.), P_{\mathcal{G}_1}(.) >_1 + < P_{\mathcal{G}_2}(.), P_{\mathcal{G}_2}(.) >_2$ is a bi-invariant metric of $\mathcal{G}$, where $P_{\mathcal{G}_i}$ is the canonical projection on $\mathcal{G}_i$.

One typical example of such a situation is the Lie group of matrices of the form $s.R$, where s is a positive scalar and R a rotation matrix (group of rotations and scalings). It can be seen as the direct product of $(\mathbb{R}_+^\star, \times)$ (commutative and non-compact) with $(SO(n), \times)$ (compact and non-commutative).

2.4 Absence of Bi-invariant Metrics for Rigid Transformations

As we have seen in the previous Subsection, bi-invariant metrics always exist for compact groups, which is the case of rotations. But when one tries to extend the use of bi-invariance metrics to more general transformation groups, one is very limited. In biomedical imaging, the simplest possible registration procedure between two anatomies uses rigid transformations. Such transformations seem quite close to rotations and one could hope for the existence of bi-invariant metrics. But we have the following general result, which is new to our knowledge[3]:

Proposition 2. *The action of the adjoint representation Ad of $SE(n)$ at the point (R, t) on an infinitesimal displacement (dR, dt) is given by:*

$$Ad(R,t).(dR,dt) = (R\, dR\, R^T,\ -R\, dR\, R^T\, t\ + R\, dt).$$

As a consequence, no bi-invariant Riemannian metric exists on the space of rigid transformations.

Proof. In the case of matrix Lie groups, we have the following formula [20] $Ad(h).dh = h.dh.h^{-1}$ for $dh \in \mathfrak{g}$. Classically, using homogeneous coordinates, the Lie group of rigid transformations is faithfully represented by the following matrix Lie group [21]:

$$(R,t) \sim \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}.$$

[3]We have recently found that the non-existence of bi-invariant Riemannian metrics for $SE(3)$ was already known in the literature [19]. However, our result does not depend on the dimension and is obtained in a very economical way, using short and abstract arguments rather than long and direct computations as in [19] for $n = 3$.

Using this, we get:

$$Ad(R,t).(dR,dt) \sim \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} . \begin{pmatrix} dR & dt \\ 0 & 0 \end{pmatrix} . \begin{pmatrix} R^T & -R^T.t \\ 0 & 1 \end{pmatrix},$$

which yields the announced formula. In this formula, the translation 't' introduces a unbounded term which prevents the adjoint group from being bounded. Applying Theorem 2, it is then clear that no bi-invariant metric exists for rigid transformations in nD (n>1). □

We thus see that the Riemannian approach based on bi-invariant metrics cannot be extended to rigid transformations, and even less so to affine transformations.

One should note that our result contradicts a statement in [9], which claimed that a bi-invariant metric existed when $n = 2$. The reference backing this claim was [21], in which it is only stated that though $SE(2)$ is non-compact, it has a bi-invariant *measure* (Chapter 7, page 92). But whereas the existence of a metric implies that of a measure (see [7], page 25; the measure can be thought of as the determinant of the metric), the existence of a measure does *not* imply the existence of a metric. This subtle mistake is of no consequence, since there truly are examples of non-compact groups which have bi-invariant metrics. As long as the group is commutative, such metrics obviously exist (think of vector spaces!). But here, the non-compactness *and* non-commutativity of $SE(n)$ forbid the existence of such a metric. Other simple illustrations of this phenomenon are given in Section 5, where we give more examples of non-compact and non-commutative Lie groups with no bi-invariant metrics.

Bi-invariant Means without Riemannian Metrics. In the sequel, we will see how it is possible to define general bi-invariant means in Lie groups *without* relying on bi-invariant Riemannian means, which can fail to exist. The key to our approach is to use the general *algebraic* properties of Lie groups, and in particular the group exponential and logarithm.

3 Fundamental Properties of the Exponential and Logarithm

Before defining general bi-invariant means in Lie groups, we detail in this Section the fundamental properties of the group exponential and logarithm. We will find these properties very useful in the sequel.

3.1 Matrix Exponential and Logarithm

The Matrix Exponential and Logarithm. Before we present the group exponential and logarithm in their full generality, let us recall the fundamental properties of the *matrix* exponential and logarithm, which correspond to the group exponential and logarithm of the Lie group of $n \times n$ invertible matrices, $GL(n)$. They are the generalization to matrices of the well-known scalar exponential and logarithm.

Definition 1. *The exponential* $\exp(M)$ *of a matrix* M *is given by* $\exp(M) = \sum_{n=0}^{\infty} \frac{M^k}{k!}$. *Let* $G \in GL(n)$. *If there exists* $M \in M(n)$ *such that* $G = \exp(M)$, *then* M *is said to be a logarithm of* N.

In general, the logarithm of a real invertible matrix may not exist, and if it exists it may not be unique. The lack of existence is a general phenomenon in connected Lie groups. One generally needs *two* exponentials to reach every element [22]. The lack of uniqueness is essentially due to the influence of rotations: rotating of an angle α is the same as rotating of an angle $\alpha + 2k\pi$ where k is an integer. Since the logarithm of a rotation matrix directly depends on its rotation angles (one angle suffices in 3D, but several angles are necessary when $n > 3$), it is not unique.

Principal Logarithm of a Matrix. When a real invertible matrix has no (complex) eigenvalue on the (closed) half line of negative real numbers, then it has a unique real logarithm whose (complex) eigenvalues have an imaginary part in $]-\pi, \pi[$ [23]. In this case this particular logarithm is well-defined and called *principal*. We will write $\log(M)$ for the principal logarithm of a matrix M whenever it is defined.

3.2 Lie Group Exponential

Let us now detail some of the fundamental properties of the group exponential and logarithm in Lie groups. For more details on these properties, see [24]. Basically, these properties are very similar to those of the matrix exponential and logarithm, which are a particular case of such mappings. One should note that this particular case is actually quite general, since most classical Lie groups can be looked upon as matrix Lie groups anyway [20]. But all Lie groups are *not* (at least directly) multiplicative matrix groups, as in the case of the Lie group structure we have recently proposed by symmetric positive-definite matrices [25]. This is the reason why we do not limit ourselves to the matrix case.

Definition 2. *Let* $\mathcal{G}$ *be a Lie group and let* v *be an tangent vector at the identity, i.e. an element of the Lie Algebra* $\mathfrak{g}$. *The group exponential of* v, *denoted* $\exp(v)$, *is given by the value at time* 1 *of the unique function* $g(t)$ *defined by the following ordinary differential equation (ODE):*

$$\begin{cases} \frac{dg}{dt} = D_e L_{g(t)}.v \\ g(0) = e. \end{cases} \tag{5}$$

Eq. (5) has particularly nice properties. $g(t)$ is in fact defined for all t, and yields a continuous *one-parameter subgroup* (also called one-parameter *Lie* subgroup), which means that $g(0) = e$, $g(t + t') = g(t).g(t') = g(t').g(t)$. v is called the *infinitesimal generator* of this subgroup. See [13], pages 27-29 for proofs of these properties. In fact, Eq. (5) is the equivalent of the matrix differential equation, which is a nice and classical *linear ODE*:

$$\begin{cases} \frac{dG}{dt} = G.V \\ G(0) = Id, \end{cases}$$

whose solution is well-known from classical ODE theory to be $G(t) = \exp(tV)$, where exp is the matrix exponential [26].

One-Parameter Subgroups vs. Group Exponential. We have just seen that for all v belonging to $\mathfrak{g}$, $\exp(t.v)$ is a one-parameter subgroup of $\mathcal{G}$: the additive subgroup $(t.V)_t$ of $\mathfrak{g}$ is mapped into a multiplicative subgroup of $\mathcal{G}$ by the exponential. Conversely, we have the interesting result that all continuous one-parameter subgroups of $\mathcal{G}$ are of this form ([12], Section V, Theorem 3.1, page 223). *This provides a simple way of computing the group exponential in situations where one-parameter subgroups are easy to obtain.*

The Exponential as a Local Diffeomorphism. Very much like the exponential map associated to a Riemannian metric, the group exponential is diffeomorphic locally around 0. More precisely, we have the following theorem:

Theorem 3. *The group exponential is a diffeomorphism from a open neighborhood of* 0 *in* $\mathfrak{g}$ *to a open neighborhood of* e *in* $\mathcal{G}$*, and its differential map at* 0 *is the identity.*

Proof. Since the exponential is a smooth mapping, the fact that its differential map is invertible at e allows for the use of the 'Implicit Function Theorem', which guarantees that it is a diffeomorphism from some open neighborhood of 0 to a open neighborhood of $\exp(0) = e$. For more details, see [13], page 28. □

This theorem implies that one can define *without ambiguity* a logarithm in a open neighborhood of e: for every g in this open neighborhood, there exists a unique v in the open neighborhood of 0 in $\mathfrak{g}$, such that $g = \exp(v)$. In the following, we will write $v = \log(g)$ for this logarithm, which is the (abstract) equivalent of the (matrix) *principal* logarithm.

Geodesic Convexity. Another useful property of the (metric) exponential map is that given any point, there exists a open neighborhood of this point, called *geodesically convex*, in which for any coupe of points, there exists a unique minimizing geodesic between them (see for example [13], page 84-85). We now prove an analogous result for the group exponential:

Theorem 4. *Let* $\Phi : \mathcal{G} \times \mathfrak{g} \to \mathcal{G} \times \mathcal{G}$*, defined by* $\Phi(g, v) = (g, g.\exp(v))$*. Then* Φ *is always locally diffeomorphic. More precisely, for all* g *in* G*, it defines a diffeomorphism from some open neighborhood of* $(g, 0)$ *to a open neighborhood of* (g, g)*.*

Proof. Since Φ is smooth, one can apply anew the 'Implicit Function Theorem' provided that the differential of Φ at $(g, 0)$ is invertible. To see this, note that we have:

$$\begin{cases} \frac{\partial \Phi}{\partial g}|_{(g,v)=(e,0)} = (Id, Id) \\ \frac{\partial \Phi}{\partial v}|_{(g,v)=(e,0)} = (0, D_e L_g \circ Id) = (0, D_e L_g), \end{cases}$$

where we have used the fact the property that the differential of the exponential at 0 is the identity (see Theorem 3). Since L_g is a diffeomorphism, its differential at e, $D_e L_g$, is always invertible. As a consequence, the differential of Φ is also always invertible, and the 'Implicit Function Theorem' applies. This proof is very similar to the proof given in [13] to show the analogous property of the metric exponential. $\square$

3.3 Group Geodesics.

Theorem 4 essentially shows that for every point g of G, there exists a open neighborhood of g in which every couple of points can be joined by a unique 'group geodesic' of the form $g.\exp(t.v)$ such that $g.\exp(v) = h$. By symmetry, the same result also holds for the geodesics of the type $\exp(t.v).g$. In fact, those two types of 'group geodesic' are the same, since we have the following result:

Theorem 5. *For all g in G, there exists a open neighborhood $\mathcal{V}_g$ of 0 in $\mathfrak{g}$ such that for all g in $\mathcal{G}$ and for all v, there exist a unique w in $\mathfrak{g}$ such that $g.\exp(t.v) = \exp(t.w).g$ for all $t \in \mathbb{R}$. More precisely, $w = Ad(g).v$. Moreover, in this open neighborhood of 0, the relationship $g.\exp(v) = \exp(w).g$ implies $w = Ad(g).v$.*

The proof of this theorem is simply based on the following relationships between the Adjoint representation, the exponential and the logarithm:

Lemma 1. *Let v be in $\mathfrak{g}$ and g in $\mathcal{G}$. Then we have:*

$$g.\exp(v).g^{-1} = \exp(Ad(g).v).$$

Also, for all g in $\mathcal{G}$, there exists a open neighborhood $\mathcal{V}_g$ of e such that for all m in $\mathcal{V}$: $\log(m)$ and $\log(g.m.g^{-1})$ are well-defined and are linked by the following relationship:

$$\log(g.m.g^{-1}) = Ad(g).\log(m).$$

These equations are simply the generalization to (abstract) Lie groups of the well-known matrix properties: $G.\exp(V).G^{-1} = \exp(G.V.G^{-1})$ and $G.\log(V).G^{-1} = \log(G.V.G^{-1})$.

Proof. The first relationship of this Lemma can be proved in the following way: $(g.\exp(t.v).g^{-1})_t$ is a continuous one-parameter subgroup, whose infinitesimal generator is $\frac{d}{dt}g.\exp(t.v).g^{-1}|_{t=0} = Ad(g).v$ (see Proposition 1.81 on page 29 of [13]). Using the fact that continuous one-parameter subgroups are of the form $\exp(t.w)$, we obtain the first equality. To prove the second (and this time *local*) equality, we see that since $\Psi_g : m \mapsto g.m.g^{-1}$ is smooth and $\Psi_g(e) = e$, there exists a open neighborhood of e where $\log(m)$ and $\log(g.m.g^{-1})$ are well-defined. Then the second equality is deduced from the first. $\square$

Proof. Proof of Theorem 5: just see that $g.\exp(t.v) = g.\exp(t.v).g^{-1}.g$ and apply Lemma 1.

$\square$

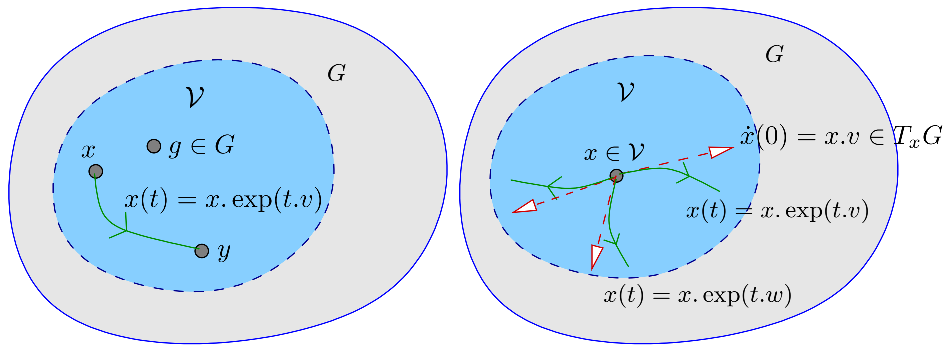

Figure 1: **Properties of Group Geodesics. Left:** group geodesic convexity, which means that for any point g of G, there exists a open neighborhood $\mathcal{V}$ of g, such that any couple of points x and y in $\mathcal{V}$ can be joined by a unique group geodesics of the form $x(t) = x.\exp(t.v)$ satisfying $x.\exp(v) = y$. Note that the geodesic is entirely contained in $\mathcal{V}$. **Right:** at any point g, a unique group geodesic is associated to any initial speed vector, provided this speed is small enough.

Definition of Group Geodesics. Essentially, we have just shown that the exponential and its translated versions can be looked upon as some sort of 'group geodesic' in a Lie group. Any couple of points can be joined by a unique 'group geodesic', provided they are close enough. This leads to the following definitions:

Definition 3. *Any continuous path of G of the form $g.\exp(t.v)$, which has the property of Theorem 5 (i.e. v is small enough) is called a* **group geodesic**. *Furthermore, an open set $\mathcal{O}$ of $\mathcal{G}$ is called* **groupwise geodesically convex** *(or GGC) if and only if any couple of points of $\mathcal{O}$ can be joined by a group geodesic. We have just shown that every g in G has a groupwise geodesically convex open neighborhood.*

To conclude this subsection, let us now present a last property for group geodesics, which generalizes the other well-known property of (metric) geodesics. The properties of group geodesics are illustrated in Fig. 1.

Proposition 3. *Let g be in G and z be in the tangent space at g. Then there exists a unique smooth path of the form $g.\exp(t.v)$ that $g(0) = g$ and $\frac{d}{dt}g|_{t=0} = z$. When z is small enough, this smooth path is a group geodesic.*

Proof. The only possible choice is $v = D_g L_{g^{-1}}.z$, since $D_g L_{g^{-1}}$ is always invertible. $\square$

3.4 Baker-Campbell-Hausdorff Formula

Before defining general bi-invariant means in Lie groups, let us focus on a last fundamental tool: the Baker-Campbell-Hausdorff formula (or BCH formula). Intuitively, this formula shows how much $\log(\exp(v).\exp(w))$ deviates from $v+w$ due to the (possible) non-commutativity of the multiplication in $\mathcal{G}$. Remarkably, this deviation can be expressed only in terms of Lie brackets between v and w [24]. We have already used in [25] to compare the traces of two different generalization of the geometric mean to symmetric positive-definite matrices.

Theorem 6. *Series form of the BCH formula ([24], Chapter VI). Let v, w be in $\mathfrak{g}$. Then they are small enough, we have the following development:*

$$\begin{array}{rcl} \log(\exp(v).\exp(w)) & = & v + w + 1/2([v, w]) \\ & & +1/12([v, [v, w]] + [w, [w, v]]) \\ & & +1/24([[v, [v, w]], w]) + O((\|v\| + \|w\|)^5). \end{array} \tag{6}$$

Following [24], let us write $H : \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ the mapping defined near 0 such that $H(v, w) = \log(\exp(v).\exp(w))$. A fundamental property of this function is the following: it is not only $\mathcal{C}^\infty$ but also *analytic* around 0, which means that $H(v, w)$ (near 0) is the sum of an absolutely converging multivariate infinite series (the usual multiplication is replaced here by the Lie bracket). This implies in particular that all the (partial) derivatives of this function are also analytic. We will use these very remarkable properties in the sequel of this article.

4 Bi-invariant Means in Lie Groups

4.1 A Geometric Definition of the Mean

Let us recall the classical definition of a mean in an *affine* space F, i.e. a space of *points*, associated to a *vector space* E such that to any couple of points M, N we can associate the vector $\overrightarrow{MN}$, which is simply the difference between the two points: $M + \overrightarrow{MN} = N$. In this context, the *barycenter* (or mean) of a system of points $(X_i)_{i=1..n}$ associated to the non-negative normalized weights (w_i) ($\sum_i w_i = 1$) is the unique point M that verifies the following equation, called *barycentric*:

$$\Sigma_i w_i \overrightarrow{MX_i} = \overrightarrow{0}. \tag{7}$$

This equation means geometrically that M is the mean of the (X_i) with respect to the *weights* (w_i). Since F is a flat space, we can get a closed form for M:

$$M = X_1 + \Sigma_i w_i \overrightarrow{X_1X_i}. \tag{8}$$

This kind of mean or averaging procedure is the direct generalization in the affine case of the arithmetic mean of real numbers. It gives a geometrical interpretation to the weighted mean: at the mean, the sum of the weighted displacements to each of the sample points is null, i.e. *the mean is at the geometrical center of the data* (more precisely at the center with respect to the weights).

Fréchet Means and Barycenters. The Fréchet mean of N points (x_i) with respect to the non-negative normalized weights w_i which induced by a Riemannian metric on a manifold is defined *implicitly* by the following barycentric equation ([7]):

$$\sum_{i=1}^{N} w_i \log_{\mathbb{E}(x_j)}(x_i) = 0, \tag{9}$$

where $\log_{\mathbb{E}(x_j)}$ is the logarithmic map at the point $\mathbb{E}(x_j)$, which is defined only locally around this point. In the particular case of bi-invariant metrics, this equation writes:

$$\sum_{i=1}^{N} w_i \log(\mathbb{E}(x_j)^{-1}.x_i) = 0, \tag{10}$$

where this time log is the inverse of the *group* exponential, defined locally around the neutral element e: the (metric) logarithmic map is expressed simply in term of group logarithm. Eq. (10) has particularly nice invariance properties: left-, right- and inverse-invariance, since it derives from a bi-invariant metric. One should note that Eq. (9) (and (10) in the bi-invariant case) provides a *geometrical* definition of the mean, exactly as in the case of affine spaces. The Fréchet mean is defined as a *barycenter*, i.e. the element positioned at the center of the data in *vectorial* sense. This situation is illustrated in Fig. 2.

The key idea developed in this Section in the following: although bi-invariant metrics may fail to exist, the group logarithm *always* exists in a Lie group and one can try to define a bi-invariant mean *directly* via Eq. (10). As will be shown in the next subsections, this equation has all the desirable invariance properties, even when bi-invariant metrics do not exist.

4.2 Stability of the Classical Iterative Scheme

In the cases where a bi-invariant metric exists, a very efficient iterative strategy can be used to solve iteratively the barycentric equation given by Eq. (10). It is given by:

1) `Initialize for example` $m_0 := x_1$.
2) `Update the estimate of the mean by:` $m_{t+1} := m_t.\exp\left(\sum_{i=1}^{N} w_i \log({m_t}^{-1}.x_i)\right)$.
3) `Test the difference between` m_{t+1} `and` m_t `to decide to stop or to loop with step 2.`

This procedure was originally proposed in [9] to compute empirically bi-invariant means in the particular case of *linear* transformations, without any proof of convergence or of the existence and uniqueness of this mean, which was not precisely defined. Here, one of our contributions is to provide a *general and precise definition* of bi-invariant means, which is *valid for any finite-dimensional real Lie group*, via Eq. (10). This will allow us to show

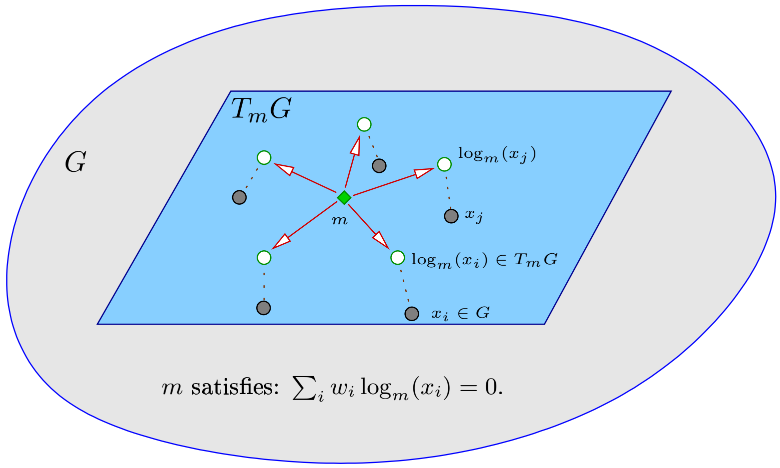

Figure 2: **Geometric Property of the Fréchet mean**. The Fréchet mean of a set of points (x_i) with non-negative (normalized) weights (w_i) satisfies a *barycentric equation*. This has a geometric interpretation: in the tangent space at the mean m, 0 (i.e. m) is precisely the barycenter of the vectors $\log_m(x_i)$ associated to the weights w_i. In this geometrical sense, m is at the center of the points x_i.

the existence and uniqueness of the bi-invariant mean provided the dispersion of the data is small enough.

Interestingly, the mapping $\Phi : m \mapsto m.\exp\left(\sum_{i=1}^N w_i \log(m^{-1}.x_i)\right)$ plays a central role in the approach presented in this article. Let us now detail some of its properties.

Proposition 4. *Let (w_i) be N fixed non-negative weights. Then mapping $\Psi : \mathfrak{g}^{N+1} \to \mathfrak{g}$ defined by $\Psi(v_1, ..., v_N, z) = \log\left(\exp(z).\exp\left(\sum_{i=1}^N w_i \log(\exp(-z).\exp(v_i))\right)\right)$ is analytic near* 0.

Proof. The multivariate nature of Ψ complicates the proof a little bit, but this comes from the simple fact that Ψ is a composition of other analytic mappings: namely the mapping H defined in Subsection 3.4, the mapping $v \mapsto -v$ and the weighted sum $(v_1, ..., v_N) \mapsto \sum_i w_i.v_i$. This suffices to ensure that near 0, Ψ is the sum an absolutely converging infinite multivariate series whose variables are the $v_1, ..., v_N$ and z. For more details on multivariate analytic functions in Lie algebras, see [24], Chapter VI. $\square$

Actually, Ψ is also an analytic function of the non-negative normalized *weights* (w_i). Since these weights live in a compact set, we can guarantee the existence and uniqueness of the bi-invariant mean *independently of the weights considered* (provided the dispersion of the data is small enough). But for more simplicity and clarity, we will skip these details in this article and consider fixed weights (w_i) in the proofs.

Corollary 2. *Let us suppose that the x_i and m are sufficiently close to e. Then we have the following development:*

$$\log(\Phi(m)) = \bar{l} + O((\sum_{i=1}^{N} \| \log(x_i)\| + \| \log(m)\|)^2), \tag{11}$$

where $\bar{l} = \sum_i w_i \log(x_i)$.

Proof. Successive applications of the BCH formula (see Subsection 3.4) yield the first term $(\bar{l})$ of the infinite series of Ψ, which is intuitively the usual arithmetic mean obtained when all the data and m commute. The bound obtained on the deviation with respect to $\bar{l}$ is a direct consequence of the fact that Ψ is analytic: the order of any remaining term of the infinite series is equal or larger to two and as a consequence the other terms can be bounded by a $O((\sum_{i=1}^{N} \| \log(x_i)\| + \| \log(m)\|)^2)$. $\square$

Corollary 2 has the following consequence:

Corollary 3. *For all α in $]0,1[$, there exists a $R > 0$ such that whenever $\| \log(x_i)\| \leq \alpha.R$ and $\| \log(m)\| \leq R$ then we also have $\| \log(\Phi(m))\| \leq R$.*

Proof. Just notice in Eq. (11) that the norm of first order term is less or equal than $\alpha.R$ and that the second-order term, which is a $O((\sum_{i=1}^{N} \| \log(x_i)\| + \| \log(m)\|)^2)$, there exists a constant C such that the second-order term is bounded in the following way:

$$\|O((\sum_{i=1}^{N} \| \log(x_i)\| + \| \log(m)\|)^2)\| \leq C.(N.\alpha^N + 1).R^2.$$

Since R^2 is a $o(R)$, $C.(N.\alpha^N + 1).R^2 \leq (1-\alpha).R$ provided that R is sufficiently small. From this we obtain $\| \log(\Phi(m))\| \leq \alpha.R + (1-\alpha).R = R$, which concludes the proof. $\square$

Corollary 3 shows that provided the x_i and m are close enough to e, we can iterate indefinitely Φ over the successive estimates of the 'mean' of the x_i. This shows that the iterative scheme presented before is *stable* and remains indefinitely well-defined when the data is close enough to e (that is, without taking numerical errors into account). For the moment, we have only considered the case where all elements are close to e. We will see in the next subsection how this extends to the general case, where all the data are only assumed to be close to one another, possibly very far from e.

4.3 Convergence: Special Case

The Bi-invariant Mean as a Fixed Point. Let α be in $]0,1[$, then, accordingly with Corollary 3, let us take a $R > 0$ such that provided for all i $\| \log(x_i) \| \leq \alpha.R$ and $\| \log(m) \| \leq R$, then $\| \log(\Phi(m)) \|$ is also non-superior to R.

Then, let us define $\Omega = \{ m \in G : \ \| \log(m) \| \leq R \}$. From Corollary 3, we know that Φ defines a mapping from Ω to Ω. Now, let us note that $\tilde{m} \in \Omega$ is a solution of Eq. (10) if and only if $\tilde{m}$ is a *fixed point* of Φ, i.e. $\Phi(\tilde{m}) = \tilde{m}$. If we want to show the existence of a solution of Eq. (10), the mathematical tool we need is therefore some sort of *fixed point Theorem.*

In fact, the mathematical literature abounds with fixed point theorems. First, let us consider Brouwer's fixed point theorem:

Theorem 7. *Brouwer's Fixed Point Theorem [27]. Let $\Psi : B^n \rightarrow B^n$ be a continuous mapping, where B^n is the n-dimensional Euclidean closed ball, i.e. $B^n = \{ x \in \mathbb{R}^n : \sum_i (x_i)^2 \leq 1 \}$. Then Ψ has at least one fixed point.*

Corollary 4. *With the assumptions made at the beginning of this subsection, then Eq. (10) has at least one solution in Ω.*

Proof. In our case, this result applies, since we can define $\Psi : \log(\Omega) \rightarrow \log(\Omega)$ by $\Psi(v) = \log(\Phi(\exp(v)))$. Since $\log(\Omega)$ is precisely a closed ball, and thus homeomorphic to the Euclidean closed ball, then Brouwer's theorem applies and guarantees the existence of at least one fixed point of Ψ, which is also a fixed point of Φ and therefore a solution of Eq. (10). □

The existence of a solution to Eq. (10) is thus guaranteed. However, in order to prove the convergence of the iterative strategy to a fixed point of Φ, the mathematical tool we need is another type of fixed point theorem. We will now recall Picard's fixed point Theorem, which is the following:

Theorem 8. *Picard's Fixed Point Theorem. Let (E, d) be a complete metric space and $f : E \rightarrow E$ be a $K-$contraction, i.e. for all x, y of E, $d(f(x), f(y)) \leq K.d(x, y)$, with $0 < K < 1$. Then f has a unique fixed point p in E and for all sequence $(x_n)_n > 0$ verifying $x_{n+1} = f(x_n)$, then $x_n \rightarrow p$ when $n \rightarrow +\infty$, with at least a $K-$linear speed of convergence.*

Proof. This is classical undergraduate topology. The idea is to prove to take any sequence verifying $x_{n+1} = f(x_n)$, to show that it is a Cauchy sequence and thus has a limit in E, which is the unique fixed point of E. Moreover, we have $d(x_{n+1}, p) \leq K.d(x_n, p)$, which shows that the speed of convergence is at least $K-$linear. □

Here, (Ω, d) is the complete metric space in which the successive evaluations of the 'mean' live. The distance d is simply given by $d(m, n) = \| \log(m) - \log(n) \|$. To obtain the existence, uniqueness of a solution of Eq. (10) and linear convergence of our iterative scheme to this point, it only remains to show that Φ is a contraction. This leads to the following Proposition:

Proposition 5. *When the R in Corollary 3 is chosen small enough, Φ is a contraction.*

Proof. Let us consider $E = \log(\Omega)$ with $\Theta : E \to E$ defined as in the proof of Corollary 4 by $\Theta(v) = \log(\Phi(\exp(v)))$. The key idea is to see that Θ is smooth with respect to $\log(m)$ *and* the $(\log(x_i))$, with the property that the norm of the differential of Θ is uniformly bounded in the following way:

$$\|D_{\log(m)}\Theta\| \leq O(\|\log(m)\| + \sum_i \|\log(x_i)\|). \tag{12}$$

In fact, Eq. (12) is a simple consequence of the fact that Ψ is analytic: $D_{\log(m)}\Theta$ is simply one of its partial derivative, which is therefore also analytic. Its value at 0 is precisely 0, and therefore all the terms of its infinite series are of order one or larger, which yields the bound in $O(\|\log(m)\| + \sum_i \|\log(x_i)\|)$.

With the bound given by Eq. (12), we can ensure that when R is small enough, there exists β in $]0,1[$ such that $\|D_{\log(m)}\Theta\| \leq \beta$ for all m. Then we have the classical bound:

$$\|\Theta(v) - \Theta(w)\| \leq (\sup_{z \in E}\|D_z\Theta\|)\|v - w\| \leq \beta\|v - w\|.$$

Since $\beta < 1$, Θ is by definition a contraction, and so is Φ.

□

Corollary 5. *As a consequence, when the data (x_i) is chosen close enough to e, there exists a open neighborhood of e in which there exists a unique solution to Eq. (10). Moreover, the iterative strategy given above always converges towards this solution, provided that the initialization to this algorithm is taken close enough to the data. Last but not least, the speed of convergence is at least linear.*

Proof. Just apply Picard's Theorem to Φ and recall that being a fixed point of Φ is equivalent to being a solution of Eq. (10). □

4.4 Convergence: General Case

In the previous subsection, we have shown the existence and uniqueness of a solution of the exponential barycentric equation living in a open neighborhood of e, as long as the data were all close enough to e. In fact, this can be greatly generalized, as shown by the following result:

Theorem 9. *Let g be in G. Then there exists a groupwise geodesically convex open neighborhood $\mathcal{V}$ of g, such that whenever the data (x_i) are in $\mathcal{V}$, then there exists a unique solution of Eq. (10) in an open neighborhood of g. Moreover, the classical iterative strategy always converges towards this solution, provided the initialization is taken close enough to g; also, the speed of convergence is at least linear.*

Proof. Just multiply the data by g^{-1} on the left to shift all the points in an adequate open neighborhood of the neutral element. Then one can run the iterative scheme to obtain the unique solution of the barycentric equation which lives close to e. Then just multiply on the left by g this solution to find the unique solution of the barycentric equation with the real data. Then note that the normal (non-shifted) iterative scheme is just the shifted version of the scheme associated to the shifted data. $\square$

This leads to the following definition:

Definition 4. *Let the $(x_i)_{i=1}^N$ be some data belonging to a small enough groupwise geodesically convex set of G. Then, for any system of (normalized) non-negative weights $(w_i)_{i=1}^N$, we call* **bi-invariant mean** *of (x_i) with respect to the weights (w_i) the unique solution (in a neighborhood of the data) of the group barycentric equation (10).*

Proposition 6. *The bi-invariant mean is left-, right- and inverse-invariant.*

Proof. The data is by hypothesis close enough to one another so that we can apply Lemma 1, so that we have:

$$\begin{aligned} Ad(m).\left(\textstyle\sum_{i=1}^N w_i \log(m^{-1}.x_i)\right) = \quad & \textstyle\sum_i w_i \log\left(m.(m^{-1}.x_i).m^{-1}\right) \\ & \textstyle\sum_{i=1}^N w_i \log(x_i.m^{-1}). \end{aligned}$$

Since $Ad(m)$ is invertible, the usual barycentric equation, which is left-invariant, is equivalent to a right-invariant barycentric equation, which shows that the barycenter is both left- and right-invariant. Now, to prove the invariance with respect by inversion, note that:

$$\begin{aligned} (-1) \times \left(\textstyle\sum_{i=1}^N w_i \log(m^{-1}.x_i))\right) = \quad & \textstyle\sum_i w_i \log\left((m^{-1}.x_i)^{-1}\right) \\ & \textstyle\sum_{i=1}^N w_i \log(x_i^{-1}.(m^{-1})^{-1}), \end{aligned}$$

which shows that whenever m is the bi-invariant mean of the x_i, m^{-1} is that of the ${x_i}^{-1}$, which is exactly inverse-invariance. $\square$

Some Comments on Bi-invariant Means. Thus, we have rigorously generalized to any real Lie group the notion of bi-invariant mean normally associated to bi-invariant Riemannian metrics, even in the case where such metrics fail to exist. This novel mean enjoys all the desirable invariance properties, and can be iteratively computed in a very efficient way.

One should note that as usual with mean in manifolds, the bi-invariant mean only exists provided the data are close enough to one another: the dispersion should not be too large. In the next section, we will see more precisely in various situations what practical limitation is imposed on the dispersion of the data. One does not seem to lose much in this regard with respect to existing Riemannian bi-invariant means: we will show for example that the bi-invariant mean of some rigid transformations exists if and only if the bi-invariant mean of their rotation parts exists.

5 Bi-invariant Means in Simple Cases

Let us now detail several insightful cases where the algebraic mean can be *explicitly* or directly computed, without using the classical iterative scheme.

5.1 Bi-invariant Mean of Two Points

There is a closed form for the bi-invariant mean of two points:

Proposition 7. *Let x be in G and y be in a GGC open neighborhood of x. Then their bi-invariant mean m with respect to the couple of weights $(1-\alpha, \alpha)$ is given by:*

$$m = x.\exp\left(\alpha \log(x^{-1}.y)\right) = x.(x^{-1}.y)^{\alpha}. \tag{13}$$

Proof. We can simply check that m is a solution to the adequate barycentric equation. We have:

$$\log(m^{-1}.x) = \log(\exp(-\alpha\log(x^{-1}.y)).x^{-1}.x) = -\alpha\log(x^{-1}.y).$$

Also, we have that:

$$\log(m^{-1}.y) = \log(\exp(-\alpha\log(x^{-1}.y)).x^{-1}.y) = \log(n^{-\alpha}.n),$$

with $n = x^{-1}.y$. Therefore:

$$\left\{\begin{array}{l} \alpha.\log(m^{-1}.y) = \log(n^{\alpha\times(1-\alpha)}) \\ (1-\alpha).\log(m^{-1}.x) = \log(n^{-\alpha\times(1-\alpha)}) = -\alpha.\log(m^{-1}.y). \end{array}\right.$$

Thus, m is the bi-invariant mean of x and y. □

Notice that the explicit formula given by Eq. (13) is quite exceptional. In general, there will be *no* closed form for the bi-invariant mean, as soon as $N > 2$. However, there are some specific groups where a closed form exists for the bi-invariant mean in *all cases*, and we will now detail some examples of this rare phenomenon.

5.2 Scalings and Translations in 1D

Here, we will devote some time to a very instructive group: *the group of scalings and translations in 1D*. The study of this (quite) simple group is relevant in the context of this work, because it is one of the most simple cases of non-compact *and* non-commutative Lie groups which does not possess any bi-invariant Riemannian metric. This group has many of the properties of rigid or affine transformations, but with only two degrees of freedom, which simplifies greatly the computations, and allows a direct (2D) geometric visualization in the plane. For these reasons, this is a highly pedagogical case. In the rest of this Subsection, we will let this group be written $ST(1)$.

Elementary Algebraic Properties of $ST(1)$.

- An element g of $ST(1)$ can be uniquely represented by a couple (λ, t) in $\mathbb{R}_+^{\star} \times \mathbb{R}$. λ corresponds to the scaling factor and t to the translation part.
- The action of $ST(1)$ on scalars is given by: $(\lambda, t).x = \lambda.x + t$ for every scalar x.

- The multiplication in $ST(1)$ is: $(\lambda', t').(\lambda, t) = (\lambda'.\lambda, \lambda'.t + t')$. $ST(1)$ is thus a *semi-direct product* between the multiplicative group $(\mathbb{R}_+^\star, \times)$ and the additive group $(\mathbb{R}, +)$. Both groups are commutative, but this semi-direct product is *not*.
- Inversion: $(\lambda, t)^{-1} = (\frac{1}{\lambda}, -\frac{t}{\lambda})$.
- $ST(1)$ can be faithfully represented by the subgroup of triangular matrices of the form:

$$\begin{pmatrix} \lambda & t \\ 0 & 1 \end{pmatrix}.$$

- The elements of the Lie algebra of $ST(1)$ are of the form $(d\lambda, dt)$, where $d\lambda$ and dt are any scalars.
- The group exponential $\exp(d\lambda, dt)$ has the following form:

$$\exp(d\lambda, dt) = \begin{cases} (e^{d\lambda}, \frac{dt}{d\lambda}(e^{d\lambda} - 1), \text{ when } d\lambda \neq 0, \\ (1, dt), \text{ when } d\lambda = 0, \end{cases}$$

  where e^λ is the scalar exponential of λ. Thus, we see that the group exponential is simply given by the scalar exponential on the scaling part, whereas the translation part mixes the multiplicative and additive influences of both components. Moreover, we see geometrically than in the upper half plane $\mathbb{R}_+ \times \mathbb{R}$, the curve given by $\exp(s.(d\lambda, dt))$ with s varying in $\mathbb{R}$ is on a ***straight line***, whose equation is $t = \frac{dt}{d\lambda}\lambda - 1$.

- $ST(1)$ is entirely groupwise geodesically convex: any two points can be joined by a unique group geodesic. In particular, the group logarithm is always well-defined and given by:

$$\log(\lambda, t) = \begin{cases} (\ln(\lambda), t.\frac{\ln(\lambda}{1-\lambda}), \text{ when } \lambda \neq 1, \\ (0, t), \text{ when } \lambda = 1, \end{cases}$$

  where $\ln(\lambda)$ is the natural (scalar) logarithm of λ. Same remark has for the exponential: we get the classical logarithm on the scaling part and a mixture of the multiplicative and additive logarithms on the translation part. We recall that in the case of an additive group such $(\mathbb{R}, +)$, both additive exponential and logarithm are simply the identity. This is what we get both for the exponential and the logarithm when there is no scaling.

  The unique group geodesic joining (λ, t) and (λ', t') of the form $(\lambda, t).\exp(s.(d\lambda, dt))$ with s in $[0, 1]$ has its parameters $(d\lambda, dt)$ given by:

$$(d\lambda, dt) = \left(\ln(\frac{\lambda'}{\lambda}), \left(\frac{t' - t}{\lambda}\right).\left(\frac{\ln(\frac{\lambda'}{\lambda})}{\frac{\lambda'}{\lambda} - 1}\right)\right). \tag{14}$$

Absence of Bi-invariant Metrics. $ST(1)$ is one of the most simple non-compact and non-commutative Lie groups. In terms of bi-invariant metrics, it exhibits the typical tendency of such Lie groups: it has no such metric. As usual (see Section 2), to see this, we use the fact it is necessary and sufficient that the adjoint representation of $ST(1)$ be not bounded to ensure that no bi-invariant metric exists for this group. To show this, we use again the classical matrix representation of $ST(1)$:

$$\begin{aligned} Ad((\lambda, t)).(d\lambda, dt) &\sim \begin{pmatrix} \lambda & t \\ 0 & 1 \end{pmatrix} . \begin{pmatrix} d\lambda & dt \\ 0 & 0 \end{pmatrix} . \begin{pmatrix} \frac{1}{\lambda} & -\frac{t}{\lambda} \\ 0 & 1 \end{pmatrix} \\ &\sim \begin{pmatrix} d\lambda & -t.d\lambda + \lambda.dt \\ 0 & 0 \end{pmatrix} \\ &= (d\lambda, -t.d\lambda + \lambda.dt). \end{aligned}$$

Both factors 't' and 'λ' in $-t.d\lambda + \lambda.dt$ are not bounded and thus $Ad(ST(1))$ cannot be bounded. Has a consequence, $ST(1)$ has no bi-invariant metric. Both $(\mathbb{R}_+{}^\star, \times)$ and $(\mathbb{R}, +)$ are commutative and thus have bi-invariant metrics, but interestingly, their semi-direct product has no such metric.

A Closed Form for the Bi-invariant Mean. We recall that the bi-invariant mean in a Lie group is defined implicitly by a barycentric equation, given by Eq. (10). Here, since we have explicit formulae for the group exponential and logarithm, one can use these formulae to try to solve directly the barycentric equation. This leads to the following result:

Proposition 8. *Let $((\lambda_i, t_i))$ be N points in $ST(1)$ and (w_i) be N non-negative (normalized) weights. Then the associated bi-invariant mean $(\bar{\lambda}, \bar{t})$ is given explicitly by:*

$$\left\{ \begin{array}{l} \bar{\lambda} = e^{\sum_i w_i \ln(\lambda_i)}, \textit{ (weighted geometric mean of scalings),} \\ \bar{t} = \frac{1}{Z}. \sum_i w_i.\alpha_i.t_i, \textit{ (weighted arithmetic mean of translations influenced by scalings),} \end{array} \right. \tag{15}$$

with:

$$\left\{ \begin{array}{l} \alpha_i = \frac{\ln(\frac{\lambda_i}{\bar{\lambda}})}{\frac{\lambda_i}{\bar{\lambda}} - 1}; \textit{ note that } \alpha_i = 1 \textit{ when } \lambda_i = \bar{\lambda}. \\ Z = \sum_i w_i.\alpha_i \end{array} \right. .$$

Proof. Just replace in the barycentric equation the exponentials and logarithms by the formulae given above. Since the scaling component is independent from the translation one, we simply obtain the geometric mean, which is the bi-invariant mean for positive numbers. The translation part can be handled simply using directly Eq. (14), which yields this simplified expression for the barycentric equation:

$$\sum_i w_i \left(\frac{t_i - \bar{t}}{\bar{\lambda}} \right) . \left(\frac{\ln(\frac{\lambda_i}{\bar{\lambda}})}{\frac{\lambda_i}{\bar{\lambda}} - 1} \right) = 0.$$

Hence the result. □

Comparison Between Group and Metric Geodesics. In Figure 3, one can visually compare the group geodesics to some of their left-invariant and right-invariant (metric) counterparts.

Interestingly, one of the left-invariant metrics on $ST(1)$ induces an isometry between this group and Poincaré half-plane model for hyperbolic geometry (see [13], page 82-83 for more details on this space). The scalar product of this scalar metric is the most simple at the $(1,0)$: it is the usual Euclidean scalar product. Geodesics take a very particular form in this case: they are the set of all the *half-circles perpendicular to the axis of translations* and of all (truncated below the axis of translations) lines perpendicular to the axis of translations (these lines can be seen as half-circles of infinite diameter anyway).

Thanks to Proposition 1, we know that the right-invariant Riemannian metric whose scalar product at $(1,0)$ is the same as the previous metric can be obtained simply by 'inverting' this left-invariant metric. As a consequence, its geodesics can be computed simply by inverting the initial conditions, computing the associated left-invariant geodesic and finally inverting it. The right-invariant geodesics visualized in Fig. 3 are by consequence some sort of 'inverted half-circles'. In fact, simple algebraic computations show that these geodesics are all *half-hyperbolas*.

One should note that the simple form taken by left-invariant geodesics is indeed exceptional. In general, there are no closed form for neither left- nor right-invariant geodesics, and group geodesics are much simpler to compute, since in most practical cases they only involve the computation of a matrix exponential and a matrix logarithm, for which very efficient methods exist [28, 29]. Another nice Lie group where left-invariant metrics (and by consequence also right-invariant metrics) take a simple (closed) form is the group of rigid transformations. See [5] for examples of left-invariant statistics on rigid transformations in the context of a statistical study of human scoliotic spines.

Extension to $ST(n)$. One can directly generalize the results obtained for $ST(1)$ to the more general group $ST(n)$ of scalings and translations in nD. Instead of being a scalar, the translation is in this general case a $n-$dimensional vector. This does not change anything: all the algebraic properties of $ST(1)$ are also valid for $ST(n)$. In particular, one can use Eq. (15) to compute bi-invariant means in $ST(n)$.

5.3 The Heisenberg Group

With the group $ST(1)$, we had seen a simple case of mixing between a 1D multiplicative group and a 1D additive group. In this subsection, we study instead a 3D group where this time 2 additive groups (one 2D and the other 1D) are mixed.

The Heisenberg group. It is the group of 3D upper triangular matrices M of the form:

$$M = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

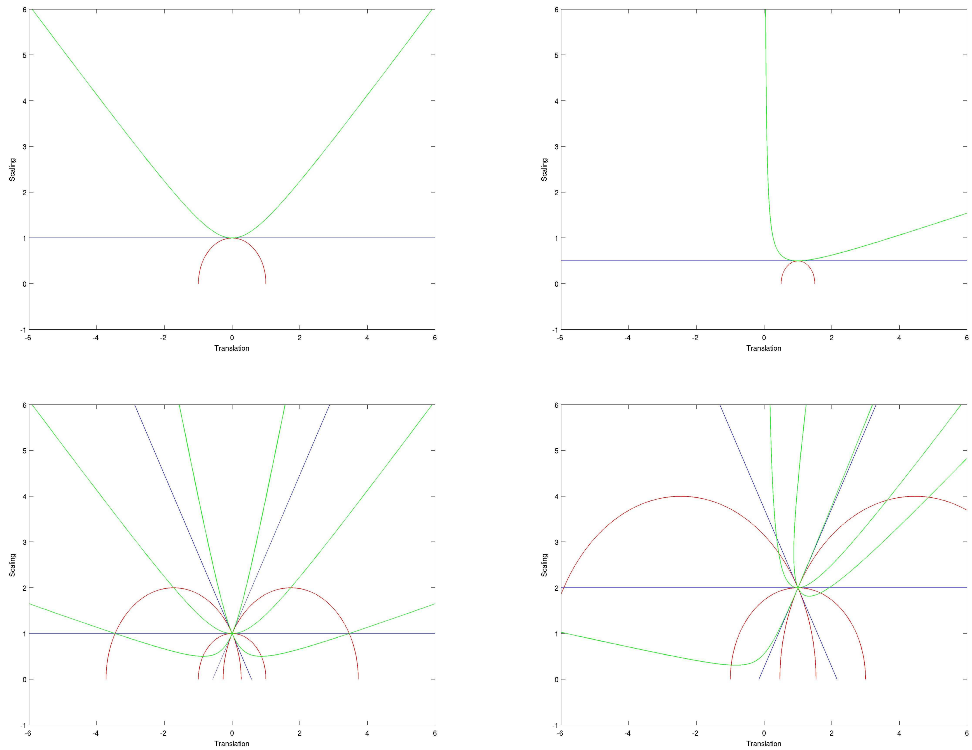

Figure 3: **Examples of geodesics in the group of scalings and translations in 1D**. **Top row**: two examples of left- and right- and group geodesics. **Bottom row**: two examples of geodesics with each time three possible orientations. **Blue**: group geodesics, **red**: left-invariant geodesics and **green**: right-invariant geodesics. Note the particular form taken by group geodesics, which are part of straight lines and of the left-invariant geodesics, which are half-circles perpendicular to the horizontal axis. Right-invariant geodesics are also given in a closed form and are in fact half-hyperbolas.

To simplify notations, we will also write (x, y, z) to represent an element of this group.

Elementary Algebraic Properties. They are the following:

- Multiplication: $(x_1, y_1, z_1).(x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2 + x_1.y_2)$. The first two parameters thus live in a 2D additive group which is independent of the third parameter, whereas the third additive parameter is influenced by the first two. The Heisenberg group is thus a *semi-direct product* between $(\mathbb{R}^2, +)$ and $(\mathbb{R}, +)$, which is *not* commutative.

- Inversion: $(x, y, z)^{-1} = (-x, -y, -z + x.y)$. Neutral element: $(0, 0, 0)$.
- As in the $ST(1)$ case, the Heisenberg group is entirely groupwise geodesically complete and we have:
$$\begin{cases} \exp((dx, dy, dz)) = (dx, dy, dz + \frac{1}{2}.dx.dy) \\ \log((x, y, z)) = (x, y, z - \frac{1}{2}x.y). \end{cases}$$
The unique group geodesic joining (x_m, y_m, z_m) and (x, y, z) of the form $(x_m, y_m, z_m). \exp(s.(dx, dy, z))$ with s in $[0, 1]$. Its parameters (dx, dy, dz) are given by:
$$(dx, dy, dz) = \left(x - x_m, y - y_m, z - z_m + \frac{1}{2}.(x_m.y_m - x.y + x_m.y - x.y_m)\right). \quad (16)$$

Bi-invariant Metrics and Bi-invariant Means. As in the $ST(1)$ case, no bi-invariant metrics exists and one has the closed form for the bi-invariant mean. Interestingly, the bi-invariant mean yields a simple arithmetic averaging of the first two parameters. The third parameter is also averaged arithmetically, except that this arithmetic mean is 'corrected' by a quadratic function of the first two parameters of the data.

Proposition 9. *The action of the adjoint representation Ad of the Heisenberg group at a point (x, y, z) on an infinitesimal displacement (dx, dy, dz) is given by:*
$$Ad(x, y, z).(dx, dy, dz) = (dx, dy, -y\, dx + x\, dy + dz).$$
As a consequence, no bi-invariance metric exists for the Heisenberg group.

Proof. Proceed exactly as in Proposition 2. □

Proposition 10. *Let $((x_i, y_i, z_i))$ be N points in the Heisenberg group and (w_i) be N non-negative (normalized) weights. Then the associated bi-invariant mean $(\bar{x}, \bar{y}, \bar{z})$ is given explicitly by:*
$$(\bar{x}, \bar{y}, \bar{z}) = \left(\sum_i w_i x_i, \sum_i w_i y_i, \sum_i w_i z_i + \frac{1}{2}\left(\bar{x}.\bar{y} - \sum_i w_i x_i.y_i\right)\right).$$

Proof. Just replace in the barycentric equation the exponentials and logarithms by the formulae given above. Since the first two components are additive and independent from the third one, their bi-invariant mean is simply their arithmetic mean. The third coefficient case can be handled simply using directly Eq. (16), which yields this simplified expression for the barycentric equation:
$$\sum_i w_i \left(z_m - z_i + \frac{1}{2}.(x_m.y_m - x_i.y_i + x_m.y_i - x_i.y_m)\right).$$
Hence the result. □

5.4 On a Subgroup of Triangular Matrices

We can generalize the results obtained on the Heisenberg group to the following subgroup of triangular matrices:

Definition 5. *Let $UT(n)$ be the group of $n \times n$ upper triangular matrices M of the form:*

$$M = \lambda.Id + N,$$

where λ is any positive scalar, Id the identity matrix and N an upper triangular nilpotent matrix ($N^n = 0$) with only zeros in its diagonal.

The Heisenberg group is the subgroup of matrices of $UT(3)$ whose λ is always equal to 1. The situation in this case is particularly nice, since thanks to the fact that N is nilpotent, one can perform *exactly* all the usual algebraic operations in $UT(n)$:

- Group exponential:

$$\begin{aligned} \exp(dM) &= \exp(d\lambda.Id + dN) = \exp(d\lambda.Id).\exp(dN) \\ &= e^{d\lambda}.\textstyle\sum_{k=0}^{n-1} \frac{dN^k}{k!}. \end{aligned}$$

- Group logarithm:

$$\begin{aligned} \log(M) &= \log(\lambda.Id + N) = \log\left((\lambda.Id).(Id + \tfrac{1}{\lambda}.N)\right) \\ &= \ln(\lambda).Id + \textstyle\sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k} (\frac{N}{\lambda})^k. \end{aligned}$$

- Inversion:

$$\begin{aligned} (M)^{-1} &= (\lambda.Id + N)^{-1} = \lambda^{-1}.(Id + \tfrac{N}{\lambda})^{-1} \\ &= \lambda^{-1}.\textstyle\sum_{k=0}^{n-1} (-1)^k.(\frac{N}{\lambda})^k. \end{aligned}$$

- Multiplication:

$$M'.M = (\lambda'.Id + N').(\lambda.Id + N) = (\lambda'.\lambda).Id + (\lambda'.N + \lambda.N' + N'.N).$$

Using these closed forms, one can derive the following equation:

$$\log(M'.M) = \ln(\lambda'.\lambda).Id + \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k}.\left(\frac{1}{\lambda}.N + \frac{1}{\lambda'}.N' + \frac{1}{\lambda'.\lambda}.N.N'\right)^k,$$

which in turns allows to compute the equation satisfied by the bi-invariant mean $\bar{M} = \bar{\lambda}.Id + \bar{N}$ in $UT(n)$:

$$\begin{gathered} -\textstyle\sum_i w_i \log(\bar{M}^{-1}.M_i) = \sum_i w_i \log(\bar{M}.{M_i}^{-1}) = 0 \\ \Longleftrightarrow \\ \textstyle\sum_i w_i \left(\ln(\bar{\lambda}.{\lambda_i}^{-1}).Id + \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{k}.\left(\frac{1}{{\lambda_i}^{-1}}.{N_i}^{-1} + \frac{1}{\bar{\lambda}}.\bar{N} + \frac{1}{\bar{\lambda}.{\lambda_i}^{-1}}.{N_i}^{-1}.\bar{N}\right)^k\right) = 0, \end{gathered} \tag{17}$$

where we ${N_i}^{-1}$ is the nilpotent part of ${M_i}^{-1}$. From Eq. (17), we see that $\bar{\lambda}$ is simply the geometric mean of the λ_i, and that the coefficient of $\bar{N}$ can be recursively computed, starting from coefficients above the diagonal. The key idea is that the k^{th} power of a nilpotent matrix N will have non-zero coefficients only in its k^{th} upper diagonal.

As a consequence, to compute the coefficients of $\bar{M}$ above the diagonal, one only needs to take into account the the following terms: $\frac{1}{{\lambda_i}^{-1}}.{N_i}^{-1} + \frac{1}{\bar{\lambda}}.\bar{N}$. These coefficients will simply be a weighted arithmetic mean of the coefficients in the data, the weights being equal to $(w_i.\frac{\lambda_i}{\bar{\lambda}})/S$ with $S = \sum_j w_j.\frac{\lambda_j}{\bar{\lambda}}$. Using this result, then one can compute the coefficients above, which are an weighted arithmetic mean of the corresponding coefficients in the data, with a quadratic correction involving the previous coefficients. The same phenomenon appears for the next set of coefficients above, with an even more complex correction involving all the previously computed coefficients. One can continue this way until all the coefficients of the mean have been effectively computed.

6 Linear Transformations

6.1 General Rigid Transformations

We recall that the Lie group of rigid transformations in the n-dimensional Euclidean space, written here $SE(n)$, is the semi-direct product between $(SO(n), \times)$ (rotations) and $(\mathbb{R}^n, +)$ (translations) defined has follows:

- An element of $SE(n)$ is uniquely represented by a couple $(R, t) \in SO(n) \times \mathbb{R}^n$ and its action on a point x of $\mathbb{R}^n$ is given by $(R, t).x = R.x + t$.
- Multiplication: $(R', t').(R, t) = (R'.R, R'.t + t')$.
- Neutral element: $(Id, 0)$, inverse: $(R^T, -R^T.t)$.
- Representation of (R, t) by a $(n+1) \times (n+1)$ matrix using homogenous coordinates:
$$\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}.$$
- Lie algebra: thanks to the matrix representation of $SE(n)$, it is simple to see that the Lie algebra of $SE(n)$ can be faithfully represented by the following vector space of matrices:
$$\begin{pmatrix} dR & dt \\ 0 & 0 \end{pmatrix},$$
where dR is any skew $n \times n$ matrix and dt any vector of $\mathbb{R}^n$. In this representation, the Lie bracket $[.,.]$ is simply given by the matrix Lie bracket: $[A, B] = A.B - B.A$.

- Group exponential: it can be computed using directly the matrix representation, or by identifying the one-parameter subgroups of $SE(n)$. It is given by:

$$\exp(dR, dt) = \left(e^{dR}, e^{dR}.(\int_0^1 e^{-u.dR} du).dt \right),$$

  where e^{dR} is the matrix exponential of dR.

Existence of the Logarithm. From Section 3, we know that it is only defined locally in a neighborhood of the neutral element $(Id, 0)$. However, since we have a faithful representation of $SE(n)$ in terms of matrices, we can use the matrix criterion for the existence of the principal logarithm: from Subsection 3.1, we know that an invertible matrix with no (complex) eigenvalue on the closed half-line of negative real numbers has unique matrix logarithm with eigenvalues having imaginary parts in $]-\pi, \pi[$. In the case of rotations, this means that the various angles of rotation (there can be *several* angles of rotation in the general n-dimensional case, whereas only one exists in 2D or 3D) of a rotation R should not go outside $]-\pi, \pi[$ if we want the logarithm of R to be well-defined. Otherwise, one cannot define a unique logarithm. This is typically the case for $-Id$ in 2D (i.e. a rotation of 180 degrees), whose two 'smallest' real logarithms are the following:

$$\begin{pmatrix} 0 & -\pi \\ \pi & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \pi \\ -\pi & 0 \end{pmatrix}.$$

Going back to $SE(n)$, we have the following result:

Proposition 11. *The logarithm of a rigid transformation (R, t) is well-defined if and only if the logarithm of its rotation part R is well-defined.*

Proof. The logarithm of (R, t) is well-defined if and only if the matrix representing (R, t) has a principal logarithm, which is equivalent to the fact that it has no eigenvalue on the closed negative line. Then, this is equivalent to the fact that R has no eigenvalue on the closed negative line, since the eigenvalues of the upper triangular matrix (in terms of blocks) $\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$ depend only on the blocks in its diagonal, i.e. only on R, and *not* t. As a consequence, the logarithm of a rigid transformation is well-defined if and only if the logarithm of its rotation part is well-defined. $\square$

Criterion for the Existence of the Bi-Invariant Mean. We have seen in Subsection 2.4 that no bi-invariant mean exists in the rigid case. One may now ask the question: is there a simple criterion for the existence of the bi-invariant mean of rigid transformations? In the case of bi-invariant metrics, one has such a criterion: the bi-invariant mean is well-defined as long as all the data is strictly included in a *regular geodesic ball* of radius r such that the geodesic ball of radius $2.r$ is still regular ([7], page 9). In the case of rotations, this is reduces to checking that within the data, there does not exist a couple R and R' such that

$R^{-1}.R'$ has an angle of rotation greater or equal to $\pi - C$ minus a small constant [30]. We have exactly the same result in the case of rigid transformations:

Theorem 10. *Let (R_i, t_i) be N rigid transformations satisfying $\sup_{i,j} d(R_i, R_j) \leq \pi - C$, where $C > 0$ and where $d(R, R')$ is the magnitude of the largest angle of rotation of $R^{-1}.R'$. Then for any set of non-negative weights (w_i), there exists a unique bi-invariant mean for (R_i, t_i).*

Proof. We have just seen that the condition imposed on rotation matrices ensures that their bi-invariant mean is well-defined, for any weights (w_i). Let use write $\bar{R}$ for that barycenter. $\bar{R}$ is included in the same geodesic ball as the data and thus satisfies the same condition on its eigenvalues than the data: $\sup_i d(R_i, \bar{R}) \leq \pi - C$.

We now have to check whether a unique translation $\bar{t}$ exists, which satisfies the following barycentric equation:

$$\sum_i w_i \log((\bar{R}, \bar{t}).(R_i, t_i)^{-1}) = 0. \tag{18}$$

We have: $(\bar{R}, \bar{t}).(R_i, t_i)^{-1} = (\bar{R}.R_i^T, \bar{R}.(-R_i^T.t_i)+\bar{t})$. Let us write $M(dR) = e^{dR}. \int_0^1 e^{-u.dR} du$. We will soon show that $M(dR)$ is always invertible provided that the norm of largest eigenvalue of dR is smaller than π. Then Eq. (18) writes in terms of translations:

$$\begin{gathered} \sum_i w_i M\left(\log(\bar{R}.R_i^T)\right)^{-1}.\left(\bar{R}.(-R_i^T.t_i) + \bar{t}\right) = 0 \\ \Longleftrightarrow \\ \left(\sum_i w_i M\left(\log(\bar{R}.R_i^T)\right)^{-1}\right).\bar{t} = \sum_i w_i M\left(\log(\bar{R}.R_i^T)\right)^{-1}.R_i^T.t_i. \end{gathered} \tag{19}$$

Thus, we see that the existence and unicity of such a $\bar{t}$ is equivalent to the invertibility of the following matrix: $\sum_i w_i \ M\left(\log(\bar{R}.R_i^T)\right)^{-1}$. Under the restriction described above on rotations, this matrix is indeed invertible, and this concludes the proof. To see this, we have the following Lemma: □

Lemma 2. *Let (dR_i) be N a skew symmetric matrices, such that the norm of their largest (complex) eigenvalue is always smaller than $\pi - C$, with $C > 0$.*

Let $M(dR)$ be equal to $e^{dR}. \int_0^1 e^{-u.dR} du$ for any skew symmetric matrix. Then for all dR_i, $M(dR_i)$ is invertible, and for any non-negative weights (w_i), $\sum_i w_i M(dR_i)^{-1}$ is also invertible.

Proof. The key idea is to see the form taken by $M(dR)$ in an appropriate orthonormal basis. From classical linear algebra, we have the following spectral decomposition for any skew symmetric matrix dR: there exists a decomposition of the geometrical space $\mathbb{R}^n$ in a direct sum of *mutually orthogonal* subspaces, which are all *stable* for dR [15].

These subspaces are of two kinds: first k (possibly equal to zero) 2-dimensional vector subspaces E_k, and second a single subspace F of dimension $n - 2.k$ (the orthogonal complement of the other subspaces). F is simply the kernel of dR, and any j in $1...k$, dR restrained to E_j is of the following form (in an appropriate orthonormal basis):

$$\begin{pmatrix} 0 & -\theta_j \\ \theta_j & 0 \end{pmatrix},$$

with $\theta_j \neq 0$, which is the j^{th} angle of rotation of the rotation matrix e^{dR}. Intuitively, *this means that any general rotation can be decomposed in k independent 2D rotations.* This spectral decomposition allows for the explicit computation of $M(dR)$ is the various subspaces mentionned above. In a E_j, $M(dR)$ is equal to:

$$\begin{pmatrix} \frac{\sin(\theta_j)}{\theta_j} & \frac{\cos(\theta_j)-1}{\theta_j} \\ -\frac{\cos(\theta_j)-1}{\theta_j} & \frac{\sin(\theta_j)}{\theta_j} \end{pmatrix}.$$

In F, dR is simply the identity. This shows that whenever for all j, $|\theta_j| < 2.\pi$, $M(dR)$ is always invertible (which is more than we need), since the determinant of the latter matrix is equal to $(\frac{\sin(\theta_j)}{\theta_j})^2 + (\frac{\cos(\theta_j)-1}{\theta_j})^2$, which is positive for $|\theta_j| < 2.\pi$. Furthermore, a direct computation shows that the inverse of $M(dR)$ takes the following form in E_j:

$$\begin{pmatrix} \frac{\theta_j.\sin(\theta_j)}{2.(1-\cos(\theta_j)} & \frac{\theta_j}{2} \\ -\frac{\theta_j}{2} & \frac{\theta_j.\sin(\theta_j)}{2.(1-\cos(\theta_j)} \end{pmatrix}.$$

For $|\theta_j| < \pi - C$, some elementary calculus shows that there exists a constant $K > 0$, such that $\frac{\theta_j.\sin(\theta_j)}{2.(1-\cos(\theta_j)} > K$. As a consequence, the latter matrix is of the following form:

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

with $a > K > 0$. This has the interesting consequence that under the assumption that for all j $|\theta_j| < \pi - C$, $M(dR)^{-1}$ has the following decomposition:

$$M(dR)^{-1} = S + A,$$

where S is a symmetric positive-definite matrices with all its eigenvalues larger than K and where A is a skew symmetric matrix. Then let us take N skew matrices whose eigenvalues are smaller than $\pi - C$. Any convex combination of the $M(dR_i)^{-1}$ writes:

$$\sum_i w_i M(dR_i)^{-1} = (\sum_i w_i S_i) + (\sum_i w_i A_i) = \tilde{S} + \tilde{A},$$

where $\tilde{S}$ is still symmetric definite positive and $\tilde{A}$ is skew symmetric. $\sum_i w_i M(dR_i)^{-1}$ is therefore invertible, since *any* matrix of the form $\tilde{S} + \tilde{A}$ is invertible. To see this, just remark that if there exists one x such that $(\tilde{S} + \tilde{A}).x = 0$, this implies $x^T.\tilde{S}.x + x^T.\tilde{A}.x = 0$. Then notice that $x^T.\tilde{A}.x = (x^T.\tilde{A}.x)^T = -x^T.\tilde{A}.x = 0$. Thus $(\tilde{S} + \tilde{A}).x = 0$ implies $x^T.\tilde{S}.x = 0$, which is equivalent ($\tilde{S}$ is symmetric positive-definite) to $x = 0$. Consequently $\tilde{S} + \tilde{A}$ is invertible and this ends the proof. □

6.2 2D Rigid Transformations

Contrary to the general case, 2D rigid transformations have a particularity: one has an *closed form* for the bi-invariant mean. The reason behing this is that $SO(2)$, the group of 2D rotations, is *commutative*. As a consequence, one can compute explicitely the bi-invariant mean of the rotation parts of the data and deduce from it the translation part using the barycentric equation, like in the proof of Theorem 10. More precisely, we have:

Proposition 12. *Let (R_i, t_i) be N 2D rigid transformations, such that there angles of rotation θ_i satisfy: $\sup_{i,j} |\theta_i - \theta_j| \leq \pi - C$. Then the bi-invariant mean $(\bar{R}, \bar{t})$ associated to the weights (w_i) is given explicitely by:*

$$\begin{cases} \bar{R} = R_1. \exp(+\sum_i w_i \log(R_1^T.R_i)) \\ \bar{t} = \sum_i w_i \, Z^{-1}.M\left(\log(\bar{R}.R_i^T)\right)^{-1}.R_i^T.t_i, \end{cases} \tag{20}$$

with the following formulae for M and Z:

$$M\left(\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}\right)^{-1} \stackrel{def}{=} \begin{pmatrix} \frac{\theta.\sin(\theta)}{2.(1-\cos(\theta)} & \frac{\theta}{2} \\ -\frac{\theta}{2} & \frac{\theta.\sin(\theta)}{2.(1-\cos(\theta)} \end{pmatrix}, Z \stackrel{def}{=} \sum_i w_i M\left(\log(\bar{R}.R_i^T)\right)^{-1}.$$

Example of Bi-Invariant Mean. Let us take a look at the example chosen in [31], page 31. Let $f_1 = (\pi/4, -\sqrt{2}/2, \sqrt{2}/2)$, $f_2 = (0, \sqrt{2}, 0)$ and $f_3 = (-\pi/4, -\sqrt{2}/2, -\sqrt{2}/2)$ be three rigid transformations. The first coefficient corresponds to the angle of rotation (chosen here in $[-\pi, \pi[$) and last two to the translation.

We can compute exactly the bi-invariant mean of these rigid transformations with (20); a left-invariant Fréchet mean can also be computed explicitely in this case thanks to the simple form taken by the corresponding geodesics (see [31] for more details), and finally, thanks to Proposition 1, the analogous right-invariant Fréchet mean can be computed by inverting the data, computing their left-invariant mean and then inverting this Fréchet mean. This yields:

- left-invariant Fréchet mean: $(0, 0, 0)$,
- bi-invariant mean: $(0, \frac{\sqrt{2}-\frac{\pi}{4}}{1+\frac{\pi}{4}.\frac{\sqrt{2}}{2-\sqrt{2}}}, 0) \simeq (0, 0.2171, 0)$,
- right-invariant Fréchet mean: $(0, \frac{\sqrt{2}}{3}, 0) \simeq (0, 0.4714, 0)$.

Interestingly, we thus see that the mean rotation is exactly the same in all three cases. But the mean translations are different, and the bi-invariant mean is situated nicely *between* the left- and right-invariant Fréchet means. This is quite intuitive, since the bi-invariant mean can be looked upon as an in-between alternative with regard to left- and right-invariant Fréchet means.

6.3 General Linear Transformations

The Bi-invariant Mean as a Geometric Mean. Interestingly, it is possible to show that in the linear group $GL(n)$, the determinant of the bi-invariant mean is equal to the *scalar geometric mean* of the determinant of the data. This mean can thus be looked upon as a *generalization to invertible linear transformations of the geometric mean of positive numbers*. Remarkably, this generalization is not the only possible one, since an other (simpler) generalization of the geometric mean exists, called the Log-Euclidean mean, which we described in [18] in the case of tensors and in [10] for affine transformations. However, the Log-Euclidean mean is restricted to linear transformations whose principal logarithm is well-defined, which is *not* the case for the bi-invariant mean.

Proposition 13. *Let T_i be N linear transformations in $GL(n)$ and let (w_i) be N (normalized) non-negative weights, such that their bi-invariant mean $\mathbb{E}(T_i)$ exists. Then, we have:*

$$\begin{cases} \textit{if } det(S_i) > 0, \textit{ for all } i, \textit{ then } det(\mathbb{E}(T_i)) = e^{\sum_i w_i . \ln(det(S_i))}, \\ \textit{if } det(S_i) < 0, \textit{ for all } i, \textit{ then } det(\mathbb{E}(T_i)) = -e^{\sum_i w_i . \ln(-det(S_i))}, \end{cases}$$

Proof. When the bi-invariant mean is well-defined, then all the data must have a determinant with the same sign, which is also the sign of $\mathbb{E}(T_i)$. Otherwise, one of the products $\mathbb{E}(T_i)^{-1}.T_i$ would have a negative determinant and its principal logarithm would fail to exist.

To prove our result, we will rely only on two ingredients: the barycentric equation (10) and the following property: $\det(M) = \exp(\text{Trace}(\log(M)))$, which holds for any square matrix with a principal logarithm. This (classical) equality can be shown for example using the Jordan (or Schur) decomposition of the matrix M.

Taking the trace of the barycentric equation and then the (scalar) exponential, we get:

$$1 = \Pi_i e^{w_i \text{Trace}\left(\ln\left(\det\left(\mathbb{E}(T_i)^{-1}.T_i\right)\right)\right)}.$$

Then, using $\det(A.B) = \det(A).\det(B)$ and $\ln(a.b) = \ln(a) + \ln(b)$ and $\det(\mathbb{E}(T_i)^{-1}.T_i) = |\det(\mathbb{E}(T_i))|^{-1}.|\det(S_i)|$, we get the geometrical interpolation of determinants:

$$1 = \left|\det\left(\mathbb{E}(T_i)\right)\right|^{-1} . \exp\left(\sum_i w_i . \ln\left(|\det(S_i)|\right)\right),$$

which yields the result.

□

Practical Computation of the Bi-invariant Mean. Whe have seen in Section 4 that an efficiently iterative scheme could be used to compute bi-invariant means. It relies on successive computations involving inversions, exponentials and logarithms. To actually compute numerically the exponential and logarithm, we recommend using modern and efficient algorithms like the 'Scaling and Squaring' method for the matrix exponential [28] and the 'Inverse Scaling and Squaring Method' [29] for the matrix logarithm.

In the case of rigid and affine transformations, one can use their representation by matrices given by homogeneous coordinates, and use the general iterative scheme on these matrices to compute their bi-invariant mean.

6.4 Tensors

Let us now say a few words about a very interesting submanifold (which is *not* a subgroup) of $GL(n)$: the set of all symmetric positive-definite matrices, denoted here $Sym_+^\star(n)$. By abuse of language, we will refer to this type of data as 'tensors'. This (nick-)name comes from 'Diffusion Tensors Imaging', which is an imaging modality whose importance has been steadily growing in the biomedical imaging commutity in past years [1].

A number of teams (including ours) proposed almost simultanously in 2004 to endow this space with *affine-invariant metrics*. These metrics provide a distance between tensors which is *completely independent* of the arbitrary choice coordinate system [32, 33, 34, 35]. In more recent work, with Pierre Fillard, we also proposed to endowed this space with a family of Riemannian metrics called 'Log-Euclidean', which also have excellent theoretical properties (e.g., a number of them are similarity-invariant) and yield very similar results, but are much simpler to use in practice.

Interestingly, the Fréchet mean associated to affine-invariant Riemannian metrics on the tensor space coincides with the bi-invariant mean of tensors! Indeed, the affine-invariant Fréchet mean $\mathbb{E}_{Aff}(S_1, ..., S_N)$ of N tensors $S_1, ..., S_N$ is defined implicitely by the following barycentric equation:

$$\sum_{i=1}^{N} \log(\mathbb{E}_{Aff}(S_1, ..., S_N)^{-1}.S_i) = 0, \tag{21}$$

which happens to be exactly our general equation (10) for bi-invariant means. Intuitively, this means that our bi-invariant mean naturally unifies into a very general framework a number of well-established notions of means for various types of data living in Lie groups (e.g tensors, rotations, translations).

7 Left-invariant Polyaffine Transformations

Before presenting our novel polyaffine framework let us briefly recall the two polyaffine frameworks we have already introduced in past years, which are described in detail in [36, 10, 11].

7.1 Polyaffine Transformations

The idea is to define transformations that exhibit a locally rigid or affine behavior, with nice invertibility properties. Following the seminal work of [37], we model such transformations by a finite number N of affine *components*. Precisely, each component i consists of an affine transformation T_i and of a non-negative *weight function* $w_i(x)$ which models its spatial extension: the influence of the i^{th} component at point x is proportional to $w_i(x)$. Furthermore, we assume that for all x, $\sum_{i=1}^{N} w_i(x) = 1$, i.e. the weights are normalized.

Fusion of Displacements. In order to obtain a global transformation from several weighted components, the classical approach to fuse the N components, given in [38], simply consists in averaging the associated displacements according to the weights:

$$T(x) = \sum_{i=1}^{N} w_i(x) T_i(x). \tag{22}$$

The transformation obtained using (22) is smooth, but this approach has one major drawback: although each component is invertible, the resulting global transformation is *not invertible* in general. To remedy this, it was proposed in [36] to rely on the averaging of some *infinitesimal* displacements associated to each affine component instead. The resulting global transformation is obtained by integrating an Ordinary Differential Equation (ODE), which is computationally more expensive but guarantees its invertibility and also yields a simple form for its inverse.

Log-Euclidean Polyaffine Transformations. However, the first polyaffine framework we proposed lacks some important properties: the inverse of a polyaffine transformation is not polyaffine in general, and the polyaffine fusion of affine components is not invariant with respect to a change of coordinate system.

This is the reason why we proposed a novel framework in [10, 11], called *Log-Euclidean polyaffine*, which overcomes these defects. We also showed that this novel type of geometrical deformations can be computed very efficiently (as well as their inverses) on regular grids, with a simple algorithm called the *Fast Polyaffine Transform*.

Let us now see what the Log-Euclidean polyaffine fusion consists of. Let (T_i) be N affine (or rigid) transformations, and let $(\log(T_i))$ be their logarithms. Using these logarithms, one can fuse the T_i infinitesimally according to the weights $w_i(x)$ with a *stationary* (or *autonomous*) ODE, called the 'Log-Euclidean polyaffine ODE'. In homogeneous coordinates, this ODE is the following:

$$\dot{x} = \sum_i w_i(x) \log(T_i).x, \tag{23}$$

which is a nice infinitesimal analogous of Eq. (22). The value at a point x of the Log-Euclidean polyaffine transformation (LEPT) defined by (23) is given by integrating (23) between time 0 and 1 with x as initial condition.

7.2 A Novel Type of Polyaffine Transformations

Why (Again) a Novel Polyaffine Framework? The properties of the Log-Euclidean polyaffine framework are excellent, and do not suffer from the defects of our first polyaffine framework. However, this Log-Euclidean framework is limited to rigid or affine transformations *whose principal logarithm is well-defined*, i.e. which are close enough to the identity. So far, we did not find this restriction limiting in our work on 3D locally affine registration [39], mainly because we perform first a global affine alignment of anatomies before any locally affine registration. Still, it would be very interesting to have an infinitesimal strategy of fusion capable of handling any type of local rigid or affine deformations (provided their dispersion is not too large, of course), *regardless of their distance to the identity.*

Left-Invariant Polyaffine Transformations (LIPTs). To define our novel polyaffine fusion, called *left-invariant*, we will rely on bi-invariant means of rigid or affine transformations.

Let (T_i) be N affine (or rigid) transformations, let $(w_i(x))$ be some non-negative weights functions. Let also (α_i) be N non-negative weights, which intuitively correspond to the *global* weights of components, whereas weight functions provide *local* information. Finally, let $\bar{T}$ be the weighted left-invariant mean of the T_i and the weights (α_i). The bi-invariant polyaffine transformation Φ associated to all of these data is defined as follows:

1. Starting from a position x_0, the following ODE (in homogeneous coordinates) is integrated during one unit of time, which yields a final position $\Psi(x_0)$:

$$\dot{x} = \sum_i w_i(x) \log(\bar{T}^{-1}.T_i).x. \qquad (24)$$

2. We obtain the value at x_0 of the LIPT Φ by computing: $\Phi(x_0) = \bar{T}.\Psi(x_0)$.

This allows to fuse infinitesimally the T_i provided only that the logarithms of the $\bar{T}^{-1}.T_i$ exist, which does not require that *any* of the logarithms of the T_i exist. The properties of this fusion are quite nice: left-invariance and affine-invariance. They are summarized in the following Proposition:

Proposition 14. *The left-invariant polyaffine fusion Φ of the components $(T_i, w_i(x))$ with respect to the global weights (α_i) has the following invariance properties:*

- *left-invariance: any left-multiplication (by an affine transformation) of the T_i results in a left-multiplication of Φ*
- *affine-invariance: the fusion does not depend on the current choice of coordinate system*
- *at any point x such that $w_i(x) = \alpha_i$ for all i, we have $\Phi(x) = \bar{T}.x$, i.e. x moves according to the mean transformation.*

Proof. Proof of left-invariance: let A be an affine transformations, and let us replace the T_i by $A.T_i$. By construction, $\bar{T}$ is bi-invariant and is replaced by $A.\bar{T}$ and thus Eq. (24) remains unchanged:

$$\dot{x} = \sum_i w_i(x) \log(\bar{T}^{-1}.A^{-1}.A.T_i).x = \sum_i w_i(x) \log(\bar{T}^{-1}.T_i).x.$$

Since the value of Φ is obtained by left-multiplying by $\bar{T}$ in a second step, Φ is replaced by $A.\Phi$, which means that our novel polyaffine fusion is left-invariant.

Affine-invariance: let us change the current coordinate system by transforming x into $y \stackrel{def}{=} A.x$ in homogeneous coordinates. This results in the following changes:

- a weight function $x \mapsto w_i(x)$ becomes $y \mapsto w_i(A^{-1}.y)$.
- an affine transformation T_i becomes $A.T_i.A^{-1}$.
- the bi-invariant mean becomes $A.\bar{T}.A^{-1}$.

In the new coordinate systems, the left-invariant polyaffine ODE becomes:

$$\dot{y} = \sum_i w_i(A^{-1}.y) \log(A.\bar{T}^{-1}.A^{-1}.A.T_i.A^{-1}).y = A.\left(\sum_i w_i(A^{-1}.y) \log(\bar{T}^{-1}.T_i)\right).A^{-1}.y,$$

which yields:

$$\frac{d}{dt}(A^{-1}.y) = \sum_i w_i(A^{-1}.y) \log(\bar{T}^{-1}.T_i).(A^{-1}.y). \tag{25}$$

This means that $x(t)$ is a solution of Eq. (24) if and only if $y(t) = A.x(t)$ is a solution of (25), i.e. of the left-invariant polyaffine ODE in the novel coordinate system. As a consequence, a change of coordinate system does not affect the left-invariant polyaffine fusion, i.e. this fusion is affine-invariant.

Finally, at a point x such that $w_i(x) = \alpha_i$, we have:

$$\sum_i w_i(x).\log(\bar{T}^{-1}.T_i).x = \left(\sum_i \alpha_i \log(\bar{T}^{-1}.T_i)\right).x = 0.x = 0,$$

by construction of the bi-invariant mean $\bar{T}$. This implies that X is a fixed point of Ψ. Therefore, we have $\Phi(x) = \bar{T}.x$ at this point. □

Log-Euclidean vs. Left-Invariant Polyaffine Transformations. The price paid for the infinitesimal fusion of local rigid or affine components regardless of their distance to the identity is the following: the fusion is not *inversion-invariant*, i.e. the inverse of a left-invariant polyaffine transormation is not left-invariant polyaffine (but *right-invariant* polyaffine in fact). However, affine-invariance is preserved, i.e. independence with respect to the choice of coordinate system.

Computing Left-Invariant Polyaffine Transformations. The first step of the left-invariant polyaffine fusion, the ODE (24) is a simple Log-Euclidean polyaffine ODE, which can be integrated very efficiently using the Fast Polyaffine Transform described in [10]. The second step is very simple to compute: it only consists in applying an affine transformation, which is the bi-invariant mean $\bar{T}$. We recall that this mean can be very efficiently numerically computed in homogeneous coordinates using the classical iterative scheme described previously in this article.

Right-Invariant Polyaffine Transformations (RIPTs). We have just defined LIPTs. What about right-invariant polyaffine transformations? With the same notations as before, one can indeed define such transformations, by slightly modifying the left-invariant polyaffine ODE:

1. Starting from a position x_0, the following ODE (in homogeneous coordinates) is integrated during one unit of time, which yields a final position $\Psi_R(x_0)$:

$$\dot{x} = \sum_i w_i(\bar{T}.x) \log(\bar{T}^{-1}.T_i).x. \tag{26}$$

2. We obtain the value at x_0 of the right-invariant polyaffine transformation Φ_R by computing: $\Phi_R(x_0) = \bar{T}.\Psi_R(x_0)$.

With exactly the same type of techniques as in the proof of Proposition 14, one can show that type of fusion is right-invariant and also affine-invariant. It is much less intuitive, since the weight functions w_i are geometrically deformed by $\bar{T}^{-1}$ before being used in the fusion. With have the interesting following relationship between LIPTs and RIPTs:

Proposition 15. *The inverse of a LIPT is the RIPT with inverted components, and vice versa.*

Proof. Inverting the left-invariant polyaffine fusion results in the following two steps:

1. multiplying by $\bar{T}^{-1}$
2. integrating the following ODE during one unit of time:

$$\dot{x} = -\sum_i w_i(x) \log(\bar{T}^{-1}.T_i).x = \sum_i w_i(x) \log({T_i}^{-1}.\bar{T}).x \tag{27}$$

Let us use the change of variable $y = \bar{T}.x$ in (27). This yields:

$$\dot{x} = \bar{T}^{-1}.\dot{y} = \sum_i w_i(\bar{T}^{-1}.y) \log({T_i}^{-1}.\bar{T}).\bar{T}^{-1}.y = \bar{T}^{-1}\left(\sum_i w_i(\bar{T}^{-1}.y). \log(\bar{T}.T_i^{-1})\right).y,$$

which yields the simpler equation:

$$\dot{y} = \sum_i w_i(\bar{T}^{-1}.y) \log(\bar{T}.T_i^{-1}).y, \tag{28}$$

which is the ODE associated to the first step of the right-invariant fusion of the inverses of the T_i with the same weights as originally.

Thus, $x(t)$ is the solution of (27), second step of the inversion of our LIPT if and only if $\bar{T}.x(t)$ is a solution of the first step of the RIPT with inverted components. As consequence, the two steps of the inversion of our LIPT are equivalent to the two steps the RIPT with inverted components. the inverse of a LIPT is therefore the RIPT with inverted components and *vice versa.*

□

And Bi-invariant Polyaffine Transformations? So far, we have not been able to define a *bi-invariant* polyaffine fusion, i.e. an infinitesimal fusion of local affine transformations which would be simultanously left- and right- and inversion-invariant. Is such a fusion possible? We do not know yet, and this will be the subject of future work.

8 Conclusions and Perspectives

In this work, we have presented a general framework to define rigorously a novel type of mean in Lie groups, called the *bi-invariant mean.* This mean enjoys many desirable *invariance properties*, which generalize to the non-linear case the properties of the arithmetic mean: it is invariant with respect to left- and right-multiplication, as well as inversion. Previously, this type of mean was only defined in Lie groups endowed with a bi-invariant Riemannian metric, like compact Lie groups such as the group of rotations [7, 8]. But Riemannian bi-invariant metrics do not always exist. In particular, we have proved in this work that such metrics do not exist in any dimension for *rigid transformations*, which form but the most simple Lie group involved in bio-medical image registration.

To overcome the lack of existence of bi-invariant Riemannian metrics for many Lie groups, we have proposed in this article to define bi-invariant means in *any finite-dimensional real Lie group* via a general *barycentric equation*, whose solution is by definition the bi-invariant mean. We have shown the existence and uniqueness of this novel type of mean, provided the dispersion of the data is small enough, and the convergence of an efficient iterative algorithm for computing this mean has also been shown. The intuition of the existence of such a mean was first given in [9] (without any precise definition), along with an efficient algorithm for computing it (without proof of convergence), in the case of *matrix* groups.

In the case of rigid transformations, we have been able to determine a simple criterion for the general existence and uniqueness of the bi-invariant mean, which happens to be the same as for rotations. We have also given closed forms for the bi-invariant mean in a number of simple but instructive cases, including 2D rigid transformations. Interestingly, for general linear transformations, we have shown that similarly to the Log-Euclidean mean, that we recently proposed in [10], the bi-invariant mean is a generalization of the (scalar) geometric mean, since the determinant of the bi-invariant mean is exactly equal to the geometric mean of the determinants of the data.

Last but not least, we have used this new type of mean to define a novel class of polyaffine transformations, called *left-invariant polyaffine*, which allows to fuse local rigid or affine components *arbitrarily far away from the identity*, contrary to Log-Euclidean polyaffine fusion, which we recently introduced in [11].

In future work, we are planning to compare the statistics obtained via the bi-invariant mean to other types of statistics on rigid or affine transformations such as Log-Euclidean statistics [10], or left-invariant Riemannian statistics [5]. These statistics could prove very useful for example to constraint locally rigid or affine registration algorithms such as the one described in [39].

References

[1] Denis Le Bihan. Diffusion MNR imaging. *Magn Reson Q*, 7:1–30, 1991.

[2] Peter J. Basser and Derek Jones. Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR in Biomedicine*, 15(7-8):456–467, 2002.

[3] E. Özarslan and T. H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetique Resonance in Medecine*, 50:955–965, 2003.

[4] 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. *Journal of Mathematical Imaging and Vision*, 2006. In press.

[5] J. Boisvert, X. Pennec, N. Ayache, H. Labelle, and F. Cheriet. 3D anatomic variability assesment of the scoliotic spine using statistics on lie groups. In *Proceedings of the Third IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, Crystal Gateway Marriott, Arlington, Virginia, USA, 2006. To appear.

[6] Alain Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.

[7] Xavier Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 2006. To appear. A preliminary version is available as INRIA RR-5093, January 2004.

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

[9] Roger P. Woods. Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. *Neuroimage*, 18(3):769–88, 2003.

[10] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A fast and log-euclidean polyaffine framework for locally affine registration. Research report, INRIA Sophia-Antipolis, March 2006.

[11] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A log-euclidean polyaffine framework for locally rigid or affine registration. In *Proceedings of the Third International Workshop on Biomedical Image Registration (WBIR'06)*, LNCS, Utrecht, The Netherlands, 9 - 11 July 2006. Springer. To appear.

[12] Shlomo Sternberg. *Lectures on Differential Geometry.* Prentice Hall Mathematics Series. Prentice Hall Inc., 1964.

[13] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian Geometry.* Springer-Verlag, 2nd edition edition, 1993.

[14] Nicolas Bourbaki. *Elements of Mathematics: Lie Groups and Lie Algebra, Chapters 1-3.* Springer-Verlag, 2nd printing edition, 1989.

[15] Serge Lang. *Algebra.* Graduate Texts in Mathematics. Springer, 3rd rev. ed. 2002. corr. 4th printing edition, 2004.

[16] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 2006. Accepted subject to minor revision.

[17] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 2006. Accepted subject to minor revision.

[18] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Fast and simple calculus on tensors in the log-Euclidean framework. In J. Duncan and G. Gerig, editors, *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part I*, volume 3749 of *LNCS*, pages 115–122, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag.

[19] Milos Zefran, Vijay Kumar, and Croke Christopher. Metrics and connections for rigid-body kinematics. *International Journal of Robotics Research,*, 18(2):243–258, 1999.

[20] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: an Elementary Introduction*, volume 222 of *Graduate Texts in Mathematics.* Springer Verlag, 2003.

[21] D.H. Sattinger and O.L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, volume 61 of *AMS.* Springer Verlag, 1986.

[22] Michael Wüstner. A connected lie group equals the square of the exponential image. *Journal of Lie Theory*, 13:307–309, 2003.

[23] M. L. Curtis. *Matrix Groups.* Springer-Verlag, New York, Heidelberg, 1979.

[24] R. Godement. *Introduction à la Théorie des Groupes de Lie, Tomes I et II.* Publications Mathématiques de l'Université Paris VII, 1982.

[25] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple computations on tensors with log-Euclidean metrics. Research Report RR-5584, INRIA, Sophia-Antipolis, France, May 2005.

[26] M. Tenenbaum and H. Pollard. *Ordinary Differential Equations.* Dover, 1985.

[27] H. Samelson. On the Brouwer fixed point theorem. *Portugal. Math.*, 22:264–268, 1963.

[28] N. J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM J. Matrix Anal. Appl.*, 26(4):1179–1193, 2005.

[29] S. Hun Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub. Approximating the logarithm of a matrix to specified accuracy. *SIAM J. Matrix Anal. Appl.*, 22(4):1112–1125, 2001.

[30] X. Pennec. Computing the mean of geometric features - application to the mean rotation. Research Report RR-3371, INRIA, March 1998.

[31] Xavier Pennec. Probabilities and statistics on Riemannian manifolds: A geometric approach. Research Report 5093, INRIA, January 2004. An extended version will appear in the Int. Journal of Mathematical Imaging and Vision.

[32] Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006. A preliminray version appeared as INRIA Research Report 5255, July 2004.

[33] P. T. Fletcher and S. C. Joshi. Principal geodesic analysisee on symmetric spaces: Statistics of diffusion tensors. In *Proc. of CVAMIA and MMBIA Workshops, Prague, Czech Republic, May 15, 2004*, LNCS 3117, pages 87–98. Springer, 2004.

[34] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor MRI. Research Report RR-5242 and RR-5243, INRIA, June 2004.

[35] M. Moakher. A differential geometry approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26:735–747, 2005.

[36] Vincent Arsigny, Xavier Pennec, and Nicholas Ayache. Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. *Medical Image Analysis*, 9(6):507–523, December 2005.

[37] J.A. Little, D.L.G. Hill, and D.J. Hawkes. Deformations incorpotationg rigid structures. *Computer Vision and Image Understanding*, 66(2):223–232, May 1996.

[38] D. Sheppard. A two-dimensionnal interpolation function for irregularly spaced data. In *23rd National Conference of the ACM*, pages 517–524. ACM Press, 1968.

[39] Olivier Commowick, Vincent Arsigny, Jimena Costa, Nicholas Ayache, and Grégoire Malandain. An efficient locally affine framework for the registration of anatomical structures. In *Proceedings of the Third IEEE International Symposium on Biomedical Imaging (ISBI 2006)*, Crystal Gateway Marriott, Arlington, Virginia, USA, 2006. To appear.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399