Defining a mean on Lie group

Nina Miolane

Imperial College London

Master's Thesis

Defining a mean on Lie groups

Author:
Nina Miolane

Supervisors:
Prof. Xavier Pennec
Prof. Kellogg Stelle

A thesis submitted in partial fulfilment of the requirements for the degree of MSc Quantum Fields and Fundamental Forces

in the

Theoretical Physics Department

September 20, 2013

IMPERIAL COLLEGE LONDON

Abstract

Theoretical Physics Department

MSc Quantum Fields and Fundamental Forces

Defining a mean on Lie groups

by Nina MIOLANE

Statistics are mostly performed in vector spaces, flat structures where the computations are linear. When one wants to generalize this setting to the non-linear structure of Lie groups, it is necessary to redefine the core concepts. For instance, the linear definition of the mean as an integral can not be used anymore. In this thesis, we investigate three possible definitions depending of the Lie group geometric structure. First, we import *on Lie groups* the notion of Riemannian center of mass (CoM) which is used to define a mean *on manifolds* and investigate when it can define a mean which is compatible with the algebraic structure. It is the case only for a small class of Lie groups. Thus we extend the CoM's definition with two others: the Riemannian exponential barycenter and the group exponential barycenter. This thesis investigates how they can define admissible means on Lie groups.

If one needs to perform statistics on a structure, it seems judicious to respect this structure. In the case of the mean of Lie group elements, it is desirable for instance that it belongs to the Lie Group and is stable by the group operations: composition and inversion. This property is ensured for the Riemannian center of mass if the metric is bi-invariant, like for compact Lie groups (e.g. rotations SO(3)). The Riemannian exponential barycenter is also an admissible definition if the pseudo-metric is bi-invariant, for example if the Lie group can be decomposed into a specific semidirect product (e.g. group of rigid transformations SE(3)). However, bi-invariant pseudo-metrics do not exist for all Lie groups. This is the case in particular for solvable Lie groups with null center like the Heisenberg group. We introduce the group exponential barycenter which is naturally consistent with the group operations, even in the infinite dimension case. On the Lie groups where these definitions are admissible, we discuss the existence and uniqueness of the mean on a global domain, called the mean maximal domain.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Xavier Pennec. Without his guidance and motivation, this master thesis would not have been possible.

I am also indebted to Imperial College and more precisely to Professor Kellogg Stelle, for his help toward the realization of this work.

I would like to thank all members of the Asclepios Team for their warm welcome at INRIA and in particular Krissie and Marine for their daily enthousiasm in our office.

I also wish to express my sincere thanks to Davy for his supportive presence during the writing of this thesis and his capacity of concentration while working which should be a model for everyone.

I do not forget how grateful I am to Andrea, for being italian and perfect, and above all for his constant support during this master thesis.

Finally, I would like to thank my family, and particularly my parents for my education which has enabled me to reach this level of study.

Contents

Abstract ii

Acknowledgements iii

Contents iv

List of Figures vii

1 Introduction 1

2 Lie group: geometric and algebraic structures 5
- 2.1 Generalizing the linear statistical framework on curved spacs . . . 5
- 2.2 Affine connection and Riemannian spaces . . . 6
  - 2.2.1 Affine connection spaces . . . 7
  - 2.2.2 Pseudo-Riemannian manifold . . . 10
- 2.3 Lie groups . . . 13
  - 2.3.1 Lie group and Lie algebra . . . 14
  - 2.3.2 Adjoint and co-adjoint representations . . . 15
  - 2.3.3 Lie group exponential and logarithm . . . 18

3 Defining the mean on a Lie group 27
- 3.1 Extension of the mean's definition . . . 27
  - 3.1.1 Relying on the pseudo-Riemannian structure . . . 28
  - 3.1.2 Relying on the algebraic structure . . . 30
- 3.2 Metric, connection and group geodesics . . . 31
  - 3.2.1 Cartan-Schouten connections . . . 32
  - 3.2.2 Normal elements . . . 34
- 3.3 The example of SE(3) . . . 36
  - 3.3.1 SE(3) and $\mathfrak{se}(3)$ . . . 36
  - 3.3.2 Geodesics of SE(3) . . . 37
  - 3.3.3 Normal elements of $\mathfrak{se}(3)$ . . . 39
  - 3.3.4 Three different means on SE(3) . . . 41
- 3.4 Admissible definitions: consistence with the Lie group structure . . . 43
  - 3.4.1 Bi-invariance . . . 43
  - 3.4.2 Special case of a positive definite metric . . . 45

4 Existence of a bi-invariant pseudo-metric 49
- 4.1 A try with Levi decomposition of Lie algebras . . . 49
  - 4.1.1 Ideals and semi-direct sums of Lie algebras . . . 49
  - 4.1.2 Simple and semi-simple Lie algebras . . . 51

4.1.3 Solvable Lie algebras . . . 53
4.2 A specific decomposition . . . 56
4.2.1 Ideals of $\mathfrak{g}$ with a bi-invariant pseudo-metric . . . 56
4.2.2 Minimal ideals of $\mathfrak{g}$ . . . 58
4.2.3 Decomposition of $\mathfrak{g}$ . . . 59
4.3 Construction of the bi-invariant pseudo-metric . . . 60
4.3.1 Lie groups with bi-invariant pseudo-metric . . . 60
4.3.2 A bi-invariant pseudo-metric on SE(3) . . . 61

5 Existence and uniqueness of the means 65
5.1 Riemannian CoMs . . . 65
5.1.1 Notions of convexity . . . 66
5.1.2 Existence and uniqueness . . . 67
5.1.3 Application to Lie groups: uniqueness domain . . . 74
5.2 Pseudo-Riemannian, group exponential barycenters . . . 76
5.2.1 Existence and uniqueness . . . 77
5.2.2 Application to Lie groups: some insights on the uniqueness domain . . . 83

6 Conclusion 85

A Scilab code 87

Bibliography 103

List of Figures

2.1 Re-interpretation of basic standard operations in a Riemannian manifold. 6
2.2 Shooting from the tangent space to the manifold using the exponential map . . . 9

3.1 Definition of an exponential barycenter, here in the Riemannian case. 29
3.2 Group (red) and riemannian left (green) and right (blue) geodesics starting at e with tangent vector [0 0 pi/3 1 0 0] 39
3.3 Group (red) and riemannian left (green) and right (blue) geodesics linking two points of SE(3) 39
3.4 Riemannian left (green), right (blue) and group geodesics starting at e with tangent vector [0 0 pi/5 0 0 1], i.e. a normal element 42
3.5 Riemannian left (green), right (blue) centers of mass and group exponential barycenter (red) for a given data set (black) 43

5.1 Bounded convex geometry for a regular geodesic ball of $\mathcal{S}^1$ 68
5.2 The open hemisphere is a CSLCG manifold. 77
5.3 Existence of an exponential barycenter for a relatively compact domain of the form $\phi < 0$ 78
5.4 Definition of U so that each x has an unique ϕ-projection on K. 80

To my parents . . .

Chapter 1

Introduction

Statistics is an established field of fundamental mathematics which has enabled notable advancements in applied sciences. It represents a convenient tool to tackle experimentally scientific problems when a theoretical model is not a priori provided. For example, the number of degree of freedom implied in the description of an organ's shape makes it difficult to compute a physical model. A possible statistical method dealing with this issue relies on medical image computing. We identify anatomically representative geometric features and study their statistical distribution across a population in order to postulate symptoms for diseases.

In order to model biological shapes without having to embed them in the Euclidean space, D'Arcy Thompson proposes in 1917 an efficient framework relying on Lie groups. He considers any anatomy as a diffeomorphic deformation of a reference shape, the *template object* [1]. The problem of performing statistics on complex geometric surfaces or volumes reduces to the creation of a statistical framework on the space of transformations, which naturally belongs to a group of smooth deformations, i.e. a Lie group. The most general group to consider is the infinite dimensional group of diffeomorphisms, whose formalism is developed in particular by Grenander and Miller ([2], [3]). However, we restrict in this master thesis to finite dimensional Lie groups and more precisely matrix Lie groups.

The most fundamental notion in statistics is the *mean*, which is usually defined as a weighted sum of the data elements. In the non-linear context of a Lie group, this linear definition has to be modified. Some alternatives have been proposed in the literatur, among them the Riemannian center of mass when the Lie group is provided with a Riemannian metric. In fact, the Riemannian framework has been extensively studied and proved to be powerful in order to compute statistics on non-linear structures ([4], [5], [6]). Hence, it represents the starting point of our approach. However, the definition of the Riemannian CoM on a Lie group appears to be consistent with the algebraic operations only if the metric is shown to be bi-invariant. This is the case for compact groups such as rotations SO(3), but a bi-invariant metric fails to exist for some basic Lie groups of image computing as the group of rigid transformations SE(3).

Hence, an extension of this mean's definition has to be provided. In this thesis, we recall the properties and limits of the Riemannian CoM and we investigate two alternatives. The first one is the *Riemannian exponential barycenter*, which represents a good definition for Lie groups provided with a bi-invariant *pseudo-metric*. It enables to extend the class of Lie groups provided with a consistent definition of mean, solving for example the case of SE(3). However, bi-invariant pseudo-metrics still don't exist on *all* Lie groups, for example the Heisenberg group can not be provided with any bi-invariant pseudo-metric. Morever, if one wants to generalize the setting to infinite dimensional Lie groups, it would be desirable to have a universal notion of mean on matrix Lie groups before extending it. To this purpose, we also define the *group exponential barycenter* of a data set on a Lie group. This one is shown to possess naturally the consistence property with the group operations. The generalization of the Riemannian CoM to the Riemannian exponential barycenter or the group exponential barycenter relies on the relaxation of the Lie group's geometric structure. From the usual Riemannian setting, with a metric or a pseudo-metric, we generalize to the affine connection framework.

We first recall in Chapter 2 concepts of Lie groups and start with the different geometric structures we can provide them with. We present affine connections, pseudo-Riemannian and Riemannian metrics, insisting on the tools which will be useful from a computational point of view. We emphasize to this aim the notions of *exponential map* and *metric or connection geodesics*. Then we turn to the algebraic side of a connected Lie group, recalling its relation with the Lie algebra. We present one of the most powerful tool for its algebraic study, the adjoint representations. Without any metric, we define the notion of *group geodesics*.

Once this setting is defined, we can delve into the problem of an *admissible definition* for the mean on a Lie group in Chapter 3. We present the definitions of the Riemannian CoM, the Riemannian exponential barycenter and the group exponential barycenter. When the Lie group is provided with the most powerful structure, i.e. the Riemannian metric, it would be desirable that these definitions provide the same result. This is the *first admissibilty condition* if one wants to define a generalization of the Riemannian CoM. We show that this is the case precisely when the metric geodesics correspond to the group geodesics. This implies that the Levi-Civita connection associated to the metric is a *Cartan-Schouten connection*. If not, we are provided with different means for the same given data set on the Lie group. We illustrate this situation, and the whole computational framework, with the example of SE(3). In the last section of this chapter, we investigate the *second admissibility condition* in order to define a mean on Lie group. It has to be consistent with the group operations, namely composition and inversion. This is the case when we can provide the Lie group with a bi-invariant metric or pseudo-metric. As the Riemannian case has been extensively studied, we presented as an example the characterization of Lie groups that admit such a *metric* [7].

We turn to the characterization of Lie groups which can be provided with a bi-invariant *pseudo-metric* in Chapter 4. To this aim, we delve into the classification of Lie groups while presenting a first attempt of a construction of a bi-invariant pseudo-metric. It relies on the Levi decomposition of a Lie group into a semidirect product of a *semisimple* and a *solvable* part. While semisimple groups can always be provided with a bi-invariant pseudo-metric, namely their Killing form, solvable groups fail to follow a general rule. We characterize the solvable part (as a co-adjoint representation) in order to conclude, before presenting a general method to construct a bi-invariant pseudo-metric when the conditions are fulfilled. We apply it to the case of rigid body transformations, where no bi-invariant metric exists but where there is a bi-invariant pseudo-metric.

Once we know which mean is admissible on which Lie groups, we investigate som existence and uniqueness properties for the three possible mean's definitions in Chapter 5. It represents the last issue before introducing them in an computational framework. We present the case of the Riemannian CoM in the first place, emphazing the problematic of convexity when dealing with the uniqueness problem. We characterize the *maximal domain* on which we have uniqueness of the COM and illustrate its derivation with the example of SO(3). In this case, it corresponds to the strict upper hemisphere. Finally, we extend the existence and uniqueness theorem to the case of pseudo-Riemannian spaces and affine connection spaces. We compute similarly the maximal domain of uniqueness for the group of rigid transfomations.

Chapter 2

Lie group: geometric and algebraic structures

2.1 Generalizing the linear statistical framework on curved spacs

Algorithms in statistics generally use the setting of (high dimensional) vector spaces. However, manifolds and Lie groups are non-linear structures. Hence, a generalization of the computational framework for Lie groups is needed. Recall that a Lie group as a manifold is still locally linear and can be approximated at a point g by its tangent plane $T_g\mathcal{G}$. Then, given a vector of the tangent plane $v \in T_g\mathcal{G}$, we may *shoot along this vector* to reach a corresponding point on the Lie group. This is the essence of the *exponential map*, $exp_g(v)$, which we'll study in this section. Similarly, we can compute the image, on the Lie group, of a straight line of the tangent space through this exponential map. It defines *geodesics*. The exponential map and its inverse, the logarithm map, associated to the geodesics of a manifold are key concepts in order to generalize linear algorithms on a Lie group structure.

With these tools, the generalization of the basic functions from the usual linear computing framework is straighforward using the notion of *bipoint* (oriented couples of points), an antecedent of vector. Indeed, one defines vectors as equivalent classes of bipoints in a Euclidean space. This is possible because the translation enables to compare what happens at two different points. In a Riemannian manifold, each vector has to remember at which point it is attached, which comes back to a bipoint. Now the exponential of a vector $\vec{gh}$ attached at g can be interpreted as the addition: $exp_g(\vec{gh}) = g + \vec{gh} = h$. On the other hand, the logarithm can be seen as the substraction: $log_g(h) = h - g = \vec{gh}$, mapping the bipoint (g,h) to the vector $\vec{gh}$ attached at g.

This reinterpretation of addition and subtraction using exponential and logarithm maps is very powerful to generalize algorithms working on vector spaces to algorithms on Riemannian manifolds. Figure 2.1 from [8] summarizes the generalization's method. It is very powerful, as most of the operations can be expressed in terms of the exponential and the logarithm maps, like the gradient descent algorithm.

	Vector space	Riemannian manifold
Subtraction	$\overrightarrow{xy} = y - x$	$\overrightarrow{xy} = \log_x(y)$
Addition	$y = x + \overrightarrow{xy}$	$y = \exp_x(\overrightarrow{xy})$
Distance	$\text{dist}(x, y) = \|y - x\|$	$\text{dist}(x, y) = \|\overrightarrow{xy}\|_x$
Mean value (implicit)	$\sum_i \overrightarrow{\bar{x}x_i} = 0$	$\sum_i \log_{\bar{x}}(x_i) = 0$
Gradient descent	$x_{t+\varepsilon} = x_t - \varepsilon \nabla C(x_t)$	$x_{t+\varepsilon} = \exp_{x_t}(-\varepsilon \nabla C(x_t))$
Linear (geodesic) interpolation	$x(t) = x_1 + t\,\overrightarrow{x_1 x_2}$	$x(t) = \exp_{x_1}(t\,\overrightarrow{x_1 x_2})$

FIGURE 2.1: Re-interpretation of basic standard operations in a Riemannian manifold.

In the following sections, we review the fundamental properties of Lie groups which are useful from a computational point of view. We specify the different notions of exponential maps and geodesics that can be defined, namely the metric exponential map, the connection exponential map and the group exponential map with their respective geodesics. The domain of definition of the exponential map depends on the space's completeness. In this thesis, all spaces are assumed to be *complete*, with regards to the definition that is appropriated in each case.

A metric space is *complete* if any Cauchy sequence in it converges. For a Riemannian manifold, we have the additional notion of *geodesic completeness*: the Riemannian manifold is geodesically complete if any maximal (inextendible) geodesic is defined for all $t \in \mathbb{R}$. The two definitions are in fact equivalent, and the manifold is called a *complete Riemannian manifold*. Hence for other structures without metric, as for affine connection spaces and for Lie groups, completeness will be defined as geodesic completeness. In all three cases, it is equivalent to the statement that the exponential map exp_g is defined on the whole tangent space $T_g\mathcal{G}$. This property is the reason why we only deal with complete spaces in this thesis. However, we should pay special attention to the definition of its *maximal injectivity domain*, which depends on the Lie group we consider and can be computed. This is linked with the maximal uniqueness domain for the corresponding definition of the mean.

2.2 Affine connection and Riemannian spaces

In this section, we denote by $T_g\mathcal{G}$ the tangent space at $g \in \mathcal{G}$ and by $T\mathcal{G}$ the tangent bundle. $\Gamma(\mathcal{G})$ is the set of vectors fields X which form the algebra of derivations of smooth functions $\phi \in \mathcal{C}^\infty(\mathcal{G})$. Indeed, we have the fundamental isomorphism of Differential Geometry at each point $g \in \mathcal{G}$:

$$\mathcal{D}_g(\mathcal{G}) \simeq T_g\mathcal{G}$$

where $\mathcal{D}_g(\mathcal{G})$ is the vector space of directional derivatives at g.

To see this, one may consider the map:

$$\begin{aligned} \chi : \qquad T_g\mathcal{G} &\longmapsto \mathcal{D}_g\mathcal{G} \\ [\mathcal{C}] &\longmapsto \chi([\mathcal{C}]) \end{aligned}$$

where $\mathcal{C}$ is a representative curve through $g = \mathcal{C}(0)$ for the vector attached at g. This map is defined by the action on functions:

$$\chi([\mathcal{C}])[\phi] \equiv \frac{d}{dt}\Big|_{t=0}(\phi \circ \mathcal{C}(t))$$

which holds for any $\phi \in \mathcal{C}^\infty(\mathcal{G})$ and is independent of the choice of the representative $\mathcal{C}$ for the vector of $T_g\mathcal{G}$. We refer to [9] to show that χ is onto and one-to-one.

In a local coordinate system, the action of a vector field X at g on ϕ is :

$$X[\phi]\Big|_g = \partial_X\phi\Big|_g = \frac{d}{dt}\Big|_{t=0}\phi(g + tX|_g).$$

When composing the derivatives, we see that $X[Y[\phi]]|_g = \partial_X\partial_Y\phi|_g$ involves second order derivations and thus does not behave like a vector (or a directional derivation). However, we can remove the second order terms by substracting $\partial_X\partial_Y$ to obtain the *Lie Bracket*:

$$[X, Y][\phi] = X[Y[\phi]] - Y[X[\phi]]$$

which is now a derivation. We may also write:

$$[X, Y] = \partial_X Y - \partial_Y X = (X^j\partial_j Y^i - Y^j\partial_j X^i)\partial_j.$$

Provided with the Lie Bracket [,], $\Gamma(\mathcal{G})$ is in fact a Lie algebra. Note that it is *not* the Lie algebra of the Lie group.

2.2.1 Affine connection spaces

A general notion of geometry on Lie groups is defined by the parallel transport. If one wants to compare data in the tangent space $T_g\mathcal{G}$ at point g, with data at another tangent space, say $T_h\mathcal{G}$, one needs a specific mapping Π_g^h between the two spaces. This is precisely the essence of parallel transport. As there is no general way to define a global linear mapping between two tangent spaces $\Pi_g^h : T_g\mathcal{G} \longmapsto T_h\mathcal{G}$, that is consistent with the composition $\Pi_g^f \circ \Pi_f^h = \Pi_g^h$, one has to specify by which path we connect g and h. It leads to the following definition.

Definition 2.1. (Parallel transport) Let γ be a curve on $\mathcal{G}$. A *parallel transport along* γ is a collection of mappings $\Pi(\gamma)_s^t : T_{\gamma(s)}\mathcal{G} \mapsto T_{\gamma(t)}\mathcal{G}$ such that:

- $\Pi(\gamma)_s^s = Id$ (identity transformation of $T_{\gamma(s)}\mathcal{G}$),
- $\Pi(\gamma)_u^t \circ \Pi(\gamma)_s^u = \Pi(\gamma)_s^t$ (consistency along the curve),
- The dependence of Π on γ, s and t is smooth.

The notion of (affine) connection is the infinitesimal version of parallel transport. It indicates how a vector transforms when it is parallely transported to an infinitesimaly close tangent space.

Let $x = \dot{\gamma}(0)$ be the tangent vector at the initial point of the curve γ and Y be a vector field. Then the quantity:

$$\lim_{t \to 0} \frac{\Pi(\gamma)_t^0 Y|_{\gamma(t)} - Y|_{\gamma(0)}}{t} = \frac{d}{dt}\Big|_{t=0} \Pi(\gamma)_t^0 Y|_{\gamma(t)} \qquad (*)$$

is independent of the curve γ. Hence we can forget about the previous expression of parallel transport and we define an affine connection without specifying a path.

Definition 2.2. (Affine connection) An *affine connection* is a bilinear map:

$$\begin{aligned} \nabla : \Gamma(\mathcal{G}) \times \Gamma(\mathcal{G}) &\longmapsto \Gamma(\mathcal{G}) \\ X, Y &\longmapsto \nabla_X Y \end{aligned}$$

with the following properties, for any $X, Y \in \Gamma(\mathcal{G})$ and $\phi \in \mathcal{C}^\infty$:

- $\nabla_{\phi X} Y = \phi \nabla_X Y$,
- $\nabla_X(\phi Y) = \partial_X(\phi Y) + \phi \nabla_X Y$ (Leibniz rule in the second variable).

One may check that the quantity $(*)$ verifies these properties. The connection ∇ is determined by its coordinates on the basis vector fields: $\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k$. The Γ_{ij}^k are called the *Christoffel symbols* of the connection.

Note that the properties satisfied by the connection are precisely those of a directional derivative of Y along X. $\nabla_X Y$ is also called the *covariant derivative* of the vector field Y along X. Its computation in a local coordinate system $(\partial_i)_i$ shows how it differs from the standard directional derivative along X:

$$\nabla_X Y = \partial_X Y + \Gamma_{ij}^k X^i Y^j \partial_k$$

where we recall that: $\partial_X Y = X^j \partial_j (Y^i) \partial_i$.

Remark 2.3. The covariant derivative along X can also be defined for any tensor field T.

If we take the covariant derivative of Y, along the tangent vector field of a curve parametrized by t, we get a function of $\mathbb{R}$.

Definition 2.4. (Covariant derivative along a curve) Let γ be a curve on $\mathcal{G}$. The *covariant derivative* of the vector field $Y \in \Gamma(\mathcal{G})$ along the curve γ is defined as:

$$\frac{DY}{dt} = \nabla_{\dot{\gamma}(t)} Y$$

Taking Y=$\dot{\gamma}$, we can study how the tangent vector field covariantly variates along its curve. If its variation is null, i.e. if it is parallely transported at any point of its curve, the curve is said to be parallel to itself. Hence we have a generalization of a straight line, which means: a definition of a geodesic.

Definition 2.5. (Connection geodesic) Let γ be a curve on $\mathcal{G}$. Then γ is a *connection geodesic* if its tangent vector field $\dot{\gamma}$ remains parallel to itself, in the sense of parallel transport along the curve. In the infinitesimal formulation of covariant derivative, this is equivalent to:

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 0$$

In a local coordinates system $(\partial_k)_k$, the equation of the geodesic is:

$$\ddot{\gamma}^k + \Gamma^k_{ij}\dot{\gamma}^i\dot{\gamma}^j = 0$$

Remark 2.6. Note that geodesics of a given connection only depend on the symmetric part $\frac{1}{2}(\nabla_X Y + \nabla_Y X)$ of ∇ (see [9]). The skew-symmetric part (i.e. the torsion, as we shall see later) only affects the parallel transport along γ of non-colinear vector fields.

Recall here our purpose of generalizing linear algorithms on the non-linear structure of manifolds. This implies precisely a generalization of straight lines, as for example this definition of geodesics in the affine connection case. Now, $T_g\mathcal{G}$ is the best linear approximation of $\mathcal{G}$ at g and the definition of an exponential map precisely defines the generalization process from linear to curved spaces. Using the connection exponential, we shoot the linear objects of the tangent space onto the manifold, for example the straight lines onto the geodesics. This shooting is illustrated in Figure 2.2.

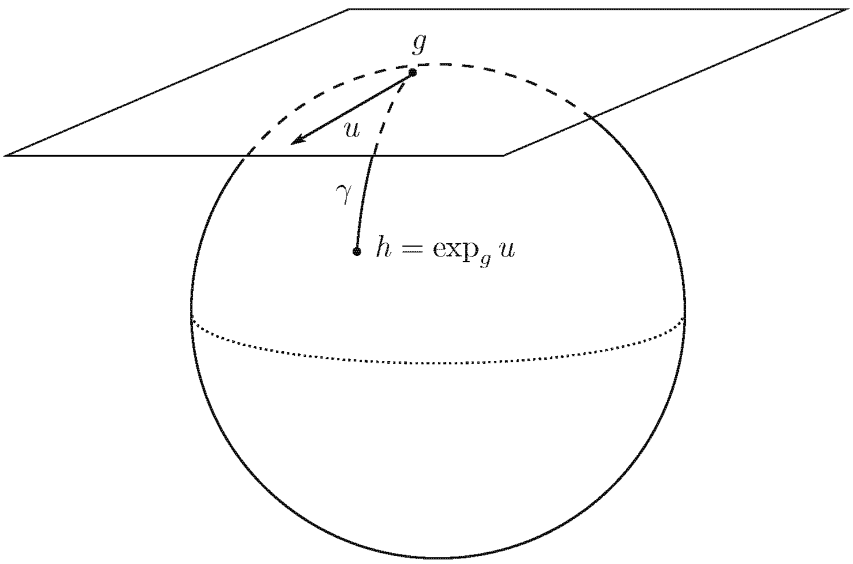

FIGURE 2.2: Shooting from the tangent space to the manifold using the exponential map

Definition 2.7. (Connection exponential map) The application mapping any vector $u \in T_g\mathcal{G}$ to the point h of $\mathcal{G}$ that is reached after a unit time by the geodesic starting at g with tangent vector u, is called the *connection exponential map* and we write: $exp_g(u) = h$.

The map is smooth (geodesics are smooth curves) and defined on the entire tangent space $T_g\mathcal{G}$ as we assumed the completeness of $\mathcal{G}$ in the sense of affine connection spaces. However, exp_g might not be onto nor one-to-one. But as its differential at 0 is the identity Id, we use the Inverse Function Theorem (finite dimensional space) to get its local injectivity [10]:

Theorem 2.8. *The connection exponential map is a diffeomorphism from an open neighborhood of 0 in $T_g\mathcal{G}$ to an open neighborhood of $exp_g(0) = g$ in $\mathcal{G}$.*

Hence, we can locally define its inverse, which is called the connection logarithm map.

Definition 2.9. (Connection logarithm) For every $g \in \mathcal{G}$, there is a open neighbordhood $\mathcal{O}_g$ where we can define the *connection logarithm* at g of any point $h \in \mathcal{O}_g$, $log_g(h)$, as the unique u in the corresponding neighbordhood of 0 in $T_g\mathcal{G}$ such that $exp_g(u) = h$. We write $u = log_g(h)$.

Remark 2.10. We don't delve into the global bijectivity domain of the connection exponential map here. However, we'll present its general meaning for the purpose of this thesis in the Riemannian case (next subsection).

Apart from the straight lines, an affine connection defines two more geometric tensors on the space: torsion and curvature.

Definition 2.11. (Torsion and Curvature) The *torsion* and *curvature* tensors are defined as:

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

The torsion tensor is obviously skew-symmetric: $T(u,v) = -T(v,u)$ and we have:

$$T(X,Y)^k = (\Gamma^k_{ij} - \Gamma^k_{ji}) X^i Y^j.$$

Hence, a symmetric connection is a torsion free connection. As for the curvature tensor, we state (without proof) its properties, as they will be useful for computations. Again, we refer to [9] for details.

Proposition 2.12. *For $g \in \mathcal{G}$ and $u, v, w, x, y \in T_g\mathcal{G}$, we have:*
(i) $R(u,v)w = -R(v,u)w$,
(ii) $R(u,v)w + R(v,w)u + R(w,u)v = 0$.

In this introductive section, we have defined geodesics, torsion and curvature without the need of any particular (pseudo-)metric. Now introduce one.

2.2.2 Pseudo-Riemannian manifold

An affine connection space is provided with the notion of parallelism but no distance is defined. This is the role of a pseudo-metric.

Definition 2.13. (Riemannian pseudo-metric) A *Riemannian pseudo-metric* is a smooth collection of definite bilinear 2-forms on tangent spaces of the manifold, i.e. it defines the dot product of tangent vectors of the same space.
We denote (p,q) its signature, $p + q = \dim \mathcal{G}$, where p is the number of strictly positive eigenvalues, and q is the number of strictly negative ones.

A Riemannian pseudo-metric of signature $(\dim \mathcal{G}, 0)$ is positive definite and called a *Riemannian metric*.

Remark 2.14. In the following, we might also denote the pseudo-metric g_{ij} in order to use Einstein summations when it clarifies the expressions.

Given $u, v \in T_g\mathcal{G}$, a Riemannian pseudo-metric enables to compute their norms and angle between them. Moreover, it defines a pseudo-distance along a curve of the manifold $\mathcal{G}$ through integration of the norm of its tangent vector. In this framework, we have another definition of geodesics, namely the *Riemannian geodesics.*

Definition 2.15. (Riemannian geodesic) A *Riemannian geodesic* $\mathcal{C}$ of $\mathcal{G}$, in affine parametrization s, is a curve minimizing the functionnal:

$$I[\mathcal{C}] = \int_{\mathcal{C}} ds \sqrt{< \dot{\gamma(x)}, \dot{\gamma(s)} > |_{\gamma(s)}} \quad \text{(i.e. the length)},$$

or, equivalenly, the functionnal:

$$K[\mathcal{C}] = \frac{1}{2} \int_{\mathcal{C}} ds < \dot{\gamma(x)}, \dot{\gamma(s)} > |_{\gamma(s)} \quad \text{(i.e. the cinetic energy)}.$$

Remark 2.16. The equivalence of the minimization of the two different functionnals is intuitive from their physical interpretations but not trivial in the computations [9]. Note that one can always reparametrize a geodesic so that its speed is 1, i.e. $|\dot{\gamma(t)}| = 1$, for all t. In the following, we only deal with such geodesics.

Now we have two possible generalizations of straight lines: the connection geodesics and the Riemannian geodesics. But, a Riemannian manifold is a special case of an affine connection space. Here we investigate the compatibility condition one should ask from a connection, in order for its connection geodesics to agree with the Riemannian ones.

Definition 2.17. For a given pseudo-metric g, a connection ∇ is *compatible with g* if at any point $g \in \mathcal{G}$ we have:

$$\nabla_X g = 0 \quad \forall X \in T_g\mathcal{G}$$

or equivalently, if we have:

$$X < Y, Z > = < \nabla_X Y, Z > + < Y, \nabla_X Z > .$$

This means that the pseudo-metric is parallely transported along any curve.

Proposition 2.18. *If a connection ∇ is compatible with the metric $<,>$, then the connection geodesics are the Riemannian geodesics.*

Once again, we refer to [9] for the proof. Now, for a given metric, we can always find a connection compatible with it. Moreover, we can characterize all the connections that are compatible with it [10].

Definition 2.19. (Levi-Civita connection) For a given pseudo-metric g, there is a unique symmetric (i.e. torsion free) connection compatible with it. It is called the *Levi-Civita connection* and can be computed with the Koszul formula as follow:

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij})$$

or:

$$\begin{aligned} < \nabla_X Y, Z > = \frac{1}{2}(X < Y, Z > + Y < X, Z > - Z < X, Y > \\ - < X, [Y, Z] > - < Y, [X, Z] > + < Z, [X, Y] >). \end{aligned}$$

Remark 2.20.

- This is consistent with the fact that connection geodesics only depend on the symmetric part of ∇. Note also that this definition/proposition gives all the connections compatible with a given metric $<,>$, we only add a term of torsion to the Levi-Civita part.
- Given a metric, we can always find a connection compatible with it. However, the converse is false. We called *metric connection*, a connection that has a metric compatible with it, *non-metric connection* if one cannot find such a metric.

Now take the Levi-Civita connection associated to $<,>$ to import the results from affine connection spaces to the pseudo-Riemannian manifold. We define the *pseudo-Riemannian exponential*, the *pseudo-Riemannian exponential chart* and the *pseudo-Riemannian logarithm*.

The pseudo-Riemannian structure itself provides two additionnal properties to Proposition 2.12 for the curvature tensor R [9]:

- $< R(x, y)u, v > = < R(x, y)v, u >$,
- $< R(x, y)u, v > = < R(u, v)x, y >$.

Moreover, we have precision about the bijectivity domain of the exponential map. Recall that in the affine connection case, we only define its inverse *locally* according to Theorem 2.8. However, in the Riemannian case, we have precisions on the global bijectivity domain of the exponential map at g [8]. It is a star-shaped domain delimited by a continuous curve $\mathcal{C}_g$ called the *tangential cut-locus* of g. Its image by the exponential map is called the *cutlocus* of g. One can prove that it is the closure of the set of points where several minimizing geodesics starting from g meet. We'll not delve into this subject here as it requires notions of conjugate points and Jacobi fields [10] which we don't introduce in this thesis.

Remark 2.21. However, these maximal domains (for each point g) are important for the problematic of this thesis. In fact, they are the maximal domains where the logarithm map exists and is unique and the definitions of the mean we'll give in the next chapter are formulated in terms of logarithms. Hence they will be valid only on the domain of bijectivity of the exponential map.

We can define the distance from a point g to its cutlocus $\mathcal{C}_g$ and the *injectivity radius* of the whole manifold.

Definition 2.22. (Injectivity radius) For each $g \in \mathcal{G}$, the *injectivity radius* of ($\mathcal{G}$,<,>) at g is:

$$\mathrm{inj}_g(\mathcal{G}, <, >) = \sup\{r : exp_g \text{is injective on } B_r(0) \subset T_g\mathcal{G}\}$$

and the injectivity radius of ($\mathcal{G}$,<,>) is:

$$\mathrm{inj}(\mathcal{G}, <, >) = \inf\left\{\mathrm{inj}_g(\mathcal{G}/\forall g \in \mathcal{G}\right\}$$

If $\mathcal{G}$ is compact, then $0 < \mathrm{inj}(\mathcal{G}, <, >) \leq \mathrm{diam}(\mathcal{G}, <, >)$. But in the general case, we may have $\mathrm{inj}(\mathcal{G}, <, >) = 0$ or $+\infty$.

Now that we have reviewed the geometric notions relatives to affine connection spaces and Riemannian manifold, we may add the algebraic structure specific of Lie groups.

2.3 Lie groups

First recall the definition:

Definition 2.23. A *Lie Group* $\mathcal{G}$ is a smooth manifold together with a compatible group structure. It is provided with an identity element e, a smooth composition law $*$ and a smooth inversion law Inv :

$$* : (g, h) \mapsto g * h \in \mathcal{G}$$

$$Inv : f \mapsto f^{(-1)} \in \mathcal{G}$$

Remark 2.24. In this thesis we'll deal only with connected, simply connected Lie groups. Motivation for this choice will be clear later.

The inverse map is a diffeomorphism of $\mathcal{G}$ seen as a manifold but also a group automorphism of $\mathcal{G}$, obviously seen as a group.

There is a canonical way to define two other diffeo/auto-morphisms families of $\mathcal{G}$, called *the left and right translations*: $L_g : f \mapsto g * f$ and $R_g : f \mapsto f * g$ in addition to a third one called *the conjugations* $C_g = f \mapsto g * f * g^{(-1)}$. One immediatly has the following relationships:

$$L_{f*g} = L_f \circ L_g \qquad R_{f*g} = R_g \circ R_f \qquad Inv \circ L_g = R_{g^{(-1)}} \circ Inv$$

$$L_g \circ R_h = R_h \circ L_g \qquad C_g = L_g \circ R_{g^{(-1)}} = R_{g^{(-1)}} \circ L_g$$

As they are smooth maps, one can differentiate them to get a linear map between two tangent spaces of the Lie group.

2.3.1 Lie group and Lie algebra

Noting that the differential $DL_g(h)$ at h of the left-translation of g, L_g, maps the tangent space $T_h\mathcal{G}$ to the tangent space $T_{L_g(h)}\mathcal{G} = T_{g*h}\mathcal{G}$, we can define the subset of left-invariant vector fields as:

Definition 2.25. (Left-invariant vector fields) A vector field $\tilde{X}$ is said to be *left-invariant* if for any $g \in \mathcal{G}$, we have:

$$\tilde{X}|_{g*h} = \tilde{X}|_{L_g(h)} = DL_g(h).\tilde{X}|_h \qquad \forall h \in \mathcal{G}$$

Hence we see that a left-invariant vector field is only determined by its value x at e and we can write in short: $\tilde{X} = DL.x$. Later, the implementation of left-invariant vector fields amounts to compute the differential of the left translations *at the identity e.*

The Lie bracket of two left-invariant vector fields is also left-invariant [9] and we denote $[x, y]$ its value at e. The left-invariant vector fields form a Lie sub-algebra of $\Gamma(\mathcal{G})$ which we identify with the tangent space at identity e provided with the inherited Lie Bracket: $\mathfrak{g} = (T_e\mathcal{G}, +, ., [., .])$.

Definition 2.26. (Lie algebra) $\mathfrak{g} = (T_e\mathcal{G}, +, ., [., .])$ is called the *Lie algebra of the group $\mathcal{G}$.*

The Lie algebra $\mathfrak{g}$ represents the structure of the Lie group $\mathcal{G}$ near its identity e.

Definition 2.27. (Structure constants) Take $(T_i)_i$ a basis of the Lie algebra $\mathfrak{g}$, taken as a vector space. Then, the constants c_{ijk} defined as:

$$[T_i, T_j] = c_{ijk}T_k$$

are called the *structure constants of the Lie algebra.*

The Lie algebra structure of the vector space $\mathfrak{g}$ is determined by these constants. As elements of the Lie algebra are left-invariant vector fields, it is desirable that these brackets don't depend on the point where we evaluate them. It is the case and that's precisely what is meant under "structure *constants*". Note that for semisimple groups, the c_{ijk} are totally antisymmetric in $[ijk]$.

Remark 2.28. By symmetry, one can also define the sub-algebra of right-invariant vector fields $\overline{X} = DR.x$ and identify it with the tangent space at e. Note that the right bracket will be the opposite of the left bracket. In fact, this special point can create misunderstandings as researches on finite dimensional Lie groups use the left bracket whereas researches on infinite dimensional Lie groups deal with the right bracket.

A *Lie group homomorphism* is a map between Lie groups respecting the Lie group structure i.e. is smooth and compatible with $*$. A *Lie algebra homomorphism* is a map between Lie algebras respecting the Lie algebra structure, i.e. is linear and compatible with [,]. One is usually more familiar with matrix groups than abstract ones. Hence, define the following Lie group and Lie algebra homomorphisms.

Definition 2.29. (Representations of $\mathcal{G}$ and $\mathfrak{g}$) A *real representation* η *of the Lie group* $\mathcal{G}$ on the real vector space V is a Lie group homomorphism:

$$\eta : \quad \mathcal{G} \longmapsto GL(V)$$

A real representation of the Lie algebra $\mathfrak{g}$ on the real vector space V is a Lie algebra homomorphism:

$$\theta : \quad \mathfrak{g} \longmapsto \mathfrak{gl}(V)$$

We will denote $\eta(g).v$ the action of the element $g \in \mathcal{G}$ on the vector $v \in V$ through the representation η, and similary $\theta(x).v$ the action of the element $x \in \mathfrak{g}$ on the vector $v \in V$ through the representation θ.

A representation η (resp. θ) of a Lie group (resp. a Lie algebra) is *faithful* if η (resp. θ) is injective. We state without proof the well-known proposition [11]:

Proposition 2.30. *The differential of a Lie group representation* η *at identity e provides a Lie algebra representation,* $\theta = d\eta|_e$*, of its Lie algebra.*

If $\mathcal{G}$ is simply connected, there is a bijection between representations η of $\mathcal{G}$ and representations $d\eta|_e$ of $\mathfrak{g}$. This is why we only consider simply connected Lie groups in this thesis.

2.3.2 Adjoint and co-adjoint representations

We go back to the auto/diffeo-morphism family of conjugations C_g, to define a special representation of $\mathcal{G}$. For $g \in \mathcal{G}$, the differential of C_g at e $DC_g(e)$ has the particularity of mapping $T_e\mathcal{G}$ to itself, hence being an automorphism of vector spaces in contrast to $DL_g(e)$ and $DR_g(e)$ which are isomorphisms of vector spaces, from $T_e\mathcal{G}$ to $T_g\mathcal{G}$. In fact, $DC_g(e)$ is more than an automorphism of vector spaces. It provides a representation of $\mathcal{G}$ on its Lie algebra i.e. take $V = \mathfrak{g}$ in the definition.

Definition 2.31. (Adjoint representation of $\mathcal{G}$) The following Lie group representation is called the *adjoint representation* of the Lie group $\mathcal{G}$:

$$\begin{aligned} Ad : \mathcal{G} &\longmapsto Aut(\mathfrak{g}) \subset GL(\mathfrak{g}) \\ g &\longmapsto Ad(g) = DC_g|_e = DL_g|_{g^{(-1)}}.DR_{g^{(-1)}}|_e = DR_{g^{(-1)}}|_g.DL_g|_e \end{aligned}$$

The subgroup $Ad(\mathcal{G})$ of $GL(\mathfrak{g})$ is called the *adjoint group.*

Ad is a Lie group homomorphism, i.e. it is smooth and compatible with $*$:

$$Ad(g_1 * g_2) = Ad(g_1).Ad(g_2).$$

For each $g \in \mathcal{G}$, $Ad(g)$ is a *Lie algebra* automorphism, i.e. $Ad(g)$ is linear, inversible with $Ad(g)^{-1} = Ad(g^{-1})$ and it respects the Lie bracket:

$$Ad(g).[x, y] = [Ad(g).x, Ad(g).y], \quad \forall x, y \in \mathfrak{g}.$$

Taking the derivative of the map Ad at e provides a representation of the Lie algebra, called the adjoint representation of the Lie algebra. However, we need to define the tangent space of $Aut(\mathfrak{g})$ at *its identity* Id, i.e. the Lie algebra of $Aut(\mathfrak{g})$. It corresponds to the derivations of $\mathfrak{g}$.

Definition 2.32. (Derivation of $\mathfrak{g}$) A *derivation of* $\mathfrak{g}$ is a linear mapping $D : \mathfrak{g} \mapsto \mathfrak{g}$ such that:

$$D[x,y] = [Dx,y] + [x,Dy] \qquad \forall x,y \in \mathfrak{g}$$

The set of derivations of $\mathfrak{g}$ is denoted $Der(\mathfrak{g})$.

Now, we have:

Proposition 2.33. *The derivations of* $\mathfrak{g}$ *represent the Lie algebra of Aut(*$\mathfrak{g}$*):*

$$Der(\mathfrak{g}) = T_{Id}Aut(\mathfrak{g}).$$

Proof. Take $D \in End(\mathfrak{g})$ be such that $exp(tD) \in Aut(\mathfrak{g})$ for $t \in \mathbb{R}$, where exp is the usual series using the product of endomorphisms $\circ$. The definition of a Lie algebra automorphism gives exp(tD)[x,y] $=$ [exp(tD)x,exp(tD)x], which we can differentiate to get:

$$D[x,y] = [Dx,y] + [x,Dy]$$

saying that D is a derivation of the Lie algebra. Conversely, if D is a derivation then, with have by induction:

$$D^n[x,y] = \sum_{k=0}^{n} \binom{n}{k} [D^k x, D^{n-k} y]$$

which shows that $exp(D) = \sum_{n=0}^{\infty} \frac{D^n}{n!}$ is an automorphism. Hence the Lie algebra of Aut($\mathfrak{g}$) is the Lie algebra Der($\mathfrak{g}$) of derivations of $\mathfrak{g}$. $\square$

This enables the following definition from the adjoint representation of the Lie group $\mathcal{G}$:

Definition 2.34. (Adjoint representation of $\mathfrak{g}$) The following Lie algebra representation is called the *adjoint representation of the Lie algebra* $\mathfrak{g}$:

$$\begin{aligned} ad = D(Ad)|_e : \mathfrak{g} &\longmapsto \mathrm{Der}(\mathfrak{g}) \\ x &\longmapsto ad(x) = [x,.] \end{aligned}$$

The subalgebra $ad(\mathfrak{g})$ is called the *inner derivations algebra.*

The particularity of the adjoint representation is that $\mathfrak{g}$ is represented on itself. It is at the same time the acting Lie algebra and the vector space it is acting on. The expression of ad as the Lie bracket will be clear in the next subsection. Let us emphasize the properties of the adjoint representation of the Lie algebra. First, ad is an Lie algebra homomorphism: it is linear and compatible with the Lie bracket:

$$ad([x,y]) = [ad(x), ad(y)] \quad \forall x,y \in \mathfrak{g},$$

which is exactly the Jacobi identity.
For $x \in \mathfrak{g}$, $ad(x)$ is a derivation, i.e. $ad(x)$ is linear and:

$$ad(x)[y,z] = [ad(x).y, z] + [y, ad(x).z] \quad \forall x, y, z \in \mathfrak{g}.$$

Provided with a definite inner product, we can present the subject of duality on $\mathfrak{g}$.

Definition 2.35. (Dual space) The *dual space* of $\mathfrak{g}$ (seen as a vector space), denoted $\mathfrak{g}^*$ is the vector space of all linear forms on $\mathfrak{g}$. More precisely, it can be defined with the inner product on $\mathfrak{g}$ using the canonical isomorphism (Riesz' theorem in finite dimension):

$$\begin{aligned}\phi : \mathfrak{g} &\longmapsto \mathfrak{g}^* \\ a &\longmapsto \phi(a) = < a, \bullet > \end{aligned}$$

Remark 2.36. As the dual space of a Lie algebra, $\mathfrak{g}^*$ inherits a specific algebraic structure: a Lie-Poisson structure. Take f_1, f_2 two linear forms on $\mathfrak{g}$ such that $f_1 = \phi(a_1), f_2 = \phi(a_2)$. The induced bracket:

$$\{f_1, f_2\} = < [a_1, a_2], \bullet >$$

defines a Lie-Poisson structure on $\mathfrak{g}^*$.

From the representation theory, we recall the following definitions.

Definition 2.37. (Co-adjoint representations) The *co-adjoint representation* of $\mathcal{G}$ is defined on $\mathfrak{g}^*$ by taking the adjoints (in the sense of duality) of the endomorphisms Ad(g):

$$\begin{aligned}Ad^* : \mathcal{G} &\longmapsto Aut(\mathfrak{g}^*) \\ g &\longmapsto Ad^*(g) \text{ such that: } \forall x \in \mathfrak{g} : < Ad(g).a, x > = < a, Ad(g)^*.x > \end{aligned}$$

We can also define the co-adjoint representation of $\mathfrak{g}$:

$$\begin{aligned}ad^* : \mathfrak{g} &\longmapsto Der(\mathfrak{g}^*) \\ x &\longmapsto ad^*(x) \text{ such that: } \forall y, z \in \mathfrak{g} : < ad(x).y, z > = < y, ad(x)^*.z > \end{aligned}$$

For $g \in \mathcal{G}$ or $x \in \mathfrak{g}$, $Ad(g)$ and $ad(x)$ have nice explicit expressions on $\mathfrak{g}$. We show here how Ad^* and ad^* can be computed in the finite dimensional case of matrices. As the adjoint (in the sense of duality) depends on the metric $<,>$ we choose, first recall the *Frobenius inner product* on matrices which we'll use several times in this thesis:

$$< X, Y > = \frac{1}{2} Tr(X^T.Y) \quad \forall X, Y \in \mathcal{M}_n(\mathbb{R}).$$

Note that it may be defined without the constant $\frac{1}{2}$ in some papers.

Proposition 2.38. *For $G \in GL(n)$, $X \in \mathfrak{gl}(n)$, and the Frobenius inner product on $\mathfrak{gl}(n)$, we have:*

$$Ad^*(G) = Ad(G^T) \quad \text{and} \quad ad^*(X) = ad(X^T)$$

Proof. Take $G \in GL(n)$, $X, Y \in \mathfrak{gl}(n)$, we have:

$$\begin{aligned} < Ad(G).X, Y > &=< G.X.G^{-1}, Y > \quad \text{(Ad in the matrix case)} \\ &= \frac{1}{2}.Tr((G.X.G^{-1})^T.Y) \quad \text{(Frobenius inner product)} \\ &= \frac{1}{2}.Tr(G^{-T}.X^T.G^T.Y) \\ &= \frac{1}{2}.Tr(X^T.G^T.Y.G^{-T}) \quad \text{(circular permutation in Tr)} \\ &=< X, Ad(G^T).Y > \end{aligned}$$

As for ad, we take $X, Y, Z \in \mathfrak{gl}(n)$ and we get:

$$\begin{aligned} < ad(X).Y, Z > &=< [X, Y], Z > \quad \text{(definition of ad)} \\ &=< X.Y - Y.X, Z > \quad \text{(matrices: the bracket is the commutator)} \\ &= \frac{1}{2}.Tr((X.Y - Y.X)^T.Z) \quad \text{(Frobenius inner product)} \\ &= \frac{1}{2}.Tr(Y^T.X^T.Z - X^T.Y^T.Z) \\ &= \frac{1}{2}.Tr(Y^T.X^T.Z - Y^T.Z.X^T) \quad \text{(circular permutation in Tr)} \\ &=< Y, [X^T, Z] > \\ &=< Y, ad(X^T).Z > \end{aligned}$$

Hence we get both results. □

Now we have set the general algebraic framework on a Lie group without having defined any connection or metric. In the next section, we show how to define a generalization of straight lines in this context.

2.3.3 Lie group exponential and logarithm

A Lie group is non-linear, while its Lie algebra $\mathfrak{g}$ is: it is the best linear approximation of $\mathcal{G}$ at e. As in the connection and Riemannian spaces, we precise this relation by defining an exponential map, the *Lie group exponential*, and first we agree on which curves on $\mathcal{G}$ should approximate the straight lines of $\mathfrak{g}$. For the group exponential, they will be the integral curves of left-invariant vector fields.

Generally speaking, for any smooth vector field X we have the following smooth ordinary differential equation:

$$\frac{dx(t)}{dt} = X|_{x(t)} \quad (*)$$

Associated with an initial condition of the type $x(0) = p_0$, the theory ODE tells us that $(*)$ has an unique solution for some interval about t=0. This solution is called the *integral curve* of the vector field X starting at p_0. We denote it $\sigma_X(t, p_0)$. Taking all curves starting at all points in

$\mathcal{G}$, we may think of σ_X as the map:

$$\begin{aligned} \sigma_X : \quad & \mathbb{R} \times \mathcal{G} \longmapsto \mathcal{G} \\ & (t, p_0) \longmapsto p = \sigma_X(t, p_0) \end{aligned}$$

We call this map the *flow* defined by the vector field X. Note that $\sigma_X(\bullet, p_0)$ might not be defined for all $t \in \mathbb{R}$. However, it is defined everywhere for left-invariant vector fields $\tilde{X}$ as they don't have zeros (DL is an isomorphism).

Integral curves belong to the *group* $\mathcal{G}$. Could an integral curve define a (one-dimensional) subgroup of $\mathcal{G}$?

Definition 2.39. A *one-parameter subgroup of* $\mathcal{G}$ is a curve on $\mathcal{G}$ defining a Lie group homomorphism from $(\mathbb{R}, +)$ to $(\mathcal{G}, *)$, i.e.:

$$\begin{aligned} \gamma : (\mathbb{R}, +) & \longmapsto (\mathcal{G}, *) \\ t & \longmapsto \gamma(t) \quad \text{such that: } \forall s, t \in \mathbb{R}, \gamma(s+t) = \gamma(s) * \gamma(t) \end{aligned}$$

A one-parameter subgroup is abelian.

The following proposition shows that there is a surjection from the Lie algebra $\mathfrak{g}$ to the set of one-parameter subgroups.

Proposition 2.40. *The integral curve going through e of a left-invariant vector field is a one-parameter subgroup.*

Proof. The flow $\gamma_x(t)$ of a left-invariant vector field $\tilde{X} = DL.x$ starting from e exists for all times. Its tangent vector is $\gamma_x(t) = DL_{\gamma_x(t)}.x$ by definition of the flow. If we fix s, we observe that the two curves $\gamma_x(s+t)$ and $\gamma_x(s) * \gamma_x(t)$ are going through point $\gamma_x(s)$ with the same tangent vector. By uniqueness of the flow, they are the same. □

Hence, one-parameter subgroups are a possible approximation of the straight lines of $\mathfrak{g}$. Note that they start at e. Hence we translate them in order to define a geodesic starting at any point of the Lie group.

Definition 2.41. (Group geodesic) The *group geodesics* are the left translations of one-parameter subgroups.

We have now three sets of geodesics: the Riemannian geodesics, the connection geodesics and the group geodesics. We have seen the condition in order to have correspondence for the two first sets in the previous section: the connection ∇ of the affine connection space has to equal the Levi-Civita connection of the metric (up to a torsion term). If, in addition, we want to match the group geodesics, we have to require that ∇ is a Cartan-Schouten connection. Details will come in the next chapter.

Note that we always have the geodesic completeness in the case of the Lie group. Hence we can define the exponential map on the whole tangent space as:

Definition 2.42. Let $\gamma_x(t)$ be the flow of a left-invariant vector field $\tilde{X} = DL.x$ starting from e. The *group exponential of x at e*, Exp(x) is defined as:

$$Exp(x) = \gamma_x(1).$$

At $g \in \mathcal{G}$, we define the group exponential of u using the corresponding left translation of the one-parameter subgroup:

$$Exp_g(u) = g * Exp(DL_{g^{-1}}.u).$$

Remark 2.43. We have an explicit formulation for the adjoint represention of $\mathcal{G}$:

$$Ad(g).y = \frac{d}{dt}\Big|_{t=0} g.Exp(ty).g^{-1}$$

and in the matrix case, the simpler formula: $Ad(R).M = R.M.R^{(-1)}$.
Moreover, the expression of $ad(x)$ as the Lie bracket is given by computations from:

$$ad(x).y = \frac{d}{ds}\Big|_{s=0} Ad(exp(sx)).y.$$

using the fact that the adjoint representation of the Lie algebra is the differential at e of the adjoint representation of the Lie group [10].

Provided with the exponential map, we can show that the set of left translations of the one-parameter subgroups equal the set of their right-translations. [7]:

Theorem 2.44. *For $x \in \mathfrak{g}$ and $g \in \mathcal{G}$, we have:*

$$g * Exp(x) = Exp(Ad(g).x) * g.$$

The differential map of the group exponential at 0 is the Identity and we have again a local theorem:

Theorem 2.45. *The group exponential is a diffeomorphism from an open neighborhood of 0 in $\mathfrak{g}$ to an open neighborhood of e in $\mathcal{G}$.*

Now, try to have a look on the maximal bijectivity domain of the *group* exponential map. We first consider the case of *matrix* Lie groups where the group exponential is precisely the matrix exponential. In general, the logarithm of a matrix may not exist or may not be unique even if it exists. Following [12], we see that we generally need two group exponential to reach an element of the Lie group. But, if a (real) invertible matrix has no (complex) eigenvalues on the closed half line of negative real numbers, then it has a unique real logarithm, which is called the *principal logarithm*.

Can we generalize this to other Lie groups? The following development is a first step in that direction: we characterize a domain that should *contain* the maximal bijectivity domain of the group exponential map.

As $Exp : \mathfrak{g} \longmapsto \mathcal{G}$ is smooth, we can differentiate it at $x \in \mathfrak{g}$ to get the linear map:

$$DExp(x) : \mathfrak{g} \longmapsto T_{Exp(x)}\mathcal{G}.$$

We define the function f:

$$\begin{aligned} f : \mathfrak{g} &\longmapsto \mathbb{R} \\ x &\longmapsto \det(\mathrm{DExp}(x)) \end{aligned}$$

we have the following definition:

Definition 2.46. We call the *conjugate locus of* e the subset C_e of $\mathcal{G}$ defined as follow:

$$C_e = f^{(-1)}(\{0\})$$

and we call Q_e the connected component of e in $f^{(-1)}(]0, +\infty[)$.

Our conjecture is the following:

Proposition 2.47. *$Q_e \subset \mathfrak{g}$ contains the maximal domain U_e for the bijectivity of Exp. Moreover, it is star-shaped at 0 and invariant under the adjoint group.*

Proof. Let us prove part of this conjecture. Take x in the open neighbordhood of 0 given by Theorem 2.45. Let ρ^* be the first positive time t for which $f(tx) = 0$, i.e. the first point in direction x where DExp is not full rank. Note that for $\alpha \in \mathbb{R}^*$, $\rho^*(\alpha.x) = \frac{\rho^*(x)}{\alpha}$. Hence, Q_e is the domain delimited by the points $\rho^*(\alpha x).\alpha x = \rho^*(x).x$. It is obviously star-shaped at 0.

Prove that Q_e is invariant under the adjoint group. Take $x \in Q_e$, such that $f(x) > 0$. We show that $f(Ad(g).x) > 0$, for all $g \in \mathcal{G}$. Take $g \in \mathcal{G}, y \in \mathfrak{g}$, we have:

$$DExp(Ad(g).x).y = \frac{d}{dt}|_{t=0} Exp(Ad(g).x + ty)$$

by definition of the differential, and recalling :

$$DExp(Ad(g).x) : \mathfrak{g} \longmapsto T_{Exp(Ad(g).x)}\mathcal{G}.$$

Now Ad(g) is an automorphism of $\mathfrak{g}$, so we can take y' the unique antecedent of y by Ad(g), such that : y=Ad(g).y'. Hence:

$$\begin{aligned} DExp(Ad(g).x).y &= \frac{d}{dt}\Big|_{t=0} Exp(Ad(g).x + tAd(g).y') \\ &= \frac{d}{dt}\Big|_{t=0} Exp(Ad(g).(x + ty')) \quad \text{(automorphism Ad(g) is linear)} \\ &= \frac{d}{dt}\Big|_{t=0} g * Exp(x + ty') * g^{(-1)} \quad \text{(see Theorem 2.44)} \\ &= \frac{d}{dt}\Big|_{t=0} C_g \circ Exp(x + ty') \end{aligned}$$

Now using the differentiate of the composition $D(f \circ g)(a) = Df(g(a)) \circ Dg(a)$, we get:

$$\begin{aligned} DExp(Ad(g).x).y &= DC_g(Exp(x)).DExp(x).y' \\ &= DC_g(Exp(x)).DExp(x).Ad(g^{-1}).y \end{aligned}$$

because $Ad(g)^{-1} = Ad(g^{-1})$. Let emphasize the nature of the objects on the last line:

- $Ad(g)$ is an automorphism from $\mathfrak{g}$ to $\mathfrak{g}$,
- $DExp(x)$ is an homomorphism from $\mathfrak{g}$ to $T_{Exp(x)}\mathcal{G}$,
- $DC_g(Exp(x))$ is an isomorphism from $T_{Exp(x)}\mathcal{G}$ to $T_{C_g(Exp(x))}\mathcal{G}$.

Hence, considering:

$$\det(DExp(Ad(g).x)) = \det(DC_g(Exp(x))).\det(DExp(x)).\det(Ad(g^{-1}))$$

it is clear that : $f(x) > 0 \Rightarrow f(Ad(g).x) > 0$ and we get the result. □

In order to get an explicit expression of U_e for a given Lie group, a first step is to compute Q_e and hence $DExp(x)$.

Theorem 2.48. *Let g be the analytic function:*

$$g(z) = \frac{1 - \exp(-z)}{z} = \sum_{m=0}^{+\infty} \frac{(-1)^m}{(m+1)!} z^m.$$

where exp is the standard exponential function. We have:

$$DExp(X).Y = DL_{Exp(X)}.g(adX).Y$$

Proof. We prove it in the matrix case, since we are dealing with matrix Lie groups in this thesis. First translate the (tangent) vector attached at $Exp(X)$ to e, in order to calculate the Lie Algebra element:

$$\begin{aligned} DL_{Exp(-X)}.DExp(X).Y &= DL_{Exp(-X)}.\left(\frac{d}{dt}\right)|_{t=0} Exp(X + tY) \\ &= DL_{Exp(-X)}.\sum_{n\geq 1}\left(\frac{d}{dt}\right)_{t=0} \frac{(X+tY)^n}{n!} \\ &= \sum_{r\geq 0, n\geq 0, 0\leq k\leq n} \frac{(-1)^r}{r!(n+1)!} X^r X^k Y X^{n-k} \end{aligned}$$

using the equivalent of Newton polynomial for non-commutative elements.
Changing the summation variables from r, n, k to m, r, k where $m = r + n$, we get:

$$\begin{aligned}
DL_{Exp(-X)}.DExp(X).Y &= \sum_{r\geq 0, m\geq 0, r+k\leq m} \frac{(-1)^r}{r!(m-r+1)!} X^{r+k} Y X^{m-r-k} \\
&= \sum_{r\geq 0, m\geq 0, r+k\leq m} \frac{(-1)^m}{r!(m+1)!} (-1)^{m-p} \binom{m+1}{r} X^p Y X^{m-p}
\end{aligned}$$

Using the following identity:

$$\sum_{r=0}^{p} (-1)^r \binom{m+1}{r} = (-1)^p \binom{m}{p}$$

and the fact that $\tilde{adX} = L_{\tilde{X}} - R_{\tilde{X}}$, so that:

$$(adX)^m.Y = \sum_{p=0}^{m} (-1)^{m-p} \binom{m}{k} X^k Y X^{m-k}.$$

we have finally:

$$\begin{aligned}
DL_{Exp(-X)}.DExp(X).Y &= \sum_{m=0}^{+\infty} \frac{(-1)^m}{(m+1)!} (adX)^m.Y \\
&= g(adX).Y
\end{aligned}$$

which ends the proof of the theorem. □

Finding Q_e amounts to solve $det(g(ad(x))) = 0$ as $DL_{Exp(x)}$ is an isomorphism. Let give some insights about how we could proceed.

If x commutes with all elements of $\mathfrak{g}$, then $\rho^*(x) = \infty$. Hence, if $\mathcal{G}$ is abelian $g(ad(x)) = Id$ for all $x \in \mathfrak{g}$ and: $Q_e = \mathfrak{g}$. Moreover, if x is nilpotent, then $\rho^*(x) = \infty$. In the case of a general x, we use the *abstract* Jordan decomposition [13] on $ad(x)$, writing it as a special sum of a nilpotent and a semisimple endomorphism:

$$ad(x) = d(x) + n(x).$$

Then we find the (unique) Jordan decomposition of $g(ad(x))$:

$$g(ad(x)) = D(x) + N(x),$$

where we have (from the computations):

$$D(x) = g(d(x)).$$

This enables to find the eventual bijectivity of $g(ad(x))$: we consider the eigenvalues $\Lambda(x)_i$ of the semisimple part $D(x)$ as the determinant is the product of these eigenvalues taken with multiplicity (by property of the Jordan decomposition). We have :

$$\Lambda(x)_i = g(\lambda(x)_i)$$

where $\lambda(x)_i$ are the eigenvalues of $d(x)$. Hence, we have to find the zeros z of the function g. They are given by:

$$z \equiv 0[2i\pi] \quad \text{with} \quad z \neq 0.$$

Hence we could conclude about the domain Q_e by an analyze of the adjoint representation.

Now define the group logarithm from $\mathcal{U}_e = Exp(U_e)$ to U_e, where we only have an inclusion $U_e \subset Q_e$ (which can be strict, for example in the case of SE(2)).

Definition 2.49. Group logarithm) For every $h \in \mathcal{U}_e$, we define the *group logarithme of h*, Log(h) as the unique $x \in U_e$ such that $Exp(x) = h$. We write $x = Log(h)$. The group logarithm at $g \in \mathcal{G}$ is defined using the left translation of the one-parameter subgroups as:

$$Log_g(h) = DL_g.Log(g^{-1} * h).$$

Note that we also have a relation of the group logarithm with the adjoint representation [7]. For h, g in the neighborhoods where the following makes sense, we have:

$$Log(g * h * g^{-1}) = Ad(g).Log(h).$$

Baker-Campbell-Hausdorff (BCH) formula The implementation of Lie group exponential and logarithm is the starting point of any programming on Lie groups. The formula is useful to speed up computations. By definition, we have:

$$Exp(x).Exp(y) = Exp(BCH(x, y))$$

Intuitively, BCH(x,y) shows how $Log(Exp(x).Exp(y))$ deviates from $x + y$ because of the non-commutativity of $\mathcal{G}$.

The expression of the BCH is given by the following theorem.

Theorem 2.50. *Take x, y small enough in $\mathfrak{g}$.*
Then the $Log(Exp(x).Exp(y))$ is analytic around 0 and we have the following development, called the BCH-formula:

$$\begin{aligned} BCH(x, y) &= Log(Exp(x).Exp(y)) \\ &= x + y + \frac{1}{2}[x, y] + \frac{1}{12}([x, [x, y]] + [y, [y, x]]) \\ &+ \frac{1}{24}[[x, [x, y]], y] + O((||x|| + ||y||)^5) \end{aligned}$$

Remark 2.51. The BCH-formula can be rewritten in the form of the *Baker-Campbell-Hausdorff-Dynkin formula*:

$$BCH(x, y) = x + \int_0^1 F(Ad_x Ad_{ty})ydt \tag{2.1}$$

for all sufficiently small x, y where $Ad_x = exp(ad_x)$, and F is the function:

$$F(t) = \frac{tlogt}{t-1}$$

which is analytic near $t = 1$ and thus can be applied to linear operators sufficiently close to the identity.

A crucial point of Theorem 2.50 is that BCH is not only $\mathcal{C}^\infty$ but also *analytic around 0.* It means that *BCH(x,y)* near 0 is the sum of an absolute converging multivariate infinite series, the usual multiplication being replaced by the Lie Bracket.

In the case of matrices, it is obvious that each term of:

$$BCH(tX, tY) = \sum_{n=0}^{+\infty} t^n.Z_n(X, Y)$$

is a homogeneous polynomial map of $\mathfrak{g}$ x $\mathfrak{g}$ into $\mathfrak{g}$ of degree n. But it is not clear why it is a *Lie polynomial map*, i.e. built from commutators of x and y. This is the statement of the BCH formula. Before proving it, we state this corollary of 2.48, which is immediate using the differential of the composition for: $t \longmapsto Exp(F(t))$.

Corollary 2.52. *If F is a map $\mathbb{R} \longmapsto \mathfrak{g}$ then:*

$$\frac{d}{dt} Exp(F(t)) = Exp(F(t)).g(adF(t))(F'(t))$$

Now we show the guidelines for the proof of Theorem 2.50, which is mostly computational.

Proof. (Proof of Theorem 2.50) In order to write the Z_n as Lie polynomials, we show an ordinary non-linear differential equation which is verified by $BCH(tX, tY)$. The Z_n will be computed inductively.

We write, for $u, v \in \mathbb{R}, X, Y \in \mathfrak{g}$:

$$Exp(uX).Exp(vY) = Exp(Z(u, v, X, Y))$$

where Z is analytic *in u, v* near $(0, 0)$ for each pair (X,Y). Moreover, Z vanishes at $u = v = 0$.

Differentiating with respect to v using Corollary 2.52, we find:

$$Exp(uX).Exp(vY).Y = Exp(Z).g(adZ).\left(\frac{\partial Z}{\partial v}\right)$$

which is, by definition of the BCH:

$$Y = g(adZ).\left(\frac{\partial Z}{\partial v}\right)$$

Inverting $g(adZ)$ will provide the needed differential equation. The last computations are technical and express a specific recursive relation between the $Z_n(X, Y)$, proving that they are indeed Lie polynomials. □

For the sake of completeness, we give (without proof) the explicit formula on the Lie polynomials $Z_n(X,Y)$, stated in the following theorem.

Theorem 2.53 (Dynkin's Formula)**.** *For $BCH(tX,tY) = \sum_{n=0}^{+\infty} t^n.Z_n(X,Y)$, we have:*

$$z_n(X,Y) = \frac{1}{n}\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}\sum\frac{[X^{p_1}Y^{q_1}...X^{p_k}Y^{q_k}]}{p_1!q_1!...p_k!q_k!},$$

with internal summation on all non-negative exponents $p_1, q_1, ...p_k, q_k$ for which:

$$\begin{gathered} p_1+q_1>0, ..., p_k+q_k>0 \\ p_1+q_1+...+p_k+q_k=n \end{gathered}$$

and:

$$\begin{aligned} &[X^{p_1}Y^{q_1}...X^{p_k}Y^{q_k}] = \\ &\qquad [...\underbrace{[X,X],X,...,X]}_{p_1},\underbrace{Y],...,Y]}_{q_1},...,\underbrace{X],...X]}_{p_k},\underbrace{Y],...,Y]}_{q_k}. \end{aligned}$$

A proof can be found in [10].

Chapter 3

Defining the mean on a Lie group

Lie groups are non-linear spaces and hence the usual definition of the mean as a weighted sum or an integral can not be used anymore. Several alternatives have been proposed in the literature, defining means in different ways. On manifolds, the most widely used is the Riemannian center of mass. In this chapter, we propose two extensions of this definition: the Riemannian exponential barycenter and the group exponential barycenter. In order for the extended definition to be valid, it has to verify two properties. First of all, if we restrict it to the case where the Riemannian CoM is also defined, it is obvious that both definitions have to correspond. Secondly, as we are generalizing to Lie groups, we want the mean to be consistent with the group operations, composition and inversion. If we transform the data set, we want the mean to transform the same way. We call a *admissible mean* a extended definition of the mean verifying these two properties. This chapter discuss the admissible means.

3.1 Extension of the mean's definition

Let $\mathcal{G}$ be a Lie group, and $\{x_i\}_i$ a data set of points in $\mathcal{G}$. In the following, we present different definition's extension of the Riemannian CoM of $\{x_i\}_i$ on the Lie group $\mathcal{G}$.

In order to define a mean on a Lie group, which is still a locally linear structure, one could think about unfolding the manifold. We could compute linear statistics on the vector space one gets., for example perform the usual mean, and then fold the space again to find the mean in the manifold. This is precisely the idea of the Log-Euclidean mean proposed in [14].

Definition 3.1. (Log-Euclidean Mean) The *Log-Euclidean mean* $\overline{x}_{LE}$ of $\{x_i\}_i$ is defined as followed:

$$\overline{x}_{LE} = Exp(\sum_{i=1}^{n} w_i Log(x_i))$$

Despite its intuitive formulation, the Log-Euclidean mean is not admissible: it is not consistent with the left and right translations when the Lie group is not abelian. Indeed, the BCH formula means that, in the general case, we have:

$$Exp(\sum_{i=1}^{n} w_i Log(g * x_i)) \neq g * Exp(\sum_{i=1}^{n} w_i Log(x_i))$$

for the left translation for instance. Hence, we directly reject the Log-Euclidean mean in this thesis.

We present two types of mean's definition: the first type depends on the pseudo-metric we defined on the Lie group, the second one is algebraic. For example, the Log-Euclidean mean would have corresponded to the second type.

3.1.1 Relying on the pseudo-Riemannian structure

The computing framework on Riemannian manifold is well developed for finite-dimensional manifolds, for example matrix Lie groups. Hence, it appears convenient to rely on this setting to define a mean. [15],[16] proposes a definition generalizing the fact that the usual esperance, i.e. in linear statistics, realizes the global minimum of the usual variance, i.e.:

$$\sigma_x^2(y) = \mathbf{E}[||\mathbf{x}, y||^2]$$

where $||.||$ is usually the Euclidean norm on the finite dimensional vector space $\mathbb{R}^n$. Replacing the norm on $\mathbb{R}^n$ by the geodesic distance of ($\mathcal{G}$, $<,>$), we define a notion of mean on ($\mathcal{G}$,$<,>$).

In the non-linear case, we might have different minima of the variance, if we have some. But if one considers *local minima* of the variance, [17] and [18] were able to prove, under some conditions, existence and uniqueness theorems (see last chapter). This is the motivation for the definition of the Riemannian centers of mass.

Definition 3.2. (Riemannian centers of mass (CoMs)) Let $\mathcal{G}$ be a Lie group provided wih a Riemannian metric $<,>$ which induces a distance dist. The *Riemannian centers of mass* of $\{x_i\}_i$, if they are some, are the points realizing the local minima of the variance:

$$\underset{y \in \mathcal{G}}{\operatorname{argmin}}\, \sigma_x^2(y) = \mathbf{E}[dist(\mathbf{x}, y)^2]$$

This is not an admissible definition for all Lie groups. The consistency of the Riemannian CoM with the group operations depends on the properties of the metric $<,>$ inducing the geodesic distance used for the variance. In order to have a mean consistent with left and right translations, we need the metric to be bi-invariant [7]. However, such a metric doesn't exist in all Lie groups as we'll see in the last section of this chapter.

Enlarging the set of admissible means, we could rely on the zeros of the variance's gradient, including maxima, minima and saddle points. This is the essence of the definition of *pseudo-Riemannian exponential barycenters* proposed by [19]. It is inspired from the affine space implicit definition of the barycenter : $\sum_i \vec{GM_i} = \vec{0}$. Note that it is now possible to use a pseudo-Riemannian metric.

Definition 3.3. (Pseudo-Riemannian exponential barycenters) The points m which are solutions of the following *pseudo-Riemannian barycenter equation*, if there are some, are called *pseudo-Riemannian exponential barycenters* of $\{x_i\}_i$:

$$\sum_i w_i.log_m(x_i) = 0 \quad \text{or, in the continuous case:} \quad \int_{\mathcal{G}} log_m(x)d\mu(x) = 0 \tag{3.1}$$

Intuitively, if we unfold the manifold at the Riemannian exponential barycenter along the Riemannian geodesics and compute the (linear) mean of the data set in the vector space, we find 0.

This is illustrated in Figure 3.1.

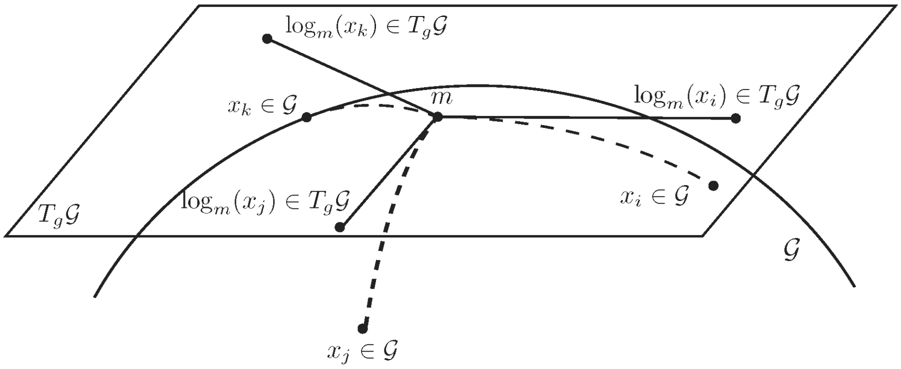

FIGURE 3.1: Definition of an exponential barycenter, here in the Riemannian case.

The previous barycentric equation corresponds precisely to the nullity of the variance's gradient and hence to the maxima, minima and saddle points we wanted, as stated in the following theorem from [6].

Theorem 3.4. *(Gradient of σ^2) Let $\mathcal{G}$ be Lie Group and μ a probability measure on $\mathcal{G}$. The variance $\sigma^2(y) = \int_{\mathcal{G}} dist(y,z)^2 d\mu(z)$ is differentiable (if it is finite) at the points y where the cut locus has a null probability measure $\mu(C(y)) = 0$. At such a point:*

$$(grad\sigma^2)(y) = -2\int_{\mathcal{G}} log_y(z)d\mu(z) \qquad if\mu(C(y)) = 0$$

Remark 3.5.

- Note that the pseudo-Riemannian exponential barycenters are either *critical points of the variance* (zeros of the gradient) or points with $\mu(C(y)) > 0$ (gradient is not defined).
- In the case of a Riemannian manifold, Riemannian CoMs form a subset of Riemannian exponential barycenters.

Once again, this definition is admissible only if <,> is bi-invariant on the Lie group. Take for example the left-translation by h of the data set x_i. Consistency of the mean m is equivalent with the following equality for any x_i (as we can choose the weights as we want):

$$log_m(x_i) = log_{h*m}(h * x_i) = v$$

As log_m and exp_m are inverse from each other, it corresponds to:

$$exp_{h*m}(v) = h * x_i \quad \text{when:} \quad exp_m(v) = x_i.$$

This is equivalent with the fact that the left-translation of a Riemannian geodesic has to be a Riemannian geodesic with same length's unit, i.e. the left-invariance of the pseudo-metric. Similarly, consistency with the right translation demands the right-invariance of the pseudo-metric. As for inversion, it is automatically verified when we have bi-invariance (see later).

Now, we may also use a pseudo-metric, not necessary a metric, in order to define a mean. We have extended the family of Lie groups we can provide with an admissible mean: those which can be provided with a bi-invariant pseudo-Riemannian metric. The characterization of Lie groups with bi-invariant pseudo-metric is the topic of Chapter 4.

3.1.2 Relying on the algebraic structure

However, the previous definition of the Riemannian exponential barycenter still reduces the Lie groups we can work on. To get free of the pseudo-metric constraint, [7] proposed a definition relying entirely on the algebraic properties of the Lie group. This mean is called the *group exponential barycenter* and corresponds formally to the (pseudo-)Riemannian exponential barycenter, but uses the group exponential map instead of the (pseudo-)Riemannian one.

Definition 3.6. (Group exponential barycenters) The points m which are solutions of the following *group barycenter equation*, if there are some, are called *group exponential barycenters* of the data set x_i.

$$\sum_i w_i.Log(m^{(-1)} * x_i) = 0 \tag{3.2}$$

In contrast to the means depending on the (pseudo-)Riemannian structure, the group exponential barycenter is naturally consistent with the group structure as stated in the following theorem:

Theorem 3.7. *(Bi- and inverse-invariance) The group exponential barycenters are left-, right- and inverse- invariant: if m is a group exponential barycenter of x_i and h an element of $\mathcal{G}$, then $h * m$ is a mean of $h * x_i$, $m * h$ is a mean of $x_i * h$ and $m^{(-1)}$ is a mean of $x_i^{(-1)}$.*

Hence, it is naturally an admissible definition as long as we can prove its existence and uniqueness. This will be treated in Section 5.2.

Proof. Take m a group exponential barycenter of x_i and h an element of $\mathcal{G}$. Then,

$$Log((h * m)^{(-1)} * h * x_i) = Log(m^{(-1)} * x_i)$$

is obviously well-defined for all x_i and $h * m$ is a solution of the group barycenter equation:

$$\sum_i w_i.Log((h * m)^{(-1)} * h * x_i) = 0 \tag{3.3}$$

which proves that $h * m$ is a group exponential barycenter of $h * x_i$, i.e. the consistency of the definition with the left translation.

In order to show the consistency with the right translation, we apply Theorem 2.44:

$$\mathrm{Ad}(m).\left(\sum_i w_i.Log(m^{-1} * x_i)\right) = \sum_i w_i.Log(x_i * m^{-1})$$

Since Ad(m) is an automorphism it is invertible, and hence the usual left-invariant barycentric equation is equivalent to a right-invariant (same arguments as above) barycentric equation. The definition is also consistent with the right translation.

To prove the consistency with the inversion, we consider:

$$(-1) \times \left(\sum_i w_i.Log(m^{-1} * x_i)\right) = \sum_i w_i.Log(x_i^{-1} * (m^{-1})^{-1})$$

using the fact that $Log(x^{-1}) = -Log(x)$. This shows that h^{-1} is a group exponential barycenter of x_i^{-1}. □

Hence we have defined the pseudo-Riemannian exponential barycenter and the group exponential barycenter in order to generalize the notion of Riemannian center of mass for as many Lie groups as possible. However, such a generalization is pertinent if the definitions give the same results in the cases where they can be both defined. This was clear for the Riemannian exponential barycenter. What about the group exponential barycenter? As the definitions rely on the exponential map (more precisely the logarithm map, its inverse), we investigate the conditions under those the Riemannian geodesics correspond to the group geodesics. This will be the case if the Levi-Civita connection associated to the metric is a *Cartan-Schouten connection*.

3.2 Metric, connection and group geodesics

First precise the definitions of left-invariant connection and metric.

Definition 3.8. (Left-invariant connection) A *left-invariant connection* is a connection ∇ such that:

$$\nabla_{DL_g X} DL_g Y = DL_g.\nabla_X Y \qquad \forall X, Y \in \Gamma(\mathcal{G}) \text{ and } g \in \mathcal{G}.$$

As a connection is completely determined by its action on the sub-algebra of left-invariant vector fields (obvious using the two axioms of an affine connection), we can restrict to this sub-algebra. Hence a left-invariant connection is only determined by its value at e and we can define a bilinear operator $\alpha(x, y)$ as:

$$\nabla_{\tilde{X}}\tilde{Y} = DL.\nabla_{\tilde{X}}\tilde{Y}\big|_e = DL.\alpha(x,y) \qquad \text{where } x = \tilde{X}\big|_e \text{ and } y = \tilde{Y}\big|_e$$

Conversely, any bilinear operator of $\mathfrak{g}$ uniquely defines a connection at e, thus a left-invariant connection on all left-invariant vector fields, and thus on all vector fields. Left-invariant connections are in bijection with bilinear operators on $\mathfrak{g}$. Moreover, we can theoretically provide any Lie group with a left-invariant connection but also computationaly as soon as we compute the jacobian of the left translation: $\nabla_{\tilde{X}}\tilde{Y} = DL.\alpha(x, y)$.

Definition 3.9. (Left-invariant pseudo-metric) A *left-invariant pseudo-metric* is a pseudo-metric $<,>$ such that:

$$< DL_h.X, DL_h.Y > |_{h*g} =< X, Y > |_g \qquad \forall X, Y \in T_g\mathcal{G}, g, h \in \mathcal{G}.$$

In other words, a left-invariant pseudo-metric is a pseudo-metric for which all left translations are isometries.

As for a left-invariant connection, a left-invariant pseudo-metric is only determined by the inner product at e. Left-invariant pseudo-metrics are hence in bijection with bilinear symmetric definite operators on $\mathfrak{g}$.

Note that we can define right-invariant connections (resp. pseudo-metrics) in the same way, and bi-invariant connections (resp. pseudo-metrics) are those with both left and right invariance.

Proposition 3.10. *(Inversion of a left-invariant metric) Let $<,>$ be a left-invariant pseudo-metric on $\mathcal{G}$. Then the inverted metric $<<,>>$ defined as:*

$$<< v, w >>_g =< DInv|_g.v, DInv|_g.w >_{g^{-1}} \qquad \forall v, w \in T_g\mathcal{G}, g \in \mathcal{G}$$

is right-invariant with $<,>_e=<<,>>_e$.

Proof. Differentiating the well-known equality $(h * g)^{-1} = g^{-1} * h^{-1}$ we get:

$$DInv|_{h*g} \circ DL_h|_g = DR_{h^{-1}}|_{g^{-1}} \circ DInv|_g,$$

which shows the right-invariance of $<<,>>$ from the left-invariance of $<,>$.
Then, $DInv|_e = -Id$ implies the equality $<,>_e=<<,>>_e$. □

3.2.1 Cartan-Schouten connections

Cartan-Schouten connections are precisely defined with the property we need in order to have consistent definitions for the mean.

Definition 3.11. (Cartan-Schouten connections) Among the left-invariant connection, the *Cartan-Schouten connections* are the ones for which geodesics going through identity are one-parameter subgroups.

Remark 3.12. Note that we could have defined the group geodesics as the geodesics of the Cartan-Schouten connections.

Theorem 3.13. *(Characterization of Cartan-Schouten connections) Among left-invariant connections, Cartan-Schouten connections are characterized by the condition:*

$$\alpha(x, x) = 0 \qquad \forall x \in \mathfrak{g}. \tag{3.4}$$

Proof. Consider the one-parameter subgroup $\gamma_x(t)$ starting from e with tangent vector $x \in \mathfrak{g}$. One-parameter subgroups are integral curves of left-invariant vector fields, so the tangent vector field is: $\dot{\gamma(t)} = DL_{\gamma(t)}.x$. Now, γ is a connection geodesic if it is autoparallel, i.e. if:

$$\nabla_{\dot{\gamma(t)}} \dot{\gamma(t)} = 0$$

By left-invariance of ∇, this is equivalent to:

$$\alpha(x, x) = 0.$$

□

Hence, Cartan-Schouten connections are the one defined by skew-symmetric bilinear α-operators on $\mathfrak{g}$.For example, the one-dimensional family of connections generated by $\alpha(x, y) = \lambda[x, y]$ satisfies the conditions of Theorem 3.13. For these connections, torsion and curvature tensors (which are themselves left-invariant) are given by:

$$T(x, y) = (2\lambda - 1)[x, y]$$

and:

$$R(x, y, z) = \lambda(\lambda - 1)[[x, y], z].$$

A pseudo-metric has same geodesics that its associated Levi-Civita connection (and all connections obtained by adding a term of torsion). Consequently, pseudo-Riemannian geodesics are group geodesics if and only if the associated (symmetric) Levi-Civita connection is a Cartan-Schouten connection.

Definition 3.14. (The canonical Cartan-Schouten connection) The left-invariant connection defined by $\alpha(x, y) = \frac{1}{2}[x, y]$ is a symmetric connection, called the *canonical Cartan-Schouten connection.*

The curvature tensor of the canonical Cartan-Schouten connection is then:

$$R(x, y, z) = -\frac{1}{4}[[x, y], z].$$

If the Levi-Civita connection of the metric is the canonical Cartan-Schouten connection, then the Riemannian geodesics equal the group geodesics and the two definitions of exponential barycenters correspond.

3.2.2 Normal elements

We have determined the pseudo-metrics whose geodesics are the group geodesics. However, there is another way to compare Riemannian geodesics and group geodesics. Given a Lie group $\mathcal{G}$, among all Riemannian geodesics going through e, which ones are one-parameter subgroups? Can we characterize them with their initial tangent vector?

Definition 3.15. (Normal elements of $\mathfrak{g}$) Let $<,>$ be a left-invariant metric on $\mathcal{G}$ and x an element of $\mathfrak{g}$. x is a *normal element of* $\mathfrak{g}$ if the one-parameter subgroup generated by x equals the geodesic starting at e with tangent vector x.

In order to characterize normal elements, we express the given left-invariant metric (more precisely the associated Levi-Civita connection) in terms of the canonical Cartan-Schouten connection.

Proposition 3.16. *A left-invariant connection is given in terms of the canonical Cartan-Schouten connection by the following formula:*

$$\alpha^L(x,y) = \frac{1}{2}[x,y] - \frac{1}{2}(ad(x)^*.y + ad(y)^*.x) \quad \forall x, y \in \mathfrak{g}$$

Proof. This is derived easily from the Koszul formula. □

Considering ad^* as a bilinear operator (depending on the metric), its symmetric part prevents the two connections of being equal for all $x, y \in \mathfrak{g}$. This means that an equation of the Lie algebra enables to read that they don't share the same set of geodesics. Can we read more? The answer is yes and we shows in the following development how to characterize geodesics that are group and metric geodesics.

Definition 3.17. (Left angular tangent vector) The *left angular tangent vector* of a curve γ is defined as:

$$x^L(t) = DL_{\gamma(t)^{-1}}.\dot{\gamma(t)} \quad \forall t \in \mathbb{R}$$

The left angular tangent vector of the Lie algebra enables to characterize group and metric geodesics. We see that a curve γ is a group geodesic if and only if:

$$x^L(t) = constant = x$$

or equivalently:

$$x\dot{(t)}^L = 0$$

as it comes directly from the definition of integral curves of a left-invariant vector field.

We also have a characterization of a connection geodesic as stated in the following proposition.

Proposition 3.18. *The curve γ is a geodesic for the left-invariant metric of Levi-Civita connection α^L if and only if we have the following dynamical equation:*

$$x\dot{(t)}^L = \alpha^L(x(t)^L, x(t)^L) \quad \forall t \in \mathbb{R}$$

This quadratic equation is called the Euler-Poincaré equation.

A derivation of this equation is given in [20]. With these two characterizations of geodesics, we can give a characterization of the normal elements of $\mathfrak{g}$.

Proposition 3.19. *The vector $x \in \mathfrak{g}$ is a normal element of $\mathfrak{g}$ iff $ad^*(x,x) = 0$.*

Proof. This is trivial using Euler-Poincarré equation and the fact:

$$\alpha^L(x,x) = -ad^*(x,x) \quad \forall x \in \mathfrak{g}.$$

□

Remark 3.20. In the case of $\mathfrak{gl}(n)$, the definition of normal elements correspond to normal matrices: a *normal matrix* is a matrix that commute with its transpose.

From the Euler-Poincaré equation, we interpret $\alpha^L(x,x)$ as an acceleration.

Definition 3.21. (Left angular acceleration vector) For a given left-invariant connection ∇ (with the corresponding α-operator), the left angular acceleration vector of γ is the covariant acceleration of the curve w.r.t ∇:

$$a^L(t) = DL_{\gamma(t)^{-1}}.\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = \alpha(x^L(t), x^L(t)) \quad \forall t \in \mathbb{R}$$

We also have:

$$a^L(t) = \alpha^L(x^L(t), x^L(t)) = -ad^*(x(t), x(t)) \quad \forall t \in \mathbb{R}$$

Remark 3.22. The previous definition depends on the connection as we took a *covariant* derivative. For example, any curve has $a^L(t) = 0$ w.r.t. the (left-invariant) canonical Cartan-Schouten connection.

Remark 3.23. For one-parameter subgroups, (i.e. group geodesics) the left angular tangent vector is constant, $x^L(t) = x$, and so is the left-angular acceleration vector $a^L(t) = \alpha^L(x,x)$. Back to the manifold, the covariant acceleration is $DL_{\gamma(t)}.\alpha^L(x,x)$ and depends on t in the general case. However, the left-invariance of the pseudo-metric implies:

- the covariant acceleration has constant norm $||\alpha^L(x,x)||$,
- the covariant acceleration is normal to the curve:

$$< \alpha^L(x,x), x >=< -ad^*(x).x, x >=< [x,x], x >= 0.$$

Despite these two properties, the covariant acceleration is not (covariantly) constant. We give here a measure of its covariant derivative, read on the Lie algebra, i.e.

$$\alpha^L(\alpha^L(x^L(t), x^L(t)), x^L(t)).$$

For one-parameter subgroup, we get in the matrix case (using $([A,B])^T = [B^T, A^T]$):

$$\begin{aligned} 2.\alpha^L(X, \alpha^L(X,X)) &= 2.\alpha^L(X, -[X^T, X]) \\ &= 2.\alpha^L(X, [X, X^T]) \\ &= [X, [X, X^T]] - [X^T, [X, X^T]] - [[X, X^T], X] \\ &= 2[X, [X, X^T]] - [X^T, [X, X^T]] \\ &= [2X - X^T, [X, X^T]] \end{aligned}$$

Hence:

$$\alpha^L(X, \alpha^L(X,X)) = [X - \frac{1}{2}.X^T, [X, X^T]],$$

which is a measure for the non-constant acceleration.

3.3 The example of SE(3)

3.3.1 SE(3) and $\mathfrak{se}(3)$

To illustrate this chapter, we consider the group of isometries of $\mathbb{R}^3$, $SE(3)$, i.e. the rotations together with the translations of the 3D real space. $SE(3)$ is defined by its action on $\mathbb{R}^3$:

$$(R,t).x = R.x + t \quad \forall x \in \mathbb{R}^3, \quad \forall (R,t) \in SE(3)$$

The group law and the group inversion are:

$$(R_1, t_1) * (R_2, t_2) = (R_1.R_2, R_1 * t_2 + t_1), \quad \forall (R_1, t_1), (R_2, t_2) \in SE(3)$$

$$(R,t)^{(-1)} = (R^{(-1)}, R^{(-1)}.(-t)), \quad \forall (R,t) \in SE(3)$$

Hence, $SE(3)$ is more precisely the semi-direct product:

$$SE(3) = SO(3) \ltimes \mathbb{R}^3$$

which amounts for the mixure of R and t terms in the translation part of the composition rule. It will be of core importance, as we shall see later.

Proposition 3.24. *The Lie algebra* $\mathfrak{se}(3)$ *is constituted by the matrices* $(A, u) \in Skew(3) \oplus \mathbb{R}^3$ *with the following Lie Bracket:*

$$[(A_1, u_1), (A_2, u_2)] = (A_1.A_2 - A_2.A_1, A_2.u_1 - A_1.u_2) \quad \forall (A_1, t_1), (A_2, t_2) \in \mathfrak{se}(3)$$

We have: $\dim(\mathfrak{se}(3)) = 6$.

Note again the mixure of the A and u terms.

A Lie algebra and its properties are determined by the Lie bracket and the structure constants. By extracting a basis from $\mathfrak{se}(3)$ and computing the structure constants, we can free us from the initial matricial construction.

Proposition 3.25. *The Lie algebra $\mathfrak{se}(3)$ is the 6 dimensional real vector space, with basis $(P_a, J'_a)_{a,a'=1..3}$ and the Lie Bracket:*

$$\begin{aligned} [P_a, P_b] &= 0, \\ [J'_a, J'_b] &= \epsilon_{a'b'c'} J'_c, \\ [P_a, J'_b] &= \epsilon_{ab'c} P_c \end{aligned}$$

where ϵ_{abc} is totally skew-symmetric in abc with $\epsilon_{123} = 1$.
The $(P_a)_a$ are the infinitesimal generators of translations, while the $(J_a)_a$ are the infinitesimal generators of rotations.

Proof. We take:

$$\begin{aligned} \forall i \in 1,2,3 : P_i = (0, e_i), &\qquad \text{with } (e_i)_i \text{ the canonical basis of } \mathbb{R}^3 \\ \forall i \in 1,2,3 : J_i = (\lambda_i, 0) &\qquad \text{with } (\lambda_i)_i \text{ the following matrices:} \end{aligned}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

and verify the previous Lie brackets by matricial computations. □

3.3.2 Geodesics of SE(3)

Now, we want to illustrate the differences between the group geodesics on SE(3) and some Riemannian geodesics. In order to do so, we provide $\mathfrak{se}(3)$ with the following inner product (i.e. Frobenius inner product on the rotation part and Euclidean inner product on the translation part):

$$< (A_1, u_1), (A_2, u_2) >= \frac{1}{2} Tr(A_1^T A_2) + u_1^T.u_2 \quad \forall (A_1, u_1), (A_2, u_2) \in \mathfrak{se}(3)$$

Translating it through left- or right-translations, we get two different Riemannian metrics on SE(3): one left-invariant and one right-invariant. Let us compute the corresponding Riemannian geodesics.

For our implementation, we use the Riemannian exponential chart associated with the *left*-invariant metric. More precisely, an element $f \in SE(3)$ is represented by its *left* Riemannian logarithm. This gives $f = \begin{pmatrix} r \\ t \end{pmatrix}$, where r is the rotation vector associated to the rotation part of an element of SE(3) and t the translation vector. We note $r = \theta.n$, where n is the unit vector of the (oriented) rotation axis and θ the rotation angle around this axis. We denote $v = \begin{pmatrix} a \\ u \end{pmatrix}$ a general element of the Lie algebra $\mathfrak{g}$. All the code for this implementation is given in Appendix.

Now, the left Riemannian geodesics γ_L in this chart are the straight lines (by definition of the chart we use):

$$\begin{cases} \gamma_L(e,v)(t) & = t.v, \\ \gamma_L(f,v)(t) & = f * \gamma_L(e, J_L(f)^{(-1)}.v)(t), \end{cases}$$

where the second line comes from the left-invariance of the Riemannian geodesics of the left-invariant metric.

As for the right Riemannian geodesics γ_R in this chart, we compute the left Riemannian geodesics associated with the inverse metric of the right one, before inverting once again. It gives:

$$\begin{cases} \gamma_R(e,v)(t) & = (t.a, t.exp(t.a).u), \\ \gamma_R(f,v)(t) & = \gamma_R(e, J_R(f)^{(-1)}.v) * f, \end{cases}$$

where the second line comes from the right-invariance of the Riemannian geodesics of the right-invariant metric.

As for the group geodesics γ_G, we compute the group exponential of v using a faithful matrix representation for SE(3) and its Lie algebra $\mathfrak{se}(3)$. In fact, $SE(3)$ can be faithfully represented by matrices of $GL(4)$ acting on $\mathbb{R}^4$ with homogeneous coordinates:

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

In this same framework, the corresponding representation of the Lie algebra is:

$$(A,u) \to \begin{pmatrix} A & u \\ 0 & 0 \end{pmatrix}$$

Now the group exponential can be computed easily using the matrice exponential. This gives:

$$\begin{cases} \gamma_G(e,v)(t) & = (t.a, u + \frac{1}{\theta^2}(1 - \frac{\sin(\theta)}{\theta})S_a.u + \frac{1}{\theta^2}(1 - cos(\theta)).S_a.u), \\ \gamma_G(f,v)(t) & = f * \gamma_G(e, J_L(f)^{(-1)}.v)(t), \end{cases}$$

where S_a is the skew-symmetric matrix generated by the vector a:

$$S_a = \begin{pmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_3 \\ a_2 & -a_3 & 0 \end{pmatrix}$$

and where the second line comes from the fact that group geodesics are in particular left invariant.

Now we can visually compare these three types of geodesics. Figure 3.2 shows a group geodesic (red), a left Riemannian geodesic (green) and a right Riemannian geodesic (blue) all starting at identity $e = \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right)$, with tangent vector $v = \left(\begin{pmatrix} 0 \\ 0 \\ \pi/3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)$, i.e. the infinitesimal generator of a rotation of $\pi/3$ around axis z, and a translation of 1 along axis x. We observe that they differ: we are precisely in the case where we can't define an admissible mean with these Riemannian metrics: they provide means that are different and different with the group exponential barycenter (as we shall see later). Interestingly, we see that the group geodesic lies between the two Riemannian ones. We observe the same thing when we plot three geodesics linking two points in SE(3), as we can see in Figure 3.3.

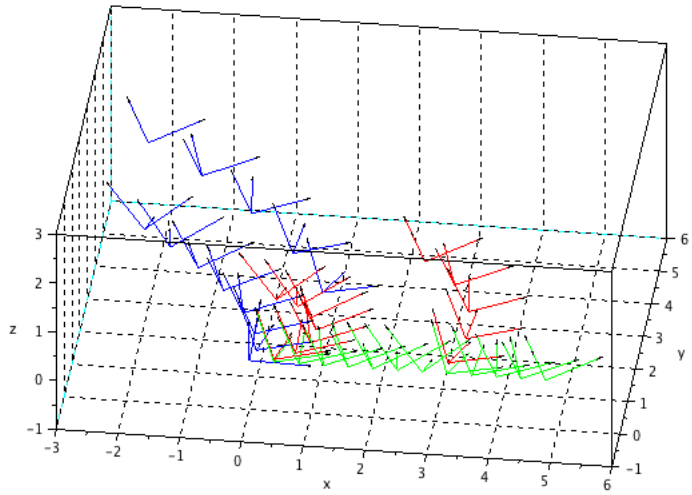

FIGURE 3.2: Group (red) and riemannian left (green) and right (blue) geodesics starting at e with tangent vector [0 0 pi/3 1 0 0]

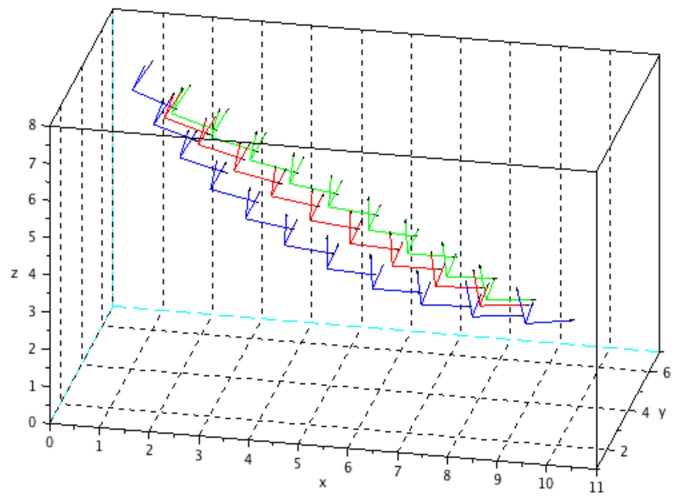

FIGURE 3.3: Group (red) and riemannian left (green) and right (blue) geodesics linking two points of SE(3)

3.3.3 Normal elements of $\mathfrak{se}(3)$

Now, we can confirm experimentally that geodesics for the normal elements of $\mathfrak{se}(3)$ are the same. First, we find analytically the normal elements of $\mathfrak{se}(3)$.

Proposition 3.26. *The normal elements of $\mathfrak{se}(3)$ are the generators of: pure rotations, pure translations and screw motions, i.e. isometries whose translation vector is parallel to the rotation axis.*

Proof. We use the characterization $ad^*(X, X) = 0$ of normal elements in the matrix case, using the matrix representation of $\mathfrak{se}(3)$. But first, we have to be careful with the following fact.

Let E be a finite dimensional vector space, V $\subset$ E a subspace of E, and u an endomorphism of E. Then u admits an adjoint u^* in E and we have:

$$< u(x), y > |_E =< x, u^*(y) > |_E \quad \forall x, y \in E$$

Now, take $x, y \in V$, and suppose that u stabilizes V. Then u admits an adjoint in V, for the restriction of $<,> |_E$ on V, i.e. $<,> |_V$: we call it $\tilde{u}^*$.

$$< u(x), y > |_V =< x, \tilde{u}^*(y) > |_V$$

Rewriting the left hand side, we get:

$$\begin{aligned}
< u(x), y > |_V &=< u(x), y > |_E \quad \text{(as u stabilizes V)} \\
&=< x, u^*(y) > |_E \quad \text{(definition of the adjoint in E)} \\
&=< x, u^*(y)|_V + u^*(y)|_{V^\perp} > |_E \quad (E = V \oplus V^\perp \text{ in finite dim for a positive definite product}) \\
&=< x, u^*(y)|_V > |_E + < u^*(y)|_{V^\perp} > |_E \\
&=< x, u^*(y)|_V > |_V + 0
\end{aligned}$$

using, for the last line, the fact that all elements belong to V for the first term, and the definition of the orthogonal for the second term.

Hence we have: $\tilde{u}^* = u^*|_V$, i.e. the orthogonal projection on V of the adjoint in E.
Applied to our case with $u = ad$, the normal elements of $\mathfrak{se}(3)$ are characterized by the condition:

$$\text{X is a normal element of } \mathfrak{se}(3) \quad \Longleftrightarrow \quad proj|_{\mathfrak{se}(3)}([X^T, X]) = 0$$

where the projection is taken regarding Frobenius inner product:

$$proj|_{\mathfrak{se}(3)}(\begin{pmatrix} A & u \\ w & c \end{pmatrix}) = \begin{pmatrix} \frac{1}{2}.(A - A^T) & u \\ 0 & 0 \end{pmatrix}$$

Indeed, we have:

$$\begin{pmatrix} A & u \\ w & c \end{pmatrix} = \begin{pmatrix} \frac{1}{2}.(A - A^T) & u \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{2}.(A - A^T) & 0 \\ w & c \end{pmatrix}$$

And:

$$< \begin{pmatrix} \frac{1}{2}.(A - A^T) & u \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}.(A - A^T) & 0 \\ w & c \end{pmatrix} > = \frac{1}{2}.Tr((\begin{pmatrix} \frac{1}{2}.(A - A^T) & u \\ 0 & 0 \end{pmatrix})^T. \begin{pmatrix} \frac{1}{2}.(A - A^T) & 0 \\ w & c \end{pmatrix})$$

$$= \frac{1}{2}.Tr(\begin{pmatrix} -\frac{1}{4}.(A - A^T).(A + A^T) & 0 \\ * & 0 \end{pmatrix})$$

$$= -\frac{1}{8}.Tr(A^2 + A^T.A - A.A^T - (A^T)^2)$$

$$= 0$$

Hence, taking $\begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix}$ in $\mathfrak{se}(3)$, we have:

$$proj|_{\mathfrak{se}(3)}([(\begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix})^T, \begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix}]) = proj|_{\mathfrak{se}(3)}(\begin{pmatrix} \Omega^T.\Omega - \Omega.\Omega^T - u.u^T & \Omega^T.u \\ u^T.\Omega & u^T.u \end{pmatrix})$$

$$= proj|_{\mathfrak{se}(3)}(\begin{pmatrix} -u.u^T & \Omega^T.u \\ u^T.\Omega & u^T.u \end{pmatrix})$$

$$= \begin{pmatrix} 0 & \Omega^T.u \\ 0 & 0 \end{pmatrix}$$

using the facts that the skew-symmetric matrix Ω commutes with its transpose and that $u.u^T$ is symmetric.
So, $\begin{pmatrix} \Omega & u \\ 0 & 0 \end{pmatrix}$ is a normal element in $\mathfrak{se}(3)$ iff $\Omega^T.u = 0$ i.e. iff $u \in Ker\Omega$ (skew-symmetry of Ω).
This is verified in the 3 following cases:

- $u = 0$, i.e. we have a generator of pure rotation,
- $\Omega = 0$ i.e. we have a generator of pure translation,
- none of the previous subcases: we have a generator of a screw motion.

Hence the result. □

As an example, we plot the three types of geodesics for the screw motion generated by: $v = \left(\begin{pmatrix} 0 \\ 0 \\ \pi/5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$, run on different simulations. This confirms that they are the same.

3.3.4 Three different means on SE(3)

The previous subsection shows that only some elements in SE(3) generate group geodesics that are also Riemannian geodesics. This means that SE(3) provides an example where the definitions of the mean do not correspond. Using the same computing framework as previously, we compare visually in this subsection:

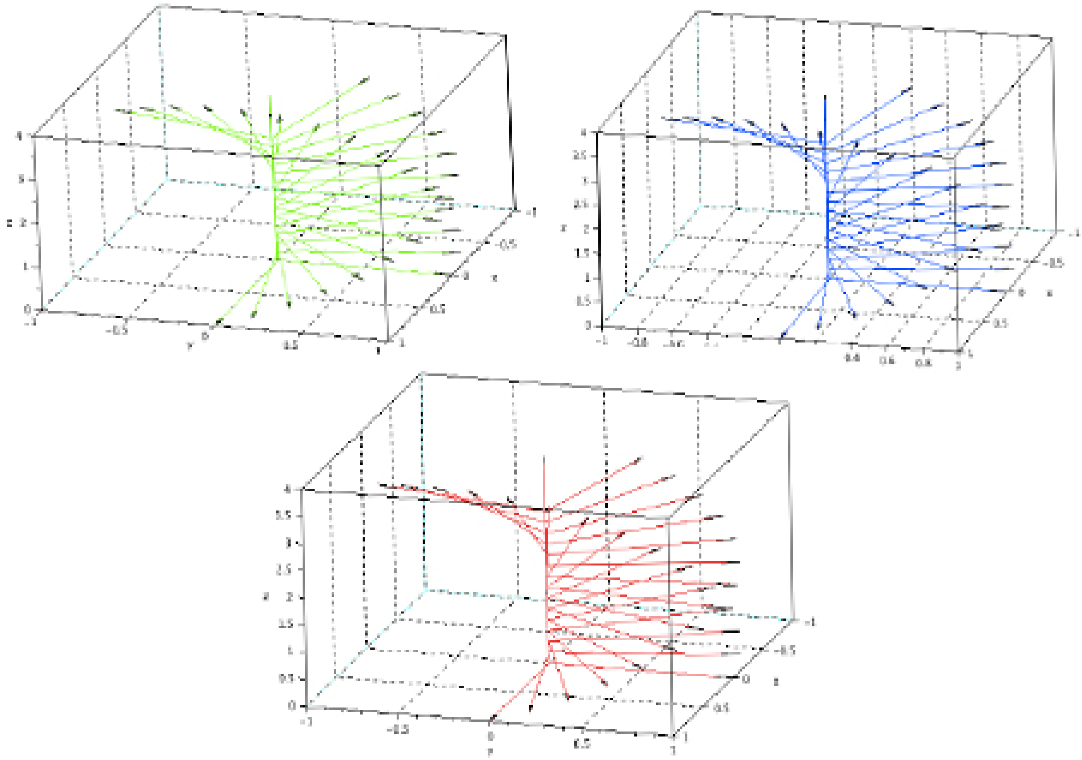

FIGURE 3.4: Riemannian left (green), right (blue) and group geodesics starting at e with tangent vector [0 0 pi/5 0 0 1], i.e. a normal element

- the Riemannian center of mass for the left-invariant metric,
- the Riemannian center of mass for the right-invariant metric,
- the group exponential barycenter.

For the computation of the Riemannian centers of mass, we implemented a Gauss-Newton algorithm for the minimization of the variance. It is widely used for non-linear least squares problems on Riemannian manifolds, because it does not need to implement the affine connection (see [5], [8] for homogeneous manifolds including SE(3), and [21] for shape spaces). The iteration can be written as follow:

Algorithm 1 (Barycentric Fixed Point Iteration on Lie groups)

- *Initialize m_0; for example with $m_0 := x_1$.*
- *Update the estimate of the mean by:*

$$m_{t+1} := exp_{m_t}\left(\sum_{i=1}^{N} w_i.log_{m_t}(x_i)\right)$$

- *Test convergence: if $||log_{m_t}(m_{t+1})||_{m_t} > \epsilon.\sigma(m_t)$ go to second step.*

The study of convergence of this algorithm is performed in [21],[22] for example.

As for the computation of the group exponential barycenter, we use the closed form given in [7]. The result of the means' computations is shown in Figure 3.5, for a data set of elements of SE(3).

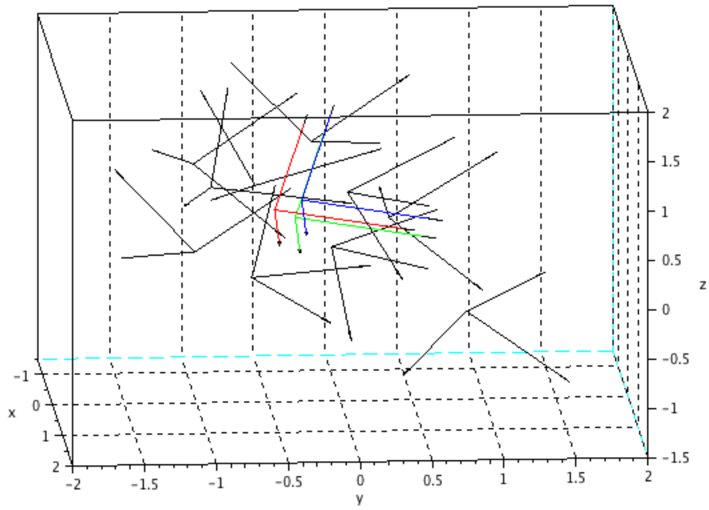

FIGURE 3.5: Riemannian left (green), right (blue) centers of mass and group exponential barycenter (red) for a given data set (black)

This confirms that they don't correspond.

3.4 Admissible definitions: consistence with the Lie group structure

In the previous section, we studied the first admissibility condition if one wants to extend the Riemannian CoM as a definition of the mean on Lie groups. Indeed, the extension has to fit the Riemannian CoM if both are defined at the same time. In this section, we analyze the second admissibility condition. This one demands that the Riemannian CoM and the definitions extending it are consistent with the group operations, i.e. that we can define a bi-invariant metric of pseudo-metric on the Lie group as we saw in the discussion from Section 3.1.

3.4.1 Bi-invariance

Hence, among those left-invariant connections and pseudo-metrics, we focus our attention on those which are bi-invariants. Let's start with the connections.

Theorem 3.27. *Among left-invariant connections, bi-invariant connections are characterized by the condition:*

$$\alpha([z,x],y) + \alpha(x,[z,y]) = [z,\alpha(x,y)] \qquad \forall x,y,z \in \mathfrak{g} \tag{3.5}$$

which is the infinitesimal version of the Ad-invariance of α*:*

$$\alpha(Ad(g).x, Ad(g.y) = Ad(g).\alpha(x,y) \qquad \forall x, y \in \mathfrak{g}, \forall g \in \mathcal{G} \tag{3.6}$$

Proof. The left-invariant connection is also right-invariant if:

$$\nabla_{DR_g.X} DR_g.Y = DR_g.\nabla_X Y$$

for any vector fields X, Y or equivalently, for any left-invariant vector fields $\tilde{X}, \tilde{Y}$. By definition of Ad, we have: $DR_{g^{-1}}.\tilde{X} = Ad(g).x$ for $\tilde{X} = DL_g.x$. The right-invariance condition is then:

$$\alpha(Ad(g).x, Ad(g).y) = Ad(g).\alpha(x,y)$$

which gives precisely the Ad invariance of α.
Take its infinitesimal version by writing $g = Exp(tz)$ and differentiating at t=0. Since:

$$\frac{d}{dt}\Big|_{t=0} Ad(Exp(tz)).x = [z, x]$$

we find the equation of the theorem. □

The one-dimensional family of (left-invariant) connections generated by $\alpha(x,y) = \lambda[x,y]$ obviously satisfies the conditions of Theorem 3.27. Consequently, every Lie group can be provided with a bi-invariant connection, for instance $\alpha(x,y) = \frac{1}{2}[x,y]$. This is the reason why the group exponential barycenter is consistent with the group operations. It is the exponential barycenter of the canonical Cartan-Schouten connection, which is bi-invariant.

However, the Riemannian CoM and the Riemannian exponential barycenter are defined using (pseudo-)metrics. We give here a characterization for bi-invariant pseudo-metrics.

Theorem 3.28. *A left-invariant pseudo-metric* $<,>$ *is bi-invariant iff one the following condition is fulfilled:*

(i)$\forall g \in \mathcal{G}$,$Ad(g)$ *is an isometry of* $\mathfrak{g}$ *for* $< . >$ *i.e.* $\quad \forall x, y \in \mathfrak{g} :< Ad(g).x, Ad(g).y >=< x, y >$

or, in its infinitesimal version:
(ii) $\forall x, y, z \in \mathfrak{g}, < ad(x).y, z > + < y, ad(x).z >= 0.$
Moreover, a bi-invariant metric is also invariant w.r.t. inversion.

Proof. (i) We show that (i) ⇔ bi-invariance.
$<,>$ is already left-invariant $g \in \mathcal{G}$: $< a, b >_g =< DL_{g^{-1}}.a, DL_{g^{-1}}.b > |_e$.
Hence, we prove: (i) ⇔ right-invariance.
Let $g \in \mathcal{G}$, $x, y \in \mathfrak{g}$. Using the left-invariance of $<,>$ we have:

$$\begin{aligned} < DR_g.x, DR_g.y >_g &=< DL_{g^{-1}}.DR_g.x, DL_{g^{-1}}.DR_g.y > |_e \\ &=< Ad(g^{-1}).x, Ad(g^{-1}).y > |_e \quad \text{as: } Ad(g^{-1}) = DC_{g^{-1}}|_e \end{aligned}$$

Hence, for all $g \in \mathcal{G}$:

$$< DR_g.x, DR_g.y >_g = < x, y > |_e \Leftrightarrow < Ad(g^{-1}).x, Ad(g^{-1}).y > |_e = < x, y > |_e$$

And we have (i) $\Leftrightarrow$ right-invariance.
(ii) Take its infinitesimal version by writing $g^{(-1)} = Exp(tz)$ and differentiating at t=0. Since:

$$\frac{d}{dt}\Big|_{t=0} Ad(Exp(tz)).x = [z, x]$$

we find the equation of the theorem. □

Moreover, the group geodesics are geodesics of such a pseudo-metric [23]. This means that the Levi-Civita connection associated to the bi-invariant metric is the canonical Cartan-Schouten connection.

In contrary of the connection case, we don't have a general family of bi-invariant pseudo-metrics on a Lie group. In fact, we'll show in Chapter 4 that some Lie groups don't possess a bi-invariant pseudo-metric.

Let focus on the necessary and sufficient condition (i). If a Lie group can be provided with a bi-invariant pseudo-metric, then $Ad(g)$ is an isometry of $\mathfrak{g}$ for this metric. Hence, for all $g \in \mathcal{G}$, $Ad(g)$ can be viewed as an element of the group $O(p, q)$, where (p,q) is the signature of the metric on $\mathfrak{g}$. This remark enables to get an efficient characterization for Lie groups with a bi-invariant metric (i.e. positive definite).

3.4.2 Special case of a positive definite metric

Let focus on the case of a positive definite metric, where we have the following proposition.

Proposition 3.29. *$O(n)$ is a maximal compact subgroup of $GL(n)$.*

Proof. Take H a compact subgroup of $GL(n)$ with $O(n) \subset H$. We want to prove $H \subset O(n)$.
Let $h \in H$, h=OS with $O \in O(n)$, and S symmetric positive definite (polar decomposition). $S = O^{-1}h \in H$. Any element h' has proper values of module 1: otherwise we can consider the divergent series h'^n or h'^{-n} which contradict the compacity assumption.
S is diagonalisable with real strictly positive proper values, hence S=Id. And h=O. □

A Lie group is said to be *relatively compact* if it is included is a compact Lie group, or equivalently if its closure is compact. Hence, Proposition 3.29 and its proof directly imply a characterization of Lie groups with bi-invariant metric from a topological property:

Theorem 3.30. *The Lie group $\mathcal{G}$ admits a bi-invariant metric if and only if its adjoint group is relatively compact.*

As Ad is continuous, the image of a compact Lie group by Ad is compact, i.e. also relatively compact.

Corollary 3.31. *Compact groups Lie groups can be provided with a bi-invariant metric* $<,>$.

Following Theorem 3.30, we conclude about the existence of a bi-invariant metric for some Lie groups.

Proposition 3.32. *SO(n) can be provided with a bi-invariant metric.*
But there is no bi-invariant metric on:

- *the group of rigid-body transformations SE(n) (for n >1),*
- *the group of scalings and translation ST(n) (for n >1),*
- *the Heisenberg group H(n) (n>1).*

Proof. SO(n) is a compact group, and the result follows from the corollary of Theorem 3.30. Now prove that SE(n) does not admit any bi-invariant metric.
We can compute easily the adjoint representation of SE(n), n>1 using the faithful representation of SE(n) in homogeneous coordinates and the fact that $Ad(M).X = M.X.M^{(-1)}$ for matrices. We get, for all $(R,t) \in SE(n)$ and $(A,u) \in \mathfrak{se}(n)$:

$$Ad(R,t).(A,u) = (R.A.R^T, -R.A.R^T.t + R.u)$$

The translation term "t" is unbounded, which prevent $Ad(\mathcal{G})$ to be bounded and hence to included in a compact. Same computations can be done for $ST(n)$ and $H(n)$ [7]. □

Interestingly, Lie groups provided with a bi-invariant metric agree on a geometric property about their curvature.

Proposition 3.33. *A Lie group provided with a bi-invariant metric has a non-negative sectionnal curvature k. Moreover, k is given by:*

$$k(x,y) = \frac{1}{4}||[x,y]||^2.$$

Proof. The sectionnal curvature in the 2-plane span(x,y) for $x, y \in \mathfrak{g}$ is defined by:

$$k(x,y) = \frac{< R(x,y)y, x >}{||x||^2||y||^2 - < x, y >^2}$$

which is left-invariant and can be computed on $\mathfrak{g}$. Moreover, k is independent on the norms of x, y or their inner product.
Hence, taking $\mathcal{G}$ a Lie group with a bi-invariant metric, i.e. in conditions of Theorem 3.28, we compute k with two orthonormal vectors of $\mathfrak{g}$:

$$k(x,y) = -\frac{1}{4} < [[x,y],y],x >$$

using the expression of the Riemann tensor: $R(x,y)z = -\frac{1}{4}[[x,y],z]$, as the Levi-Civita connection of a bi-invariant metric is the canonical Cartan Schouten connection. Now, Theorem

3.28 (ii) allows to move one bracket from the left to the right in the inner product. Hence, the sectionnal curvature reduces to:

$$k(x,y) = \frac{1}{4}||[x,y]||^2$$

which is hence non-negative. □

In this chapter we concluded on the first admissiblity condition in order to generalize the definition of mean. We have presented the second admissibility condition, i.e. the consistency of the mean's definition with the Lie group's operations. We got an efficient characterization of the existence of a bi-invariant *metric* on a Lie group. The case of a bi-invariant *pseudo-metric* is more complicated, and we present it in the next chapter.

Chapter 4

Existence of a bi-invariant pseudo-metric

In this chapter, $\mathcal{G}$ is a finite dimensional Lie group of real Lie algebra $\mathfrak{g}$. Through the characterization of Lie groups that can be provided with a bi-invariant pseudo-metric, we derive conditions in terms of their Lie algebras. In fact, the *Lie functor* gives the correspondence between the two structures (see [10] for details).

4.1 A try with Levi decomposition of Lie algebras

In order to characterize the Lie groups with bi-invariant pseudo-metric, we aim to reduce the problem to the different classes of the usual Lie group classification. The basic idea of classification of elements with a certain mathematical structure is:
(i) to find the elementary bricks of this structure,
(ii) to show how all elements of the given structure can be built from the elementary bricks.
This is what we present in the first section.

4.1.1 Ideals and semi-direct sums of Lie algebras

The first step in this approach is to find convenient substructures to decompose the Lie algebra $\mathfrak{g}$. This is the essence of the *ideals* of $\mathfrak{g}$.

Definition 4.1. (Ideal of $\mathfrak{g}$) An *ideal* $\mathfrak{h}$ of the Lie algebra $\mathfrak{g}$ is a Lie sub-algebra with the "absorbance property":

$$\forall x \in \mathfrak{h}, \text{ we have: } \forall y \in \mathfrak{g}, \quad ad(x).y = [x, y] \in \mathfrak{h}$$

If I and I' are two ideals of $\mathfrak{g}$, then I∩I', [I,I'], I⊕I', $\mathfrak{g}$/I are also ideals of $\mathfrak{g}$.

Remark 4.2. The associated structure for Lie groups is *normal subgroup.*

All the ideals of a given Lie algebra can be computed from its commutation relations, as we see in the following example.

Example 4.1. *Recall that* $\mathfrak{se}(3)$ *is defined by the commutation relations:*

$$[P_a, P_b] = 0,$$
$$[J_a, J_b] = \epsilon_{abc} J_c,$$
$$[P_a, J_b] = \epsilon_{abc} P_c.$$

We see that the only proper (i.e. not 0 and not $\mathfrak{g}$*) ideal of* $\mathfrak{se}(3)$ *is the 3-dimensionnal sub-algebra* $\mathfrak{p}$=$Span(P_a)_a$.

Some ideals provide insights about the "algebraic organization" of $\mathfrak{g}$, as for example the center and the derived algebra of $\mathfrak{g}$. The *center of a Lie algebra* $Z(\mathfrak{g})$ is the set of elements commuting with everything.

$$\begin{aligned} Z(\mathfrak{g}) &= \{x \in \mathfrak{g} / \forall x' \in \mathfrak{g}, [x, x'] = 0\} \\ &= \text{Ker ad} \end{aligned}$$

As $Z(\mathfrak{g})$= Ker *ad*, where *ad* is a representation of the Lie algebra $\mathfrak{g}$(i.e. an algebra homomorphism), $Z(\mathfrak{g})$ is an ideal of $\mathfrak{g}$. It governs whether or not the adjoint representation is faithful. Intuitively, it measures how close the Lie algebra is to be abelian.

The *derived Lie algebra* of $\mathfrak{g}$, which is also an ideal of $\mathfrak{g}$ is defined as:

$$\mathcal{D}\mathfrak{g} = Span[x, y], \quad \forall x, y \in \mathfrak{g}$$

It gives the set of elements of $\mathfrak{g}$ that can be written in terms of Lie bracket.

The construction of a Lie algebra from these elementary bricks is done through the *direct sum* or the *semi-direct sum* of Lie algebras.

Definition 4.3. (Direct sum of Lie algebras) We say that $\mathfrak{g}$ is a *direct sum of Lie algebras* $\mathfrak{g}_1$ and $\mathfrak{g}_2$, and we write it $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ if we have:
(i) $\mathfrak{g}$ is a direct sum in terms of vector spaces,
(ii) $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are ideals of $\mathfrak{g}$.

The second conditions implies that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ commute with each other, i.e. $[\mathfrak{g}_1, \mathfrak{g}_2]$=0. As a Lie algebra is the adjoint representation of itself, we can see this decomposition as a Lie algebra representation decomposition. In fact, ideals are *sub-representations of the adjoint representation* of $\mathfrak{g}$, as their defining property precisely corresponds to the stability of the subspace w.r.t. $ad(x)$, for all $x \in \mathfrak{g}$.

Now, from the representation theory, we recall the following definitions. If W has exactly two sub-representations, namely the trivial subspace 0 and W itself, then the representation is said to be *irreducible*; if W has a proper non-trivial sub-representation, the representation is said to be *reducible*. If W can be decomposed as a direct sum of irreducible sub-representations, it is said to be *completely reducible*.

For vector spaces, every subspace has a direct complement (even in infinite dimension if we assume the axiom of choice). If it were the case for Lie algebras, this would mean that the adjoint representation is always completely reducible. But it is not the case in general: every ideal does not necessary have a direct complement *which is also an ideal*. We called $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ *a semi-direct sum of Lie algebras* if only $\mathfrak{g}_2$ is an ideal. More precisely, we define:

Definition 4.4. (Semi-direct sum of Lie algebras) We say that $\mathfrak{g}$ is a semi-direct sum of Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ w.r.t. θ, and we write it $\mathfrak{g} = \mathfrak{g}_1 \oplus_\theta \mathfrak{g}_2$ if we have:
(i) $\mathfrak{g}$ is a direct sum in terms of vector spaces,
(ii) θ is a representation of $\mathfrak{g}_1$ such that:

$$\theta : \quad \mathfrak{g}_1 \longmapsto \mathrm{Der}(\mathfrak{g}_2),$$

(iii) the Lie bracket writes:

$$[x_1 + x_2, y_1 + y_2] = [x_1, y_1] + [x_2, y_2] + \theta(x_1).y_2 - \theta(x_2).y_1.$$

Under those conditions, only $\mathfrak{g}_2$ is an ideal of $\mathfrak{g}$ in the general case.

Example 4.2. $\mathfrak{se}(3)$ *decomposes as the semi-direct sum:* $\mathfrak{se}(3) = \mathfrak{so}(3) \oplus \mathfrak{p}$*. The nature of* θ *will be studied later.*

4.1.2 Simple and semi-simple Lie algebras

We have introduced the possible decomposition of a Lie algebra: as a direct sum or a semi-direct sum. However, what are the possible elementary bricks of these decompositions, for example the irreducible sub-representations? We define:

Definition 4.5. (Minimal ideals of $\mathfrak{g}$) An ideal I of $\mathfrak{g}$ is *minimal* in $\mathfrak{g}$ if it doesn't contain any ideal of $\mathfrak{g}$ other than 0 and itself.

This corresponds precisely to the definition of an *irreducible* sub-representation of the adjoint representation. Now we can characterize the Lie algebras that are completely reducible.

Definition 4.6. (Simple Lie algebras) A Lie algebra is *simple* if it doesn't admit proper ideal and it is not abelian.

Note that a simple ideal I of $\mathfrak{g}$ is also a minimal ideal of $\mathfrak{g}$, hence an irreducible sub-representation of the adjoint representation. But a minimal ideal I in $\mathfrak{g}$ may contain a proper ideal *of itself*, that is not an ideal of $\mathfrak{g}$. Simplicity is a intrinsic property, whereas minimality depends on the global Lie algebra $\mathfrak{g}$.

Hovewer, the complete reducibility property is linked to the notion of simplicity, or more precisely of semisimplicity.

Definition 4.7. (Semisimple Lie algebras) A Lie algebra is *semisimple* if it doesn't admit proper abelian ideal.
Obviously, a simple Lie algebra is semisimple.

Remark 4.8. For a semisimple or simple Lie algebra: $Z(\mathfrak{g}) = \{0\}$. Hence the adjoint representation is always faithful.

The complete reducibility of semisimple Lie algebras comes from Weyl theorem [24].

Theorem 4.9. *(Weyl Theorem) Every finite-dimensional representation of a semisimple Lie algebra is completely reducible. That is for every invariant subspace of the representation there is an invariant complement.*

Applied to the (finite dimensional) adjoint representation of the (finite dimensional) semisimple Lie algebra $\mathfrak{g}$, this shows that $\mathfrak{g}$ can be decomposed as a direct sum of minimal ideals.

In addition to this complete reducibility property, let us show that a semisimple Lie algebra can always be provided with a bi-invariant pseudo-metric. Indeed it is given by the Killing form, defined as follow.

Definition 4.10. (Killing form of a Lie algebra) The *Killing form of the Lie algebra* $\mathfrak{g}$ is the symmetric bilinear form defined as:

$$K(x, y) = Tr(ad(x) \circ ad(y)) \qquad \forall x, y \in \mathfrak{g}$$

In order for the Killing for to be a bi-invariant pseudo-metric, it has to be bi-invariant and non-degenerate. Let us show these two properties and first the bi-invariance.

Proposition 4.11. *(Properties of the Killing form)*
(i) K is invariant under $Aut(\mathfrak{g})$*:* $\forall \phi \in Aut(\mathfrak{g}), K(\phi(x), \phi(y)) = K(x, y)$.
(ii) K generates a bi-invariant symmetric bilinear form on $\mathcal{G}$.

Proof. (i). Let $\phi \in Aut(\mathfrak{g})$. ϕ respects the Lie bracket, which can be written:

$$\phi([x, y]) = [\phi(x), \phi(y)] \quad \forall x, y \in \mathfrak{g} \text{ i.e. } \phi \circ ad(x) = ad(\phi(x)) \circ \phi \quad \forall x \in \mathfrak{g}$$

Hence we have: $ad(\phi(x)) = \phi \circ ad(x) \circ \phi^{-1}$ which implies (i) by cyclicity of the trace.
This means that all automorphisms are isometries w.r.t. the Killing form.
(ii) K bilinear by linearity of Tr and ad, and also symmetric because of the cyclicity of Tr. Now prove its bi-invariance.
We make K left-invariant by propagating it at every $g \in \mathcal{G}$ through left translations:

$$K_g(a, b) = K(DL_{g^{-1}}.a, DL_{g^{-1}}.b).$$

The right-invariance comes from Theorem 3.28, (i), noting that it does not require the form to be non-degenerate at this point. For all $g \in \mathcal{G}$, Ad(g) is an automorphism of $\mathfrak{g}$ and hence an isometry for the Killing form. The Killing form is thus bi-invariant. □

Now the Killing form has the bi-invariance property which is required. Moreover it is non-degenerate in the semisimple case, as stated by the Cartan criterion (proof can be found in [10]).

Theorem 4.12. *(Cartan criterion)*

$$\mathfrak{g} \text{ semi-simple} \iff K \text{ non-degenerate.}$$

Example 4.3. *(Example of Killing forms) Let $\mathcal{G} = GL_n$, $\mathfrak{g} = \mathfrak{gl_n} = \mathcal{M}_n$ provided with the commutator as the Lie bracket.*
Then, the corresponding Killing form is:

$$K_{\mathfrak{gl_n}}(X,Y) = 2nTr(X.Y) - 2Tr(X).Tr(Y), \quad \forall X, Y \in \mathfrak{gl_n}$$

For $\mathfrak{so}(n)$ we have:

$$K_{\mathfrak{so}(n)}(X,Y) = (n-2)Tr(X.Y).$$

Hence, we solved our problem for Lie algebras whose adjoint representation is completely reducible, i.e. semisimple Lie algebras. The Killing form of a semisimple Lie algebra $\mathfrak{g}$ is non-degenerate and provides a bi-invariant pseudo-metric on the Lie algebra $\mathfrak{g}$. What about Lie algebras that are not completely reducible? Indeed, if the direct sum of Lie algebras is not possible, the semi-direct sum always exists: this is the essence of the following theorem [25].

Theorem 4.13. *(Levi decomposition) Any Lie algebra $\mathfrak{g}$ can be written as the semi-direct sum $\mathfrak{g} = Rad \oplus \mathfrak{s}$ where $\mathfrak{s}$ is a semisimple Lie sub-algebra of $\mathfrak{g}$, called a Levi-subalgebra and Rad is an ideal called the radical of $\mathfrak{g}$. The Levi sub-algebra always exists but may not be unique.*

Note how this decomposition differs from those of Cartan, Iwasawa or the root decomposition of Lie algebras: they only concern semisimple Lie algebras, on which our problem is already solved. Hence, we don't delve into those in this thesis and focus on the Levi decomposition.

The Levi decomposition invite us:

- to see if we can define a bi-invariant pseudo-metric on the Radical,
- and to try to recombine (in some way) this metric on the global Lie group $\mathfrak{g} = Rad \oplus \mathfrak{s}$, with the Killing form on the semisimple part.

We would have a method to construct a bi-invariant pseudo-metric on any algebra. Hence, let have a look on the radical of $\mathfrak{g}$, which is in fact the maximal *solvable* ideal of $\mathfrak{g}$.

4.1.3 Solvable Lie algebras

In this subsection, we precise the properties of the Radical givenby the Levi decomposition Theorem. This leads to the definition of new categories of Lie algebras which rely on the following well-known series.

Definition 4.14. (The derived series and the central series of $\mathfrak{g}$)
The *derived series* of $\mathfrak{g}$ is defined as follow:

$$\mathcal{D}^1(\mathfrak{g}) = \mathfrak{g} \quad \text{and} \quad \forall n, \mathcal{D}^{n+1}(\mathfrak{g}) = \mathcal{D}(\mathcal{D}^n(\mathfrak{g})) = [\mathcal{D}^n(\mathfrak{g}), \mathcal{D}^n(\mathfrak{g})]$$

Each $\mathcal{D}^n\mathfrak{g}$ is an ideal of $\mathfrak{g}$.
The *central decreasing series* of $\mathfrak{g}$ is defined as follow:

$$\xi_1(\mathfrak{g}) = \mathfrak{g} \quad \text{and} \quad \forall n, \xi_{n+1}(\mathfrak{g}) = [\mathfrak{g}, \xi_n(\mathfrak{g})]$$

The *central ascending series* of $\mathfrak{g}$ is defined as follow:

$$\xi(\mathfrak{g})^1 = 0 \quad \text{and} \quad \forall n, \xi^{n+1}(\mathfrak{g}) = \{x \in \mathfrak{g}/\forall y \in \mathfrak{g}, [x,y] \in \xi^n(\mathfrak{g})\}$$

For both central decreasing series and derived series $\mathfrak{g}$ becomes more and more abelian through the iterations. We have also:

$$\mathcal{D}\mathfrak{g} = \xi_2(\mathfrak{g}).$$

Hence, if $\mathcal{D}\mathfrak{g} = \mathfrak{g}$, then both series are constants with all terms equal $\mathfrak{g}$.

For the central ascending series, the elements become less and less abelian. We have also:

$$Z(\mathfrak{g}) = \xi^2(\mathfrak{g}).$$

Hence, if the center is $\{0\}$ (i.e. no elements that commute with everything), then the central ascending series is constant equal to $\{0\}$.

Example 4.4. *In the case of* $\mathfrak{se}(3)$*, the center is:*

$$Z(\mathfrak{se}(3)) = (P_a)_a$$

Consequently, the adjoint representation is not faithful. As we have also: $\xi^3(\mathfrak{g}) = (P_a)_a$ *directly from the definition, the central ascending serie is constant, for* $n \geq 2$*, and equal to the ideal* $(P_a)_a$*. Then, the derived Lie algebra is:*

$$\mathcal{D}(\mathfrak{se}(3)) = \mathfrak{se}(3)$$

Both derived and central series are constants and equal the whole Lie algebra $\mathfrak{se}(3)$*. Moreover, any element of* $\mathfrak{se}(3)$ *can be written as a commutator.*

These series are used to define another category of Lie algebra, with null intersection with the semisimple one.

Definition 4.15. (Solvable and nilpotent Lie algebras) A Lie algebra $\mathfrak{g}$ is *solvable* if there exists n s.t. $\mathcal{D}^n(\mathfrak{g}) = \{0\}$. A Lie algebra $\mathfrak{g}$ is *nilpotent.* if there exists n s.t. $\xi_n(\mathfrak{g}) = \{0\}$ or, equivalently, if there exists n' s.t. $\xi^{n'}(\mathfrak{g}) = \mathfrak{g}$

The following proposition comes directly from the definitions of the central and derived series.

Proposition 4.16. *An abelian Lie algebra is nilpotent. A nilpotent Lie algebra is solvable.*

Moreover, we have:

Proposition 4.17. *(Measures of commutativity) Let* $\mathfrak{n}, \mathfrak{a}$ *be some nilpotent and abelian Lie algebras,*

$$Z(\mathfrak{a}) = \mathfrak{a}$$
$$Z(\mathfrak{n}) \neq 0$$

Proof. The first equality is the characterization of $\mathfrak{a}$ being abelian. The second one is obviously by taking the smallest integer m verifying the defining condition of $\mathfrak{n}$ being nilpotent:

$$Z(\mathfrak{n}) = \xi_{m-1}(\mathfrak{n})$$

and hence, never 0. □

Remark 4.18. For an abelian Lie algebra, the adjoint representation is trivial, for a nilpotent Lie algebra it is never faithful. For a solvable Lie algebra, it can be faithful or not.

Solvable and semisimple Lie algebras correspond to disjoint categories of Lie algebras. With the following precision on the Radical from Levi-decomposition, they also enable to characterize any Lie algebra with the Levi decomposition.

Definition 4.19. (Radical of $\mathfrak{g}$) The radical $Rad(\mathfrak{g})$ of $\mathfrak{g}$ is its maximal solvable ideal.

In fact, now we can see the Levi decomposition as a semi-direct product of semisimple and solvable parts. Now, our problem reduces to the question: can we put a bi-invariant pseudo-metric on the solvable part? If yes, we could think of an eventual recombinaison of the bi-invariant pseudo-metrics on the solvable and the semisimple part. But we have (see [10]):

Theorem 4.20. *(Killing-Cartan criterion)*

$$\mathfrak{g} \textit{ solvable} \quad \Longleftrightarrow \quad K(\mathfrak{g}, \mathcal{D}(\mathfrak{g})) = 0.$$

If $\mathcal{D}(\mathfrak{g}) \neq 0$, then the Killing form is degenerate. On the other hand, $mathcalD(\mathfrak{g}) = 0$ means that $\mathfrak{g}$ is abelian, which implies that the adjoint represention is trivial and the Killing form constantly 0. Hence, the Killing form can't define a bi-invariant pseudo-metric on a solvable Lie algebra.

In fact, we can't find a general way to define a bi-invariant pseudo-metric on solvable Lie algebras. Some solvable Lie algebra simply don't admit a bi-invariant pseudo-metric, for example the Heisenberg group H_{2m+1} as shown in the next section. But, even if the solvable sub-algebra of the Levi decomposition doesn't admit a bi-invariant pseudo-metric, there still can be a way to globally define one. For example, the semi-direct product $\mathbb{R} \ltimes H_{2m+1}$ defines an *oscillator group* which posses a bi-invariant pseudo-metric [26]. Thus, it seems that the Levi decomposition with the usual Lie algebra classification semisimple/solvable is not the most pertinent decomposition for our problem.

Hence, we investigate in the next section another type of decomposition, where we still use the vocabulary of simple, semisimple, abelian, nilpotent and solvable Lie algebras. We present necessary conditions to be fulfilled when one Lie group admits a bi-invariant pseudo-metric and hence reject some Lie groups. Then we will characterize all Lie groups admitting a bi-invariant pseudo-metric following the work of [27],[28].

4.2 A specific decomposition

Assume that the Lie algebra $\mathfrak{g}$ is *provided with a bi-invariant pseudo-metric* $<,>$. Hence, we have an inner product on the Lie algebra, verifying the conditions of Theorem 3.28. This metric structure enables new properties for the ideals of $\mathfrak{g}$ and new types of decomposition of the Lie algebra.

4.2.1 Ideals of $\mathfrak{g}$ with a bi-invariant pseudo-metric

Now that we have a pseudo-metric, we could think of decomposing any Lie algebra into irreducible representations of the adjoint representation, i.e. minimal ideals, by taking the orthogonal of each minimal ideal we find. However, it is not that simple and we have to make a distinction between *degenerate* and *non-degenerate* ideals.

Lemma 4.21. *If I is an ideal of $\mathfrak{g}$ then its orthogonal $I^\perp$ w.r.t. the bi-invariant $<,>$ is also an ideal of $\mathfrak{g}$. Moreover, we have $[I, I^\perp] = 0$ i.e. the second axiom for a direct sum of Lie algebras.*

Proof. It is direct using (ii) of Theorem 3.28. □

However, this lemma doesn't enable to decompose any Lie algebra with bi-invariant $<,>$ onto a direct sum of minimal ideals. If $<,>$ is not positive or negative definite, then the sum of an ideal and its orthogonal is not necessary direct, in the sense of vector spaces. Some vectors have norm 0 without being null: for the Lorentz pseudo-metric for instance, they correspond to momenta of massless particles. Such vectors are orthogonal to themselves: if they belong to an ideal, they also belong to its orthogonal.

Definition 4.22. (Degenerate and completely degenerate ideals) An ideal I of $\mathfrak{g}$ is *degenerate* (or *isotropic*) if $I \cap I^\perp \neq 0$. It is *completely degenerate* (or *completely isotropic*) if $I \subset I^\perp$.

Note that a degenerate ideal of dimension 1 is totally degenerate.we have $I \cap I^\perp \neq 0$ and necessary $I \cap I^\perp = I$ since I has dimension 1. This implies $I \subset I^\perp$. Note also that a totally degenerate ideal is abelian. $[I, I^\perp]=0$.

Following from the previous discussion, we can only write $\mathfrak{g} = I \oplus I^\perp$ in the non-degenerate case. However, we always have $\dim I + \dim I^\perp = \dim \mathfrak{g}$ as $I^\perp$ is the intersection of kernels of independent linear forms $< j, \bullet >$, where j are the basis vectors of I.

In order to characterize Lie algebras that can be provided with a bi-invariant pseudo-metric, we investigate first necessary conditions. The following concern the central series of such a Lie algebra [27], [28].

Proposition 4.23. *If $\mathfrak{g}$ admits a bi-invariant pseudo-metric $<,>$, we have:*

$$\forall n \in \mathbb{N}, \quad \xi^n(\mathfrak{g})^\perp = \xi_n(\mathfrak{g}),$$

and hence:

$$\dim \xi^n(\mathfrak{g}) + \dim \xi_n(\mathfrak{g}) = \dim \mathfrak{g}.$$

Proof. We show the equality by induction on $n \geq 1$. It is obvious for $n = 1$ as $<,>$ is non-degenerate.
Now suppose we have $\xi^n(\mathfrak{g})^\perp = \xi_n(\mathfrak{g})$ for some n and prove it for n+1.
We have:

$$\begin{aligned} z \in \xi^{n+1}(\mathfrak{g}) &\Leftrightarrow \forall a \in \mathfrak{g}, [a,z] \in \xi^n(\mathfrak{g}) \quad \text{(definition of central increasing series)} \\ &\Leftrightarrow \forall a \in \mathfrak{g}, [a,z] \in \xi_n(\mathfrak{g})^\perp \quad \text{(induction)} \\ &\Leftrightarrow z \in \xi_{n+1}(\mathfrak{g})^\perp \end{aligned}$$

where the last line is obtained from Theorem 3.28 (ii):

$$\text{for all } a, z \in \mathfrak{g}, x \in \xi_n(\mathfrak{g} : \quad < [a,z], x >=< z, [a,x] >$$

and the non-degeneracy of $<,>$.
Hence we get the result for all n. The equality on dimensions follows. □

This condition already enables to recognize Lie groups that cannot be provided with bi-invariant pseudo-metric. For example, applying it with $n = 2$ gives:

$$[\mathfrak{g}, \mathfrak{g}]^\perp = Z(\mathfrak{g})$$

and:

$$\dim Z(\mathfrak{g}) + \dim \mathcal{D}(\mathfrak{g}) = \dim \mathfrak{g}.$$

We can conclude that solvable Lie algebras with $Z(\mathfrak{g}) = 0$ don't admit any bi-invariant pseudo-metric. This is precisely what happens for the Heisenberg algebra [27].

Let us go on with another necessary condition.

Proposition 4.24. *If $\mathfrak{g}$ admits a bi-invariant pseudo-metric $<,>$, we have:*
(i) I is an ideal of dimension 1 $\Rightarrow I \subset Z_{\mathfrak{g}}(\mathfrak{g})$, i.e. I is central.
(ii) I is a non-degenerate abelian ideal $\Rightarrow I \subset Z_{\mathfrak{g}}(\mathfrak{g})$, i.e. I is central.

Proof. (i) Take $I = \mathbb{R}z$ an ideal of dimension 1.
If I is non degenerate, we can write: $\mathfrak{g} = I \oplus I^\perp$ with $[I, I^\perp]$=0. As I is abelian, we also have:

$[I, I] = 0$. Hence I commutes with everything, i.e. is central by definition.
If I is degenerate, then it is totally degenerate as it has dimension 1. For all x $\in \mathfrak{g}$, since I is an ideal of dimension 1, we have $[a, x] = \alpha(x).a$ where α is a linear form on $\mathfrak{g}$. The condition (ii) of Theorem 3.28 writes:

$$\alpha(x) < a, y > +\alpha(y) < a, x >= 0$$

Suppose α is not equally 0. Then, $x \in \mathrm{Ker}\alpha$ is necessary orthogonal to I: <a,x>=0. We take z such that <z,a>=1. Then, $\alpha(z) \neq 0$. But the previous identity gives:

$$2\alpha(z) < a, z >= 0$$

which is a contradiction. Hencewe have to conclude that α is equally 0.
(ii) Take an abelian non-degenerate ideal I. We write: $\mathfrak{g} = I \oplus I^{\perp}$. Now, take $a \in i$, $x \in \mathfrak{g}$, we have:

$$[a, x] = [a, b] + [a, b'] = 0 + [a, b']$$

as I is abelian. The second bracket is also 0 because we always have $[I, I^{\perp}] = 0$, see Lemma 4.21. Hence, I is central. □

4.2.2 Minimal ideals of $\mathfrak{g}$

Within our attempt of the construction of a bi-invariant pseudo-metric using the Levi decomposition, the only minimal ideals we considered were the simple ideals. Indeed, they are crucial in the usual decomposition of Lie algebras.

But we saw that the Levi decomposition of a Lie algebra was not pertinent for our problem. Hence, we create our own decomposition and go back to the most general notion of minimal ideals to play the role of elementary bricks. We characterize them in the case of $\mathfrak{g}$ provided with a bi-invariant pseudo-metric through the following dichotomie.

Lemma 4.25. *Take I a minimal ideal of $\mathfrak{g}$.*
(i) If I is non-degenerate, then I is a simple ideal or of dimension 1,
(ii) If I is degenerate, then it is completely degenerate (and also abelian.)

Note the link between minimal and simple ideals in the non-degenerate case, which justifies our approach in Section 4.1.

Proof. (i) In this case we can write $\mathfrak{g} = I \oplus I^{\perp}$ and any ideal of I is also an ideal of $\mathfrak{g}$.Indeed, take A an ideal of I, and $x \in \mathfrak{g}$. We have:

$$[x, a] = [x_I, a] + [x_{I^{\perp}}, a]$$

using the non-degeneracy of I: the sum $I \oplus I^{\perp}$ is direct in terms of vector spaces. The second term is null because $[I, I^{\perp}] = 0$ for a direct sum of Lie algebras. As A is an ideal of I, we have $[x_I, a] \in A$ and hence $[x, a] \in A$. So any ideal of I is also an ideal of $\mathfrak{g}$ in the non-degenerate case and I is simple because it is a minimal ideal of $\mathfrak{g}$.

(ii) In this case, $J = I \cap I^{\perp}$ is also an ideal of $\mathfrak{g}$, included in a minimal ideal: I. Hence $J = I$ and $I \subset I^{\perp}$. As $I^{\perp} \subset Z_{\mathfrak{g}}(I)$, I is abelian. □

This lemma gives a first insight about the construction of the bi-invariant pseudo-metric. If $\mathfrak{g}$ could be decomposed onto its minimal ideals, we could construct the bi-invariant pseudo-metric it is provided with. According to Lemma 4.25, any minimal ideal is a semisimple or an abelian ideal, ie two classes on which we know a construction. But do we actually have such a decomposition? The answer is no: the decomposition is slightly more complicated.

4.2.3 Decomposition of $\mathfrak{g}$

In this section, we'll prove the following theorem [27], which corresponds to the pertinent decomposition of a Lie algebra with regards to the existence of a bi-invariant pseudo-metric.

Theorem 4.26. *If $\mathfrak{g}$ admits a bi-invariant pseudo-metric, then $\mathfrak{g}$ can be decomposed into a direct orthogonal sum of three ideals $\mathfrak{g}_1, \mathfrak{g}_2, \mathfrak{g}_3$ such that:*
(1) $\mathfrak{g}_1$ is the bigest semi-simple ideal of $\mathfrak{g}$,
(2) $\mathfrak{g}_2 = \oplus_p^{\perp} I_p$, direct orthogonal sum of one-dimensional ideals of $\mathfrak{g}$,
(3) $\mathfrak{g}_3 = \oplus_k^{\perp} J_k$, direct orthogonal sum of non-simple ideals of $\mathfrak{g}$, with only degenerate proper ideals.

In order to prove this theorem, we first need the following lemma.

Lemma 4.27. *If $\mathfrak{g}$ admits a bi-invariant pseudo-metric, then semisimple ideals are non-degenerate.*

Proof. Take I an ideal of $\mathfrak{g}$. Elements of $I \cap I^{\perp}$ are orthogonal to themselves, i.e. $I \cap I^{\perp} \subset (I \cap I^{\perp})^{\perp}$. This means that $I \cap I^{\perp}$ is totally degenerate, and hence abelian. If I is semisimple, then $I \cap I^{\perp}$ is also semisimple and it is 0 as it is abelian. Consequently, a semisimple ideal is non-degenerate. □

Proof. (Proof of Theorem 4.26) The proof constructs the corresponding ideals using to the previous lemmas.
Take $\mathfrak{g}_1$ to be the bigest semi-simple ideal of $\mathfrak{g}$ (can be $\mathfrak{g}$ or 0).
Following Lemma 4.27 with the semisimple ideal $\mathfrak{g}_1$ we write:

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_1^{\perp}.$$

Now $\mathfrak{g}_1^{\perp}$ is obviously non-semisimple. Because we have a direct sum of Lie algebras, ideals of $\mathfrak{g}_1^{\perp}$ are ideals of $\mathfrak{g}$. Hence, they can't be simple otherwise they would be in $\mathfrak{g}_1$. From the dichotomy of Lemma 4.25, minimal ideals are either non degenerate ideals of dimension 1 or totally degenerate abelian ideals.

Take I_1 a non-degenerate *minimal ideal* of $\mathfrak{g}_1^{\perp}$ (of dimension 1) and write:

$$\mathfrak{g}_1^{\perp} = I_1 \oplus^{\perp} I_1^{\perp}.$$

Minimal ideals of $I_1^\perp$ are also minimal ideals of $\mathfrak{g}_1^\perp$, and we repeat the process until there are no more non-degenerate minimal ideals:

$$\mathfrak{g}_1^\perp = I_1 \oplus^\perp I_2 \oplus^\perp ...I_p \oplus^\perp \mathfrak{g}_3$$

where:

- the I_p are all non-degenerate ideals of $\mathfrak{g}_1$,
- $\mathfrak{g}_3$ is the orthogonal of I_p in I_{p-1}, hence an ideal of I_{p-1} and hence an ideal of $\mathfrak{g}_1$.

Moreover, any minimal ideal of $\mathfrak{g}_3$ is degenerate.

Let go to the decomposition of $\mathfrak{g}_3$. Take J_1 a minimal non-degenerate ideal of $\mathfrak{g}_3$ (different from a non-degenerate *minimal ideal* of $\mathfrak{g}_3$) and write:

$$\mathfrak{g}_3 = J_1 \oplus \perp J_1^\perp$$

Hence we get a decomposition where the J_k are non-simple with all ideals degenerate. □

If we can characterize the non-simple algebra J_k with bi-invariant pseudo-metric so that J_k has only degenerate ideals, then we can characterize the Lie algebras $\mathfrak{g}$ which admit bi-invariant pseudo-metric. Indeed, the previous decomposition leads to the construction of the pseudo-metric on all $\mathfrak{g}$, as it is written as a orthogonal sum of parts where we know how to define a bi-invariant pseudo-metric.

4.3 Construction of the bi-invariant pseudo-metric

4.3.1 Lie groups with bi-invariant pseudo-metric

Now, we finally go to the characterization of non-simple Lie groups which admit a bi-invariant pseudo-metric for which any proper ideal of the Lie algebra is degenerate and we give a method to explicitely construct it.

Theorem 4.28. *Let $\mathcal{G}$ be a simply connected non simple Lie group with Lie algebra $\mathfrak{g}$. Then $\mathcal{G}$ can be provided with a bi-invariant pseudo-metric such that any proper ideal of its Lie algebra is totally degenerate if and only if $\mathcal{G}$ is isomorphic to $K \ltimes B^*$, where K is simple, non abelian and simply connected and the semi-direct product is given by the co-adjoint representation η.*

Remark 4.29. The structure $K \ltimes B^*$ corresponds in fact to a very specific construction called *double extensions*. More details about these can be found in [28].

For a proof of this theorem, we refer to [27]. Note that it is in fact a generalization of the statement defining the oscillator groups [26]. However, let us explicit the construction of the bi-invariant pseudo-metric under the conditions of Theorem 4.28.

Let $\mathcal{G}$ be a Lie group under the condition of Theorem 4.28. Its decomposition is:

$$\mathcal{G} = K \ltimes B^*$$

where B^* is the dual vector space of B, the Lie algebra of K.

The co-adjoint representation of the *Lie group* K on B^* is:

$$(\eta(k))(\alpha).b = \alpha(Ad(k^{(-1)}.b)$$

so that the semi-direct product group $G = K \ltimes B^*$ (where B^* is the abelian group for $+$) has composition law:

$$(\alpha, k) * (\beta, k') = (\alpha + (\eta(k))(\beta), k.k')$$

according to the theorem.

Then, the left-invariant pseudo-metric $<,>$:

$$< (\alpha, b), (\beta, b') >= \alpha(b') + \beta(b)$$

is bi-invariant on $\mathcal{G}$ [27]. We'll verify it in the case of SE(3) in the next subsection, in order to compute on a more concret example.

Remark 4.30. In terms of Lie algebras, taking the differential at e of η, the corresponding co-adjoint representation of the *Lie algebra* B on B^* is:

$$(\theta(b))(\alpha).b' = -\alpha[b, b'].$$

Moreover, the corresponding semi-direct sum of Lie algebras is :

$$\mathfrak{g} = B \oplus B^*$$

where B^* is taken as abelian, with Lie bracket:

$$[(\alpha, b), (\beta, b')] = (\theta(\beta)b - \theta(b')\alpha, [b, b']).$$

4.3.2 A bi-invariant pseudo-metric on SE(3)

This result allowed us to build a pseudo-metric bi-invariant on SE(3), as stated in the following theorem. In fact, the action of SO(3) on $\mathbb{R}^3$ defining the semidirect product of SE(3) is precisely the co-adjoint action of Theorem 4.28.

Theorem 4.31. *The semi-direct product $\ltimes_*$ defining the group of SE(3)=SO(3)$\ltimes_*\mathbb{R}^3$ is the co-adjoint action of the simple group SO(3) on $\mathbb{R}^3$, taken as the dual of the Lie algebra of SO(3), namely $\mathfrak{so}(3)^*$.*

Proof. We know that the action of $SO(3)$ on $\mathbb{R}^3$ defining the semi-direct product of $SE(3)$ is: $R * t = R.t$. However, t has to be taken as an element $\tilde{t}$ of $\mathfrak{so}(3)^*$, i.e. a linear form on skew-symmetric matrices. We write:

$$R * \tilde{t}.A =< R.t, A > \quad \text{for all } A \in \mathfrak{so}(3)$$

where <,> is the usual Euclidean inner product on $\mathbb{R}^3$, when we represent the skew-symmetric matrix A by a vector a such that $A = S_a$.

On the other hand, the co-adjoint action of $SO(3)$ is, by definition: $\eta(R) * \tilde{t}.A =< t, R^T.A.R >$.

Do we have equality for these actions? Rewriting the skew-symmetric matrices in terms of the associated vectors, and noting that:

$$R^T.A.R = S_{R^T.a},$$

we find:

$$\eta(R) * \tilde{t}.A =< t, R^T.a >=< R.t, a >= R * \tilde{t}.A \quad \forall A$$

Indeed, the actions correspond and we are in the conditions of Theorem 4.28. □

Now a bi-invariant pseudo-metric on $SE(3)$ is given by the construction of the previous subsection. We explicit it and check its bi-invariance.

Theorem 4.32. *The following pseudo-metric of signature (3,3) is bi-invariant on SE(3):*

$$< (A, u), (B, v) >= a^T.v + b^T.u$$

where a, b are the unique vectors associated to the skew-symmetric matrices A and B.

Proof. In the principal chart, where $v = (a, u)$, this pseudo-metric has matrix: $Q = \begin{pmatrix} 0 & Id_3 \\ Id_3 & 0 \end{pmatrix}$ which is diagonalisable with diagonal: $D = \begin{pmatrix} Id_3 & 0 \\ 0 & -Id_3 \end{pmatrix}$ Hence it is not degenerate and has signature (3,3).

Now show that this pseudo-metric is bi-invariant using (i) of Theorem 3.28, i.e. proving that $Ad(R, t)$ is an isometry for all $(R, t) \in SE(3)$.
Take $(R, t) \in SE(3)$, and recall from Chapter 3, Section 3.4, that:

$$Ad(R, t).(A, u) = (R.A.R^T, -R.A.R^T.t + R.u)$$

For the computations, we denote any skew-symmetric matrix with its associated vector, i.e. for example: $A = S_a$. As $-R.A.R^T = S_{-R.a}$, the adjoint action can be rewritten:

$$Ad(R, t).(a, u) = (R.a, (-R.a) \wedge t + R.u).$$

Compute:

$$\begin{aligned} < Ad(R,t).(a,u), Ad(R,t).(b,v) > &= (R.a)^T.\,((-R.b)\wedge t + R.v) + (R.b)^T.\,((-R.a)\wedge t + R.u) \\ &= a^T.v + b^T.u + (R.a)^T.(-R.b)\wedge t + (R.b)^T.(-R.a)\wedge t \end{aligned}$$

Here we recognize the expression of the determinant in $\mathbb{R}^3$: $[x,y,z] = x^T.(y \wedge z)$. Its total skew-symmetry gives finally:

$$< Ad(R,t).(a,u), Ad(R,t).(b,v) >= a^T.v + b^T.u$$

Hence, $Ad(R,t)$ is an isometry which shows the bi-invariance of $<,>$. □

Hence, we have characterized in this chapter the Lie groups that admit a bi-invariant pseudo-metric, i.e. the Lie groups where we have an admissible definition of the mean as a Riemannian exponential barycenter. The characterization we found is certainly less explicit at first glance that in the case of (positive definite) bi-invariant metrics. However, it is computationally efficient, as we illustrated with the example of SE(3).

Chapter 5

Existence and uniqueness of the means

For now on, we have presented the geometric structures on Lie groups (with their exponential/logarithm maps) that allow to write the definition of the Riemannian exponential barycenter and the group exponential barycenter. Then, we have investigated conditions in order for these means to be admissible, also as extensions of the Riemannian center of mass on manifolds. Now, assuming that we are on the right structure with the right admissiblity conditions for the mean's definition, do we have existence and uniqueness of the mean? This is the purpose of this chapter, where we first investigate the case of the Riemannian CoM as a guideline, before extending to the two other means. Note that we'll often write about probability measures with compact support, which is the case for our data sets of discrete points.

5.1 Riemannian CoMs

In this section, $\mathcal{G}$ is a Lie group endowed with a bi-invariant metric, i.e. positive definite. The Riemannian CoM, as defined in Section 3.1, is an admissible definition for a mean on the Lie group. We investigate here the existence and uniqueness conditions for this mean. In Section 5.1.1, we prove the following geometric theorem :

Theorem 5.1. *The Riemannian CoM of a given data set* $\{x_i\}$ *exists and is unique on any regular geodesic ball of* $\mathcal{G}$.

In Section 5.1.2, we characterize the so-called *regular geodesic balls* for a given Lie group with a bi-invariant metric, investigating the example of SO(3) in Section 5.1.3. Once again, we deal with the geometric notions first, adding the Lie group properties in a second time only.

5.1.1 Notions of convexity

In this section we deal with topological notions. Hence define the canonical objects for this kind of study, *the geodesic balls.*

Definition 5.2. (Geodesic ball and geodesic sphere) Let $B_\rho(0)$ the ball of radius ρ in $T_p\mathcal{M}$. If $\rho < inj_p(\mathcal{M}, <,>)$, we call $\mathcal{B}_\rho(p) = exp_p\{B_\rho(0)\}$ the geodesic ball of center p and radius ρ.

Remark 5.3. According to Hopf-Rinow theorem [10], any geodesic ball is complete.

Now define some notions of convexity. The following definition for an affine manifold should not be mixed up with a property of complete spaces, where there is an *unique minimal geodesic* linking two points. As for convexity, the unique geodesic is not necessary minimizing.

Definition 5.4. (Convexity of an affine manifold) An affine manifold $(\mathcal{M}, \nabla)$ is said to be *convex* if for every pair of points $p, q \in \mathcal{M}$, there exists an unique geodesic, defined using ∇, joining p and q, and this geodesic depends smoothly on p and q.

Canonical convex elements of a complete Riemannian manifold are the following geodesic balls.

Definition 5.5. (Regular geodesic ball) Let $(\mathcal{M}, <,>)$ be a complete Riemannian manifold, and $\mathcal{B} = \overline{\mathcal{B}_\rho(p)}$ a closed geodesic ball of center p and radius ρ. $\mathcal{B}$ is a *regular geodesic ball* if:
(i) $\sqrt{\kappa}\rho < \frac{1}{2}\pi$,
(ii) the cutlocus of the center p does not meet $\mathcal{B}$, ie $\rho < inj_p(\mathcal{M}, <,>)$,
where κ is the supremum of sectional curvatures of $\mathcal{M}$ in $\mathcal{B}$ (can be $+\infty$), or 0 if the supremum is negative.

Theorem 5.6. *A regular geodesic ball $\mathcal{B}$ of a complete affine manifold $(\mathcal{M}, <,>)$ is convex.*

The proof can be found in [18]. In order to find the regular geodesic balls of a manifold, we need its injectivity radius $inj_\mathcal{M}$. A consequence of Rauch comparison theorem (see [29]) gives a lower bound on $inj_\mathcal{M}$ in the case of a *finite* κ. Roughly speaking, it links κ to the rate at which geodesics spread apart. For small (or negative) κ, geodesics tend to spread from each other, while for large κ, geodesics tend to converge. In this last case, two geodesics starting from the same point x might converge toward each other and meet at a point y. If y is precisely their first meeting point, the length along them between x and y is the injectivity radius of x. The following corollary quantifies this discussion:

Corollary 5.7. *Let $(\mathcal{M}, <,>)$ be a Riemannian manifold, with κ a finite supremum of its sectional curvatures. We have:*
(i) If $\kappa \leq 0$ then $inj_\mathcal{M} = \infty$,
(ii) If $\kappa > 0$, then $inj_\mathcal{M} > \frac{\pi}{\sqrt{\kappa}}$.

Hence, for manifolds with bounded sectional curvatures κ, a regular geodesic ball is a closed geodesic ball such that $\sqrt{\kappa}\rho < \frac{1}{2}\pi$, with the previous notations, i.e. we don't need the condition on $\mathrm{inj}_\mathcal{M}$ anymore. This will be useful for the characterization of regular geodesic balls in Lie groups.

The convexity of the manifold should not be mixed up with its eventual *convex geometry*, determined by a convex function, called a *separating function*. We have the following definitions.

Definition 5.8. (Convex function) Let $(\mathcal{M}, \nabla)$ an affine manifold. A function $\phi : \mathcal{M} \longmapsto \mathbb{R}$ is a *convex function* if:

$$\phi \circ \gamma = \mathbb{R} \longmapsto \mathbb{R}$$

is a convex function for all geodesics γ. More precisely, if γ is a geodesic connecting p to q, we have:

$$\phi \circ \gamma(t) < t\phi(x) + (1-t)\phi(y) \qquad \forall t \in [0,1]$$

Lemma 5.9. *For a* $\mathcal{C}^2$ *function* ϕ*, this is equivalent to* $\frac{D}{dt}\frac{d}{dt}\phi \circ \gamma(t) \geq 0$.

Definition 5.10. (Separating function) Let $(\mathcal{M}, \nabla)$ be an affine manifold and $\mathcal{M} \times \mathcal{M}$ the product manifold with product connection. A function $\phi : \mathcal{M} \times \mathcal{M} \longmapsto \mathbb{R}$ is a separating function on $\mathcal{M}$ if it is convex and vanishing exactly on the diagonal Δ of $\mathcal{M} \times \mathcal{M}$.

Definition 5.11. (Convex geometry) Let $(\mathcal{M}, <,>)$ be a complete Riemannian manifold and $\mathcal{B}$ a compact submanifold with same dimension. The domain $\mathcal{B}$ is said to have a *weak convex geometry* if there is a continuous non-negative convex function:

$$\Psi : \mathcal{B} \times \mathcal{B} \longmapsto [0, +\infty[$$

vanishing precisely on the diagonal $\Delta = \{(p,p)/p \in \mathcal{B}\}$.
If Ψ can be chosen to be bounded, we say that $\mathcal{B}$ has a *bounded convex geometry.*

Remark 5.12. An affine manifold with weak convex geometry is not necessary convex. For example, $\mathbb{R}^m \setminus \{0\}$ is not convex even though the distance function is a separating function.

Intuitively, this means that the geometry of $\mathcal{M}$ can be specified by the smooth convex function Ψ instead of the metric distance. For instance, if $\mathcal{M}$ has a non-positive sectionnal curvature $\kappa \leq 0$, then suitable candidates are $\Psi(p,q) = \text{dist}(p,q)$ or $\Psi(p,q) = \text{dist}^2(p,q)$. Hence, manifold with $\kappa \leq 0$ have convex geometry. Complete, simply connected, Riemannian manifolds with $\kappa \leq 0$ are called *Cartan-Hadamard* manifolds.

However, this case doesn't interess us because of Proposition 3.33: a Lie group provided with a bi-invariant metric has non-negative sectional curvature $\kappa \geq 0$. Hence, it requires more work to construct a separating function and to show its eventual convex geometry. For such manifolds, we use the regular geodesic balls as intermediates and prove the uniqueness of the Riemannian CoM on them.

5.1.2 Existence and uniqueness

Now that we have defined notions convexity, we prove the existence and uniqueness of the Riemannian CoM. The existence will be straighforward, but the uniqueness requires two more lemmas. As one could expect, they deal with convexity. The first lemma shows that a regular geodesic ball is not only convex (see Theorem 5.6), but has also a weak convex geometry.

Lemma 5.13. *Let $(\mathcal{M}, <,>)$ be a complete Riemannian manifold, with geodesic distance dist. Let $\mathcal{B}$ be a regular geodesic ball of radius ρ and center o, with positive upper curvature bound κ such that $\sqrt{\kappa}\rho < \frac{1}{2}\pi$. Then $\mathcal{B}$ has a bounded convex geometry, provided by the following separating function:*

$$\Psi^{(\kappa)}_{\nu,\tilde{h}}(x,y) = \Big(\frac{1 - cos(\kappa dist(x,y))}{cos(\kappa dist(x,o))cos(\kappa dist(y,o)) - \tilde{h}^2}\Big)^{\nu+1}$$

with conditions $\nu \geq 1$ and $2\nu\tilde{h}^2(h^2 - \tilde{h}^2) \geq 1$, $h = cos(\kappa R)$ and $\tilde{h} \in]0, h[$.

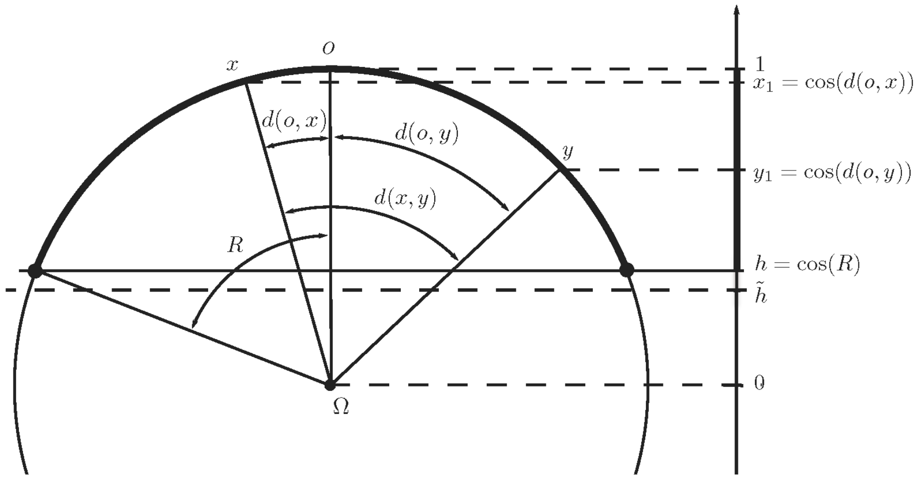

FIGURE 5.1: Bounded convex geometry for a regular geodesic ball of $\mathcal{S}^1$

Proof. This proof generalizes the proof of [18], where a separating function is given for a small hemisphere $S^n_{h,+}$. We consider regular geodesic balls with upper curvature bound 1, as the general case can be obtained by a scaling argument.

First, it is clear that $\Phi^{(\kappa)}_{\nu,\tilde{h}}$ is non-negative. We have the non-negativity of the numerator:

$$1 - \cos(\mathrm{dist}(x,y)) \geq 0,$$

and:

$$\cos(\mathrm{dist}(x,o)) \geq h \quad \text{and} \quad \cos(\mathrm{dist}(y,o)) \geq h$$

which implies the non-negativity of the denominator:

$$\cos(\mathrm{dist}(x,o))\cos(\mathrm{dist}(y,o)) - \tilde{h}^2 \geq h^2 - \tilde{h}^2 > 0.$$

Composed with the ν-power function, $\nu \geq 1$, we get the non-negativity of $\Phi^{(\kappa)}_{\nu,\tilde{h}}$.

Moreover, $\Phi^{(\kappa)}_{\nu,\tilde{h}}$ is bounded:

$$\Phi^{(\kappa)}_{\nu,\tilde{h}}(x,y) \leq \Big(\frac{2}{h^2 - \tilde{h}^2}\Big)^{\nu+1}$$

and it is clear that $\Phi^{(\kappa)}_{\nu,\tilde{h}}$ is continuous and vanishes only on the diagonal Δ.We denote $\Phi = \Phi^{(\kappa)}_{\nu,\tilde{h}}$ in the following development.

Let show the convexity of Φ by considering its second derivative $[\Phi'']_0$ taken along an arbitrary geodesic in $\mathcal{B} \times \mathcal{B}$, which is just a coordinate pair of two arbitrary geodesics of $\mathcal{B}$.

Frst define some notations. We take $x, y \in \mathcal{B}$ and two arbitrary geodesics λ, μ passing through x and y respectively, such that $\gamma = (\lambda, \mu)$ is a geodesic of $\mathcal{B} \times \mathcal{B}$ going through (x, y). We have:

$$\lambda(0) = x, \quad \lambda'(0) = u$$
$$\mu(0) = y, \quad \mu'(0) = v.$$

We also define an intermediate function Ψ as:

$$\Psi(x, y) = \frac{p}{q}$$

where:

$$p = 1 - \cos(\text{dist}(x, y))$$
$$q = \cos(\text{dist}(x, o)) \cos(\text{dist}(y, o)) - \tilde{h}^2.$$

In the following development, all these functions will be taken as functions of t, the parameter of the geodesic γ.

We show a lower bound on the intermediate function $[\Psi'']_0$, before considering $[\Phi'']_0$. Compute the derivatives of Ψ at t=0:

$$[\Psi']_0 = \frac{d}{dt}\Big|_{t=0} \Psi(\lambda(t), \mu(t)) = \frac{p'q - qp'}{q^2} = \frac{p' - \Psi q'}{q} \qquad (*)$$

Derive once again:

$$\begin{aligned} [\Psi'']_0 &= \frac{d^2}{dt^2}\Big|_{t=0} \Psi(\lambda(t), \mu(t)) \\ &= \frac{p''q - pq''}{q^2} - 2\frac{p'q - pq'}{q^2}\frac{q'}{q} \\ &= \frac{p'' - \Psi q''}{q} - \frac{2\Psi' q'}{q} \quad \text{using } p' = \Psi q' + q\Psi' \text{ from } (*). \end{aligned}$$

At this point, we have to compute the derivatives of p, q w.r.t. the variable t. The derivation of composition gives:

$$p' = -\sin(\text{dist}(x, y))[< \text{grad dist}(x, y), u > + < \text{grad dist}(x, y), v >]$$

$$\begin{aligned} q' = \sin(\text{dist}(x, o)). \cos(\text{dist}(y, o)). &< \text{grad dist}(x, o), u > \\ &+ \cos(\text{dist}(x, o)). \sin(\text{dist}(y, o)). < \text{grad dist}(y, o), v > \end{aligned}$$

We use the notations:

$$\begin{aligned}
x_1 &= \cos(\mathrm{dist}(x,o)),\\
y_1 &= \cos(\mathrm{dist}(y,o)),\\
u_1 &= \frac{d}{dt}\Big|_{t=0}\cos(\mathrm{dist}(\lambda(t),o)) = \sin(\mathrm{dist}(x,o)). < \mathrm{grad}\ \mathrm{dist}(x,o), u >,\\
v_1 &= \frac{d}{dt}\Big|_{t=0}\cos(\mathrm{dist}(\mu(t),o)) = \sin(\mathrm{dist}(y,o)). < \mathrm{grad}\ \mathrm{dist}(y,o), v > .
\end{aligned}$$

Hence, we have: $q' = y_1.u_1 + x_1.v_1$.

We derive one again and use the following inequalities whose computation can be found in [30].

$$p'' \geq \frac{(p')^2}{2p} - (|u|^2+|v|^2)p = \frac{(\Psi q' + q\Psi')^2}{2\Psi q} - (|u|^2+|v|^2)\Psi q,$$

$$q'' \leq 2u_1v_1 - (|u|^2+|v|^2)(q+\tilde{h}^2)$$

where we express everything in terms of Ψ and q (no p).

Hence, we can proceed to the lower bound of $[\Psi'']_0$:

$$\begin{aligned}
[\Psi'']_0 &= \frac{p''}{q} - \frac{\Psi q''}{q} - \frac{2\Psi' q'}{q}\\
&\geq \frac{(\Psi q' + q\Psi')^2}{2\Psi q^2} - (|u|^2+|v|^2)\Psi - 2\frac{\Psi u_1 v_1}{q}\\
&\qquad\qquad + (|u|^2+|v|^2)\Psi + \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) - 2\frac{\Psi' q'}{q}\\
&= \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi q' + q\Psi')^2}{2\Psi q^2} - 2\frac{\Psi u_1 v_1}{q} - 2\frac{\Psi' q'}{q}\\
&= \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi q' + q\Psi')^2}{2\Psi q^2}\\
&\qquad - \frac{(\Psi q')^2}{2\Psi q^2} + [\frac{(\Psi q')^2}{2\Psi q^2} - 2\frac{\Psi u_1 v_1}{q}] - 2\frac{\Psi' q'}{q}
\end{aligned}$$

where we have added and substracted $\frac{(\Psi q')^2}{2\Psi q^2}$ on the last line, in order to consider the term under brackets:

$$[\frac{(\Psi q')^2}{2\Psi q^2} - 2\frac{\Psi u_1 v_1}{q}] = \frac{\Psi}{2q^2}\left(y_1^2u_1^2 - 2(x_1y_1 - 2\tilde{h}^2)u_1v_1 + x_1^2v_1^2\right),$$

using the previous expression of q'.

We recognize a quadratic form in u_1, v_1, which has discriminant:

$$\begin{aligned}
\Delta &= 4\big(\frac{\Psi}{2q^2}\big)^2[(x_1y_1 - 2\tilde{h}^2)^2 - x_1^2y_1^2]\\
&\leq 0 \quad \text{as } x_1y_1 \geq h^2 > 0
\end{aligned}$$

The coefficients of u_1^2 and v_1^2 are positives, so the previous quadratic form is in fact positive-definite.

This allows us to precise the lower bound on Ψ'':

$$
\begin{aligned}
[\Psi'']_0 &\geq \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi q' + q\Psi')^2}{2\Psi q^2} - \frac{(\Psi q')^2}{2\Psi q^2} - 2\frac{\Psi' q'}{q} \\
&= \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi q' + q\Psi')^2}{2\Psi q^2} - \frac{(\Psi q')^2}{2\Psi q^2} - 2\frac{\Psi' q'}{q} \\
&= \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi q')^2 + 2(\Psi q')(q\Psi') + (q\Psi)^2}{2\Psi q^2} - \frac{(\Psi q')^2}{2\Psi q^2} - 2\frac{\Psi' q'}{q} \\
&= \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) + \frac{(\Psi')^2}{2\Psi} - \frac{\Psi' q'}{q} \\
&\geq \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) - \frac{\Psi' q'}{q}
\end{aligned}
$$

This lower bound is not necessary non-negative, so Ψ is not necessary convex. However, recall that Ψ is only an intermediate function. Now, we have to find a transformation ϕ such that $\Phi(x, y) = \phi(\Psi(x, y))$ is convex and of the form given by the theorem.

Compute the derivatives of Φ at $t = 0$:

$$[\Phi']_0 = \frac{d}{dt}\Big|_{t=0} \Phi(\lambda(t), \mu(t)) = \phi'(\psi).[\Psi']_0$$

$$[\Phi'']_0 = \frac{d^2}{dt^2}\Phi(\lambda(t), \mu(t))\big|_{t=0} = \phi''(\psi)([\Psi']_0)^2 + \phi'(\psi)[\Psi']_0$$

We take $\phi(\psi) = \psi^\nu$ for some $\nu \geq 1$. Hence:

$$\phi'(\psi) = \nu\psi^{\nu-1} > 0 \quad \text{and} \quad \phi''(\psi) = \frac{\nu}{\psi}\phi' > 0$$

if $\psi > 0$.

Now, taking into account the lower bound on $[\Psi'']_0$, we have

$$[\Phi'']_0 = \phi'\big[\frac{\nu}{\psi}(\psi')^2 + [\Psi'']_0\big]$$

and hence:

$$[\Phi'']_0 \geq \phi'\big[\frac{\nu}{\psi}(\psi')^2 + \frac{\tilde{h}^2\Psi}{q}(|u|^2+|v|^2) - \frac{\Psi' q'}{q}\big]$$

But:

$$q' = y_1.u_1 + x_1.v_1$$

which implies:

$$(q')^2 \leq 2(|u|^2+|v|^2)$$

and hence:

$$[\Phi'']_0 \geq \phi'\big[\frac{\nu}{\psi}(\psi')^2 + \frac{\tilde{h}^2\psi}{2q}(q')^2 - \frac{\Psi' q'}{q}\big]$$

Now, the expression in brackets is a quadratic form in q', ψ' with discriminant:

$$(\phi')^2(\frac{1}{q^2} - 2\frac{\tilde{h}^2\nu}{q} = (\frac{\phi'}{q})^2(1 - 2\nu\tilde{h}^2 q) \leq (\frac{\phi'}{q})^2(1 - 2\nu\tilde{h}^2(h^2 - \tilde{h}^2))$$

This discriminant is negative if we impose: $2\nu\tilde{h}^2(h^2 - \tilde{h}^2) \geq 1$. It implies that the quadratic form is positive-definite, as the coefficients in q', ψ' are positives. Hence:

$$[\Phi'']_0 \geq 0 \quad \text{under the condition } \psi > 0 \text{, i.e. everywhere off-diagonal.}$$

But we can extend it to $\psi = 0$, i.e. Δ (a strict submanifold) since Φ is a $\mathcal{C}^2$ function on $\mathcal{B} \times \mathcal{B}$. Hence, we get the convexity of Φ everywhere. □

The second lemma we need for the uniqueness of the Riemannian CoM is the following Jensen's type inequality.

Lemma 5.14. *Let $(\mathcal{M}, <,>)$ a complete Riemannian manifold, $\mathcal{B}$ a geodesic ball in $\mathcal{M}$ and ϕ a smooth convex bounded function on $\mathcal{M}$. If μ is a probability measure with compact support contained in the interior of $\mathcal{B}$, and if m is a Riemannian center of mass of μ, then:*

$$\phi(m) \leq \int \phi(x)\mu(dx)$$

Proof. Let γ be a geodesic starting at m. By convexity of the function ϕ, we have:

$$\phi(\gamma(t)) \geq \phi(\gamma(0)) + t. < \text{grad}\phi, \frac{d\gamma}{ds}(0) > \quad \forall t \geq 0$$

Take $x \in \mathcal{B}$. As $\mathcal{B}$ is a complete connected manifold, there is a unique *minimizing* geodesic γ connecting m and x. Using this geodesic in the previous equation, we have:

$$\phi(x) \geq \phi(\gamma(0)) + \text{dist}(x, m). < \text{grad}\phi, \frac{d\gamma}{ds}(0) >$$

Now, as m is a Riemannian CoM, it is a local minimum of the variance:

$$\sigma_x^2(y) = \mathbf{E}[\text{dist}(\mathbf{x}, y)]^2 = \int_{\mathcal{M}} \text{dist}(x, y)^2 \mu(dx)$$

and hence a zero of its gradient. We have:

$$\text{grad}\sigma(m) = \int \text{dist}(m, x)\text{grad}_1\text{dist}(m, x)\mu(dx) = 0$$

where grad_1 dist is the gradient of dist w.r.t. the first argument. Noting that:

$$\frac{d\gamma}{ds}(0) = \text{grad}_1\text{dist}(m, x)|_q,$$

we can now integrate our inequation on $\mathcal{B}$ over μ to get:

$$\begin{aligned}\int_{\mathcal{B}} \phi(x)\mu(dx) &\geq \int_{\mathcal{B}} \phi(\gamma(0))\mu(dx) + < \text{grad}\phi, \int_{\mathcal{B}} \frac{d\gamma}{ds}(0).\text{dist}(x,m)\mu(dx) > \\ &= \phi(m) + < \text{grad}\phi, \text{grad}\sigma(m) > \\ &= \phi(m)\end{aligned}$$

The integration on $\mathcal{M}$ is equivalent to the integration on $\mathcal{B}$ as the support of μ is contained in $\mathcal{B}$ and therefore:

$$\phi(m) \leq \int_{\mathcal{M}} \phi(x)\mu(dx)$$

□

We are finally able to prove the theorem stated at the beginning of the section, i.e. existence and uniqueness for the Riemannian CoM.

Theorem 5.15. *(Existence and uniqueness for Riemannian CoM) Let $(\mathcal{M}, <,>)$ be a complete Riemannian manifold and $\mathcal{B}$ a regular geodesic ball. If μ is a probability measure with compact support in $\mathcal{B}$, then there is one and only one Riemannian CoM m of μ within $\mathcal{B}$.*

Proof. Existence The gradient of the variance on a point $x \in \partial\mathcal{B}$ is inward pointing: it shows the variation of the distance of point x to a point belonging to $\mathcal{B}$. Hence, there is at least one local minimum of the variance in $\mathcal{B}$, i.e. at least one Riemannian CoM.
Uniqueness It follows from Lemma 5.13 and Lemma 5.14. Suppose m and m' are two Riemannian CoMs for μ in $\mathcal{B}$. Then (m, m') is a Riemannian CoM of the image of μ under the diagonal map:

$$\delta : \mathcal{B} \longmapsto \mathcal{B} \times \mathcal{B}.$$

Let Φ be a continuous non-negative bounded convex function vanishing only on the diagonal Δ of $\mathcal{B} \times \mathcal{B}$, whose existence is given by Lemma 5.13. Then, by Lemma 5.14:

$$\Phi(m, m') \leq \int \Phi(p, p')\delta(dp, dp') = \int \Phi(p,p)\mu(dp) = 0$$

by definition of the diagonal map δ and the fact that Φ vanishes on the diagonal. Hence, m=m' which proves the uniqueness. □

We have shown that the Riemannian CoM of a data set exists and is unique *on any regular geodesic ball* of a complete Riemannian manifold. Now we want to find the *maximal domain* on which we have uniqueness. For the specific case of $\mathcal{S}^n$, [18] shows that the strict upper hemisphere $\mathcal{S}^n_+$ does not possess convex geometry, whereas any small hemisphere has. Radius of $\mathcal{S}^n_+$ is given by the limit of acceptable radii for regular geodesic ball. Intuitively, it would represent the maximal domain where we can get a unique Riemannian CoM. More precisely, this domain $\mathcal{U}$ is the union (even infinite, as the infinite union of open sets is still open) of regular geodesic balls interiors:

$$\mathcal{U} = \bigcup_n \mathring{\mathcal{B}}_n.$$

The proof of existence and uniqueness on a domain like $\mathcal{U}$ will be done in a more general setting in the next section.

5.1.3 Application to Lie groups: uniqueness domain

Now take into account the algebraic properties of the Lie group. We compute global domains where we have existence and uniqueness of the Riemannian CoM.

Given a Lie group $\mathcal{G}$ endowed with a bi-invariant metric, we want to characterize its regular geodesic balls which provide such global domains as stated in Theorem 5.15. Recalling the definition of a regular geodesic ball, we compute the lowest upper bound on the sectionnal curvature κ (eventually $+\infty$) and the injectivity radius $\text{inj}_{\mathcal{G}}$.

Start with κ. For a Lie group endowed with a bi-invariant metric $<,>$, the sectional curvature is bi-invariant and can be computed at e. It is given by:

$$k(x,y) = \frac{1}{4}||[x,y]||^2 = \frac{1}{4}||ad(x).y||^2$$

where we take x, y two orthonormal vectors of the Lie algebra.

Now, assume that $ad(x) : y \longmapsto ad(x).y$ is a *continuous* linear map. This is true in particulary for finite dimensional Lie groups, as the matrix groups we consider. Then, $ad(x)$ possess a subordinate norm $|||.||| < \infty$, relative to the bi-invariant one $||.||$. Assume that the linear map $ad : x \longmapsto ad(x)$ is also *continuous*, we define its norm $||||.||||< \infty$, subordinate to $|||.|||$ and $||.||$. Now, we have:

$$\begin{aligned}
k(x,y) &= \frac{1}{4}||ad(x).y||^2 \\
&\leq \frac{1}{4}|||ad(x)|||^2.||y||^2 \\
&= \frac{1}{4}|||ad(x)|||^2 \quad \text{(as y has norm 1)} \\
&\leq \frac{1}{4}||||ad||||^2.||x||^2 \quad \text{(as x has norm 1)} \\
&= \frac{1}{4}||||ad||||^2
\end{aligned}$$

Hence, under the previous continuity assumptions, we reduce to the case of Lie groups with (positive) bounded sectional curvatures. For now on, we consider this special case. This implies that the characterization of regular geodesic balls only depends on κ (see Corollary 5.7), i.e. there is no need of $\text{inj}_{\mathcal{G}}$. From the previous computation:

$$\kappa = \frac{1}{4}||||ad||||^2$$

is the lowest upper bound of k for a Lie group endowed with bi-invariant metric, by properties of the subordinate norms.

A straightforward method to compute κ uses the structures constants. Take T_i, T_j two orthonormal basis vectors of the Lie algebra, we have:

$$\begin{aligned} k(T_i, T_j) &= \frac{1}{4}||[T_i, T_j]||^2 \\ &= \frac{1}{4}||c_{ijk}T_k||^2 \quad \text{using Einstein sumation notation} \\ &= \frac{1}{4}\sum_k |c_{ijk}|^2 \quad \text{as the } (T_k)_k \text{ form an orthonormal basis} \end{aligned}$$

Then we get $\kappa = \sup_i \sup_j \sum_k |c_{ijk}|^2$.

However, we want to see precisely what happens with computations on subordinate norms. To this purpose, take $<,>$ the Euclidean norm on $\mathfrak{g}$ (i.e. Frobenius norm if we have a matrix space). We have the well-know equaliy: $||ad(x)||_2 = \rho(ad(x))$, with ρ the spectral radius of $ad(x)$, i.e. the supremum of the norms of its eigenvalues. As we have chosen the half of the usual Frobenius norm in our setting, we have in fact:

$$||ad(x)||_2 = 2.\rho(ad(x))$$

Hence, $||||ad|||| = 2. \sup_{||x||_2=1} \rho(ad(x))$. With our choice of Frobenius norm, we have the lowest upper bound:

$$\kappa = \sup_{||x||_2=1} \rho(ad(x)).$$

Example 5.1. *We compute it on the example of SO(3). We take the usual orthonormal basis of* $\mathfrak{so}(3)$ *for the Frobenius inner product:* $< X, Y >= \frac{1}{2}Tr(X^T.Y)$*:*

$$J_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

As $[J_i, J_j] = \varepsilon_{ij}{}^k J_k$*, i.e.* $(ad(J_i))_{kj} = \varepsilon_{ijk}$*, we have:*

$$ad(J_1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad ad(J_2) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad ad(J_3) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

For each element of the orthonormal basis: $\rho(ad(J_i)) = 1$*. However,* $\kappa = \sup_{||X||_2=1} \rho(ad(X))$ *is the smallest majorant of the* $\rho(ad(X))$*, for* $X \in \mathfrak{so}(3)$ *of norm 1. Take such a X, and its* *decomposition onto the basis:* $X = w_i J_i$ *where* $\sqrt{\sum_i |w_i|^2} = 1$. *We have:*

$$\begin{aligned}
\rho(ad(X))^2 &= \frac{1}{4}|||ad(w_i J_i)|||^2 \\
&= \frac{1}{4}|||w_i.ad(J_i)|||^2 \quad \textit{by linearity of ad} \\
&\leq \frac{1}{4}|w_i|^2.|||ad(J_i)|||^2 \quad \textit{using the triangle inequality on } |||.||| \\
&= \frac{1}{4}.4.\sum_i |w_i|^2 \quad \textit{because } |||ad(J_i)||| = 2 \textit{ as computed before} \\
&= 1 \quad \textit{by assumption on the } w_i
\end{aligned}$$

1 is a majorant of $\rho(ad(X))$ *which is reached for the* J_i*: it is the smallest majorant. Hence* $\kappa = 1$*, which is the well-known result for SO(3).*

The regular geodesic balls of SO(3) are those of radius $\rho < \frac{\pi}{2}$*. Hence, we have existence and uniqueness of the Riemannian CoM for any data set included in such a ball. As we shall see, we also have existence and uniqueness for the limit domain* $\mathcal{U}$*:*

$$\mathcal{U} = \bigcup_{n \geq 1} \mathring{\mathcal{B}}_{(1-\frac{1}{n})\frac{\pi}{2}}(e)$$

which is precisely the strict upper hemisphere of [18] in this case. It is indeed maximal. If we take the upper hemisphere, we could have two different means of a data set, for example if we take only data on the equator.

This setting is computationnally efficient. However, the Riemannian CoM is an admissible definition of mean for Lie groups endowed with bi-invariant *metric*. This condition is restrictive as we saw in Chapter 3. In the next section, we generalize the results above in the case of (pseudo-)Riemannian and a group exponential barycenters.

5.2 Pseudo-Riemannian, group exponential barycenters

Both pseudo-Riemannian exponential barycenters and group exponential barycenters are defined by a exponential barycentric equation:

$$\sum_i w_i.\mathbf{log_m}(x_i) = 0 \tag{5.1}$$

where log_m, when it is defined (see Chapter 2), is taken to be either the pseudo-Riemannian logarithm or the group logarithm.

As they are generalizations of the Riemannian CoMs, we can think of a generalization of Theorem 5.15 for them. Note that we don't have any metric anymore for the group barycenter case. However, in both cases, we are provided with an affine connection structure: take the Levi-Civita connection associated to the pseudo-Riemannian metric in one case and the canonical Cartan-Schouten connection in the other case. In fact, the definitions of convexity, weak convexity geometry, and geodesic balls still hold for an affine manifold [31].

5.2.1 Existence and uniqueness

We prove the existence and uniqueness of the exponential barycenters on other domains than the regular geodesic balls of the previous section. In fact, regular geodesic balls need the definition of the sectionnal curvature which we don't have in the case of an affine manifold.

Definition 5.16. An affine manifold $(\mathcal{M}, \nabla)$ is said to have a *p-convex geometry* for some $p \in 2\mathbb{N}$ if there is a smooth separating function ϕ, such that:

$$c.d^p \leq \phi \leq C.d^p$$

for some constants $0 < c < C$ and some auxiliary Riemannian distance function d.

Example 5.2. *For instance, any regular geodesic ball of the sphere $\mathcal{S}^n$ has a p-convex geometry with p depending on its radius [31].*

Definition 5.17. A convex affine manifold $(\mathcal{M}, \nabla)$ is said to be *CSLCG (convex, with semilocal convex geometry)* if every compact subset K of $\mathcal{M}$ has a relatively compact convex neighborhood U_K which has p-convex geomtry for some $p \in 2\mathbb{N}$ depening on K.
Equivalently, a convex affine manifold $(\mathcal{M}, \nabla)$ is CSLCG if there exists an increasing sequence $(\mathcal{U}_n)_{n\geq 1}$ of relatively compact open convex subsets of $\mathcal{M}$ such that:
(i) $\mathcal{M} = \bigcup_{n\geq 1} \mathcal{U}_n$,
(ii) $\forall n \geq 1, \quad \mathcal{U}_n$ has a p-convex geometry for some $p \in 2\mathbb{N}$, depending on n.

Example 5.3. *For instance the open hemisphere $\mathcal{S}^n_+$ is a CSLCG manifold, as we can take $\mathcal{U}_n = \mathring{\mathcal{B}}_{\rho_n}(p)$ with $\rho_n = (1 - \frac{1}{n})\frac{\pi}{2}$, see Figure 5.2.*

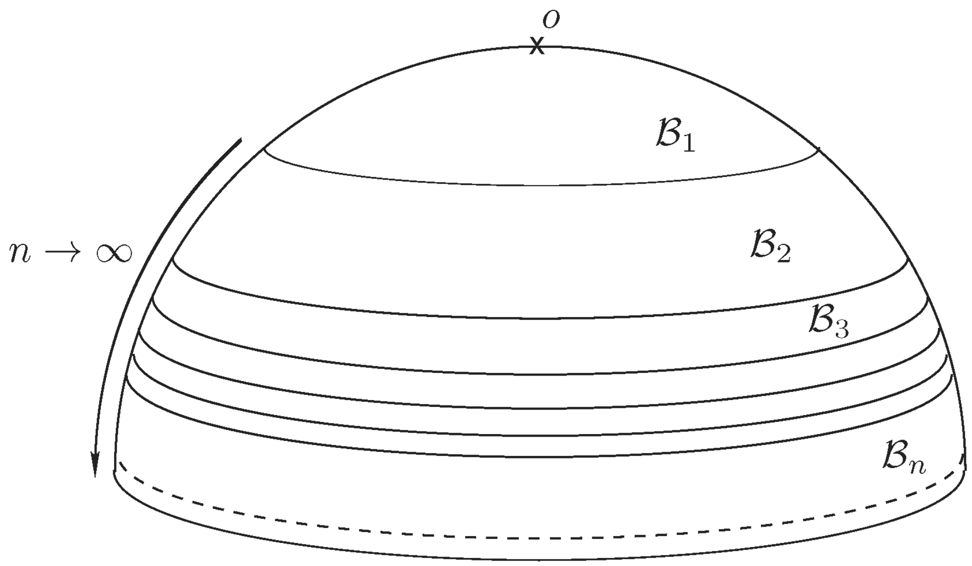

FIGURE 5.2: The open hemisphere is a CSLCG manifold.

Hence, the notion of CSLCG manifold is precisely the notion of maximal domain we need for the uniqueness of the exponential barycenter. Indeed, we can't have more than this as for example the upper (non-strict) hemisphere (for example of the Lie group SO(3)) doesn't represent a domain on which we have uniqueness. And the following theorem proves that we have uniqueness on a CSLCG manifold.

Theorem 5.18. *Let $\mathcal{M}$ be a CSLCG manifold. Then every probability measure μ on $\mathcal{M}$ with a compact support has a unique exponential barycenter.*

Remark 5.19. The proof of this theorem shows the remark below Theorem 5.15.

First we prove the following proposition, taken from [19]:

Proposition 5.20. *Any probability measure on $\mathcal{M}$ with a support relatively compact on the form $\phi < 0$ when ϕ is a convex $\mathcal{C}^1$ function has (at least) an exponential barycenter.*

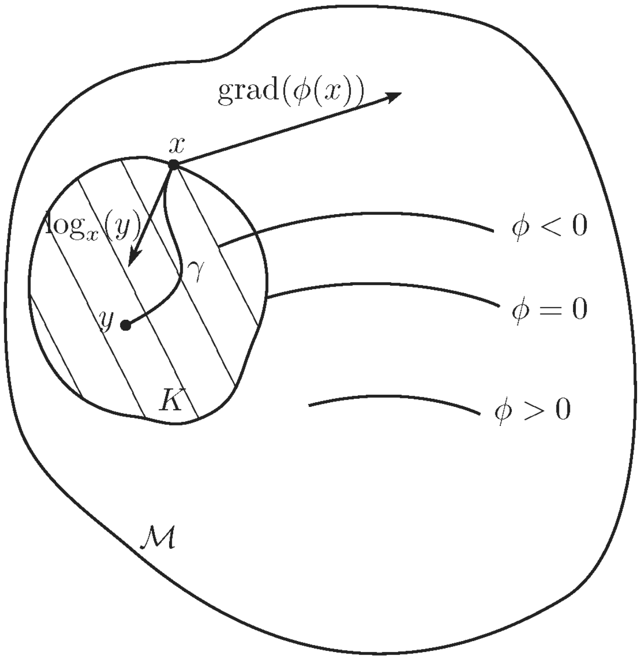

Figure 5.3: Existence of an exponential barycenter for a relatively compact domain of the form $\phi < 0$

Proof. Let $<,>$ be an auxiliary metric, independent of the affine connection ∇ defining the exponential. We denote K the compact K$=\phi \leq 0$. Let $x \in \partial K$ and $y \in K$. We have $\phi(x) = 0$ and $\phi(y) \leq 0$. Take γ the geodesic linking x and y. ϕ being convex, we have:

$$\begin{aligned} \phi(\gamma(t)) &\leq t\phi(x) + (1-t)\phi(y) \quad \forall t \in [0,1] \\ &= 0 + (1-t)\phi(y) \\ &\leq 0 \end{aligned}$$

Hence the geodesic γ is included in K. Now, differentiating at x, we get:

$$(*) \quad \frac{d}{dt}\Big|_{t=0} \phi(\gamma(t)) =< \exp_x^{-1}(y), \mathrm{grad}\phi(x) > \leq 0$$

because γ is going into $\mathring{K}$ while ϕ is decreasing from ∂K to K.

Now, we have:

$$| < \exp_x^{-1}(y), \mathrm{grad}\phi(x) > | \leq ||\mathrm{grad}\phi(x)||.|| \exp_x^{-1}(y)|| \leq c(K)$$

where c(K) is a positive constant (i.e. integrable on the compact K) depending only the compact K (i.e. independent of x). Moreover, the functions:

$$x \longmapsto < \exp_x^{-1}(y), \mathrm{grad}\phi(x) >$$

and:

$$y \longmapsto < \exp_x^{-1}(y), \mathrm{grad}\phi(x) >$$

are continuous. The previous majoration implies that the function which integrates inequality $(*)$:

$$\int_{y \in K} < \exp_x^{-1}(y), \mathrm{grad}\phi(x) > \mu(dy)$$

is well-defined and continuous. Moreover, it is negative because of $(*)$.
Hence, the vector field:

$$\int_y \exp_x^{-1}(y)\mu(dy)$$

is going inward K at any point on ∂K. Hence, it admits a zero in $\mathring{K}$ which defines an exponential barycenter for $\phi < 0$. $\square$

Now we need an additional lemma before going to the existence/uniqueness theorem.

Lemma 5.21. *Let $\mathcal{M}$ be a CSLCG manifold. Every compact convex subset K of $\mathcal{M}$ has a convex neighborhood U with a non-negative $\mathcal{C}^1$ convex function ϕ_K such that $\phi_K^{-1}(\{0\}) = K$.*

Proof. Let K be a compact convex subset of $\mathcal{M}$. As $\mathcal{M}$ is a CSLCG manifold, there exist a relatively compact convex neighborhood U_K of K, which has a p-convex geometry for some $p \in 2\mathbb{N}$. Fix U_K and take ϕ the corresponding separating function.

Define, for $x \in U_K$:

$$\phi_K(x) = \inf \{\phi(x, y), y \in K\}$$

i.e. the ϕ-distance from x to the compact K.

As K is compact, we have: $\phi_K^{-1}(\{0\}) = K$. Moreover, ϕ_K is obviously non-negative by non-negativity of ϕ.

We want to prove that ϕ_K has the properties given by the lemma, i.e. convexity and $\mathcal{C}^1$.

First, show that ϕ_K is convex on U_K. Take $x, x' \in U_K$, $y \in K$ and γ the unique geodesic linking x to x' by convexity of U_K. As ϕ is convex, we have:

$$(*) \quad \phi(\gamma(t), y) \leq t.\phi(x, y) + (1 - t)\phi(x', y) \quad \forall t \in [0, 1]$$

as $(\gamma(t), y(t))$ with $y(t) = y$, $\forall t$ is a geodesic of $U_K \times U_K$.

As $y \in K$, the definition of ϕ_K gives the following minorant for the left-hand side of $(*)$:

$$\phi_K(\gamma(t))$$

and hence is a minorant of the right-hand side:

$$\phi_K(\gamma(t)) \leq t.\phi(x, y) + (1 - t)\phi(x', y) \quad \forall t \in [0, 1].$$

But:

$$t.\phi_K(x)) + (1 - t)\phi_(x')$$

is the largest minorant of the right-hand side of $(*)$ by definition of the infimum. Hence:

$$\phi_K(\gamma(t)) \leq t.\phi_K(x)) + (1 - t)\phi_(x') \quad \forall t \in [0, 1$$

which proves the convexity of ϕ_K on U_K.

Next, show that ϕ_K is $\mathcal{C}^1$ on U_K. We already know that it is continuous as it is convex on the open set U_K [32].

We first prove that there exists a neighborhood $U \subset U_K$ of K, on which there is a unique point $p(x) \in K$ such that $\phi_K(x) = \phi(x, p(x))$ for each $x \in U$ as shown in Figure 5.4. This means that each $x \in U$ admits a unique ϕ-projection on the compact K. We choose U of the form:

$$U = \{x \in U_K, \phi_K(x) < \varepsilon\}$$

where we impose a condition on ε.

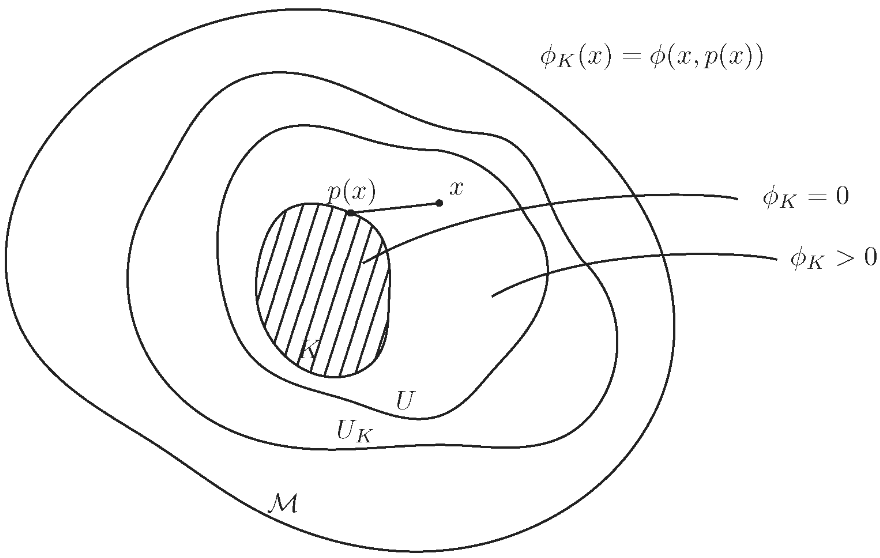

FIGURE 5.4: Definition of U so that each x has an unique ϕ-projection on K.

Take $x \in U_K$, and suppose there are two ϕ-projections for x on K, $y_1, y_2 \in K$:

$$\phi_K(x) = \phi(x, y_1) = \phi(x, y_2)$$

Then, by convexity of ϕ, the whole geodesic linking y_1 to y_2 in K (convex) will also verify: $\phi(x, y(t)) = \phi_K(x)$, $\forall t \in [0,1]$.

Now, by Hadamard lemma, we can write in the global chart:

$$\begin{aligned}\phi(x,y) =& a_{i_1 \dots i_p}(x) \prod_{j=1}^{p} (y^{i_j} - x^{i_j}) \\ &+ b_{i_1 \dots i_{p+1}}(x,y) \prod_{j=1}^{p+1} (y^{i_j} - x^{i_j})\end{aligned}$$

where all $a_{i_1 \dots i_p}, b_{i_1 \dots i_{p+1}}$ are smooth functions and p is the index of p-convexity of ϕ.

Consider the function $f(t) = \phi(x, y(t))$ which is a constant function equal to $\phi_K(x)$. We can differentiate it p times using the expression in the global chart. Taking into account that (y(t)) verifies the geodesic equation in the global chart:

$$\ddot{y}^i = -\Gamma^i_{jk} \dot{y}^j \dot{y}^k,$$

which implies that all its derivatives can be expressed in terms of $\dot{y}^i$, we have:

$$f^{(p)}(f) = a_{i_1 \dots i_p}(x) \prod_{j=1}^{p} (\dot{y}^{i_j} + g(y(t), \dot{y}(t), x),$$

where g is a smooth function: the product of $b_{i_1 \dots i_{p+1}}$ with power functions, composed with the smooth geodesic function $t \mapsto y(t)$ are smooths.

Now we minore $f^{(p)}$:

$$f^{(p)}(t) \geq a_{i_1 \dots i_p}(x) \prod_{j=1}^{p} \dot{y}^{i_j} - |g(y(t), \dot{y}(t), x)|$$

But the Hadamard lemma and explicit calculations give:

$$|g(y, x, z)| \leq C' ||y - x|| ||z||^p,$$

and by property of ϕ we have also:

$$a_{i_1 \dots i_p}(x) \prod_{j=1}^{p} z^{i_j} \geq c||z||^p.$$

Inserting this in the minoration of $f^{(p)}$, we get:

$$f^{(p)}(t) \geq c||\dot{y}^{i_j}||^p - C'||y(t) - x||.||\dot{y}(t)||^p.$$

Hence, if we choose $x \in U$ with ε such that $\varepsilon \in]0, \frac{c^{p+1}}{C'^p}[$ we have:

$$f^{(p)}(t) \geq (c - (\frac{\varepsilon}{c})^{\frac{1}{p}} C')||\dot{y}(t)||^p.$$

As $f^{(p)}(t) = 0$ and $c - (\frac{\varepsilon}{c} > 0$, we get $\dot{y}(t) = 0$, $\forall t$ and hence $y_1 = y_2$ for $x \in U$.

Now we focus only on the new neighborhood U, which will be the one given by the lemma, and show that the (well-defined) ϕ-projection p is continuous on U. Suppose it is not continuous. Let $(x_n)_n$ a convergent sequence with limit $x \in U$ such that $(p(x_n))_n$ doest not converge to $p(x)$. Since $(p(x_n))_n$ is a sequence of K compact we can assume:

$$p(x_n) \to y \in K \setminus \{p(x)\},$$

choosing a subsequence if necessary.

By continuity of ϕ, we have:

$$\phi_K(x_n) = \phi(x_n, p(x_n)) \to \phi(x, y).$$

On the other hand, ϕ_K is itself continuous: $\phi_K(x_n) \to \phi_K(x)$. By the uniqueness of the ϕ-projection in U, we get $y = p(x)$, which is a contradiction. Hence p is continuous.

Finally we can prove that ϕ_K is $\mathcal{C}^1$. Denote by $d_g\phi_K$ the Gâteaux-differential of ϕ and recall that, for any $x \in U$ and $v \in T_x\mathcal{M}$:

$$d_g\phi_K(x).v = \lim_{t \downarrow 0} \frac{\phi_K(\exp_x(tv) - \phi_K(x)}{t}$$

and that $d_g\phi_K(x)$ is convex on $T_x\mathcal{M}$, [32]. Take $x \in U$ and $v \in T_x\mathcal{M}$. If t is small enough, we have:

$$\begin{aligned}\phi_K(\exp_x tv) - \phi_K(x) &= \phi(\exp tv, p(\exp_x tv)) - \phi(x, p(x)) \\ &\leq \phi(\exp_x tv, p(x)) - \phi(x, p(x)).\end{aligned}$$

This implies:

$$d_g\phi_K(x).v \leq D\phi_{p(x)}(x).v$$

where ϕ_y denotes the map $\phi(\bullet, y)$. Since $d_g\phi_K(x)$ is convex and $D\phi_{p(x)}(x)$ is linear, we have in fact an equality:

$$d_g\phi_K(x) = D\phi_{p(x)}(x) \quad \text{on } T_x\mathcal{M}.$$

Now the differentiability of ϕ_k comes from:

$$0 \leq \phi_K(\exp_x v) - \phi_K(x) - d\phi_{p(x)}(x).v \leq \phi_{p(x)}(\exp_x v) - \phi_{p(x)}(x) - d\phi_{p(x)}(x).v$$

and the fact that the right-hand side goes to 0 when $v \to 0$. The $\mathcal{C}^1$ property finally comes from the continuity of p as ϕ is smooth.

Hence we have constructed a open convex neighborhood U of K and a funtion ϕ_K on it with the desired properties given by the lemma. $\square$

We are now able to prove Theorem 5.18, using the same kind of arguments as in the proof of Theorem 5.15.

Proof. (Proof of Theorem 5.18) First show the uniqueness, which is only a slight modification of the Riemannian CoM's case. Take m and m' two exponential barycenters and the compact $K = \text{supp}(\mu) \bigcup \{m, m'\}$. Then (m, m') is also an exponential barycenter of the image of μ under the diagonal map : $\delta : \mathcal{B} \longmapsto \mathcal{B} \times \mathcal{B}$. As $\mathcal{M}$ is CSLCG, K has a relatively compact convex neighborhood U_K which has p-convex geometr for some $p \in 2\mathbb{N}$. Take ϕ the corresponding separating function, with d the auxiliary Riemannian distance.

We have:

$$\begin{aligned}
d^p(m, m') &\leq \frac{1}{c}\phi(m, m') \quad \text{by p-convex geometry of } U_K \\
&\leq \frac{1}{c}\int_{U_K} \phi(x, x')d\mu(x, x') \quad \text{using 5.14} \\
&= \frac{1}{c}\int_{U_K} \phi(x, x)d\mu(x) \quad \text{because } d\mu(x, x') = 0 \text{ if } x \neq x' \\
&= 0 \quad \text{as } \phi \text{ is vanishing on the diagonal}
\end{aligned}$$

and this gives $m = m'$ and uniqueness of the exponential barycenter.

Now, show the existence of the exponential barycenter. Take K a convex compact subset of $\mathcal{M}$. Let ϕ_K be the $\mathcal{C}^1$ non-negative convex function given by lemma 5.21. ϕ_K is defined on a relatively compact open neighborhood U_K of K, such that $\phi_K^{-1}(\{0\}) = K$.

Let $\varepsilon > 0$ satisfy $\phi_K^{-1}([0, \varepsilon[) \subset U$. We apply 5.20 to the function $\phi_k - \frac{\varepsilon}{2}$. Hence, μ has an exponential barycenter in $\phi_K^{-1}([0, \frac{\varepsilon}{2}[)$, which shows the existence. $\square$

Hence we have found the maximal domain of uniqueness for the exponential barycenters in affine connection spaces. What remains if we add the algebraic properties of the Lie group?

5.2.2 Application to Lie groups: some insights on the uniqueness domain

Now we want to construct the maximal domain where a Lie group is CSLCG. In the Riemannian case, recall that it corresponded to the union of interiors of its regular geodesic balls, which were themselves characterized by algebraic properties of the Lie group. Can we find an equivalent domain for the affine connection case? Our conjecture is yes and we provide some insights about it in this last subsection.

Hence, consider $\mathcal{G}$ as a Lie group with the canonical Cartan-Schouten connection, i.e. as an affine connection space. Take an auxiliary Riemannian metric $<,>_{\text{aux}}$. Any regular geodesic ball of this auxiliary metric will have p-convex geometry. Then we can consider the union of interiors of regular geodesic balls as in the Riemannian case as an attempt for the maximal domain.

However, as $<,>_{\text{aux}}$ is only an *auxiliary* metric, we need this union to be independent of $<,>_{\text{aux}}$. This might be not trivial as reguler geodesic balls are defined by condition on the sectionnal curvature and the injectivity radius, both depending on the metric.

But we could think about taking the union of all regular geodesic balls of *all possible auxiliary metrics* (in a more rigorous sense, to be defined). However, we didn't go further in that direction in the context of this thesis and we leave this issue to the reader.

Hence, in this chapter we investigated the existence and uniqueness property of the admissible mean's definitions. We conclude about the case of the Riemannian CoM and illustrated it with the example of SO(3). We gave a generalization of the approach for affine connection spaces, defining the central notion of CSLCG manifolds. This should enable to conclude about all Lie groups, following the ideas we gave in the last subsection.

Chapter 6

Conclusion

This thesis focuses on the definition of the mean for curved spaces, in particular for Lie groups. The Riemannian center of mass was already used in the literatur in order to define the mean on *manifolds*. As a Lie group is a special case of manifold, we imported this definition and defined two extensions of it: the Riemannian exponential barycenter and the group exponential barycenter.

These three definitions need special geometric structures for the Lie group, which we introduced in Chapter 2. Then, we investigated the compatibility of all definitions, and more precisely: if two or three of them are defined on the Lie group, do they give the same result for the mean? This led us to the comparison of the geodesics of the corresponding geometric structures and the definition of Cartan-Schouten connections, which we illustrated with the example of SE(3). Then we turn to the consistency of these means with the *Lie group's algebraic structure*, which is not trivial as we imported a mean's definition from a *manifold* structure. It led to the characterization of the existence of bi-invariant metrics or pseudo-metrics on Lie groups. Indeed, all Lie groups don't admit a bi-invariant metric or a bi-invariant pseudo-metric. Lie groups that can be provided with a bi-invariant metric are exactly the one whose adjoint group is relatively compact. We gave the example of SO(3) for this case. Lie groups that can be provided with a bi-invariant pseudo-metric are the one which admit a special decomposition, which was the topic of Chapter 4. We gave the example of SE(3) for this case. Finally, when the mean's definitions are well defined, compatible with each other, and consistent with the Lie group's structure, we investigated their existence and uniqueness (Chapter 5). We presented the well-studied case of the Riemannian CoM as an example and showed that it exists and is unique on any regular geodesic ball of the manifold. We gave some insights about the regular geodesic balls of a Lie group provided with a bi-invariant metric and studied SO(3) as an example. Lastly, we extended the properties of existence and uniqueness in the case of affine connection spaces for the exponential barycenters and we gave some ideas about how to apply it for Lie groups.

The thesis discussed means on Lie groups, but mainly on *finite dimensional* Lie groups. However, the possible applications of this setting in statistics might use *infinite dimensional* Lie groups, as for example the general group of diffeomorphisms. We might wonder what will be still valid in our approach. For example, we don't have any Inverse Function Theorem anymore nor the Riesz

Theorem i.e. not necessary the existence of adjoint (in the sense of duality) of endomorphisms. Hence, it is possible that a whole new setting for the mean on infinite dimensional Lie groups would have to be defined in the infinite dimension case.

Appendix A

Scilab code

This scilab code has provided the figures of Section 3.3.

```
//################### SE(3) ###################


//####### Preliminaries

//Computation of the closest rotation matrix R of a given matrix M (avoids
    computational errors.)
function [R]=Rot(M)
    [u,s,v]=svd(M);
    R=u*v';
    if det(R)<0 then
        s(1,1)=1;
        s(2,2)=1;
        s(3,3)=-1;
        R=u*s*v';
    end
endfunction

//Test if a matrix is skewsymmetric
function testskew(A)
    Err=A+A';
    err=trace(Err*Err');
    if err>1e-12 then
        disp("warning: not a skew-symmetric matrix");
    end
endfunction
```

```
//############# SE(3): principal chart f=(r,t), vector of R^6 ()unfolding along
     the L Riemannian geodesics)
//r= (theta,n) where theta[%pi]

// Conversion of the elements f of the principal chart to matrices +
    regularization functions.

function [S]=Skew(r)
    S=[[0, -r(3), r(2)];
      [r(3), 0, -r(1)];
      [-r(2), r(1), 0]];
endfunction

function [u]=regrot(r)
    phi=norm(r);
    u=r;
    if phi<>0 then
        k0=int(phi/(2*%pi)+1/2);
        u=(phi-2*%pi*k0)*r/phi;
        end
endfunction

function [ff]=reg(f)
    ff=zeros(6,1);
    r=f(1:3,1);
    t=f(4:6,1);
    r=regrot(r);
    ff(1:3)=r;
    ff(4:6)=t;
endfunction

function [r]=RotVect(R)
    M=R*R'-eye(3,3);
    if trace(M)>1e-20 then R=Rot(R); end
    c=(trace(R)-1)/2;
    if c>1 then c=1; end
    if c<-1 then c=-1; end
    theta=acos(c);

    if theta<1e-5 then
        fact=0.5*(1+theta^2/6);
        Sr=fact*(R-R');
        r=[Sr(3,2), Sr(1,3), Sr(2,1)]';
    elseif abs(theta-%pi)<1e-5 then
```

```
  for i=1:3
          sq=1+(R(i,i)-1)/(1-c);
          if (sq<0) sq=0; end
          if (sq>1) sq=1; end
          r(i)=sqrt(sq);
      end
      //Normalize rotation vector
      Norm_r=norm(r);
      r=r*theta/Norm_r;
      //set signs with off-diagonal terms of n.n^T
      if (R(1,2)+R(2,1)<0) r(2)=-r(2); end
      if (R(1,3)+R(3,1)<0) r(3)=-r(3); end

      //Determine wether r=+/- theta*n
      sin_r(1)=(R(3,2)-R(2,3));
      sin_r(2)=(R(1,3)-R(3,1));
      sin_r(3)=(R(2,1)-R(1,2));
      //Determine the most significant term
      k=1;
      if (abs(sin_r(2))>abs(sin_r(k))) k=2; end
      if (abs(sin_r(3))>abs(sin_r(k))) k=3; end
      //Choose the sign
      if (sin_r(k)*r(k)<0) r=-r;
    else
      fact=0.5*theta/sin(theta);
      Sr=fact*(R-R');
      r=[Sr(3,2), Sr(1,3), Sr(2,1)]';
    end
endfunction

function [R]=RotMat(r)
    r=regrot(r);
    theta=norm(r);
    //disp("theta=");disp(theta);
    Sr=Skew(r);
    if theta<1e-5 then
        s=1-theta^2/6;
        k=1/2-theta^2;
        R=eye(3,3)+s*Sr+k*Sr^2;
    else
    R=eye(3,3)+(sin(theta)/theta)*Sr+(1-cos(theta))/(theta^2)*Sr^2;
    end
endfunction
```

```
//Differentials of the left and right translations for SO(3) in the principal
    chart

function [Jl]=Jrot_L(r)
    r=regrot(r);
    theta=norm(r);
    if theta<1e-5 then
        phi=1-(theta^2)/12;
        w=1/12+theta^2/720;
    elseif abs((theta-%pi))<1e-5 then
        phi=theta*(%pi-theta)/4;
        w=(1-phi)/theta^2;
    else phi=(theta/2)/(tan(theta/2));
        w=(1-phi)/theta^2;
        end,
Jl=phi*eye(3,3)+w*r*r'+Skew(r)/2;
endfunction


function [Jl]=Jrot_R(r)
    r=regrot(r);
    theta=norm(r);
    if theta<1e-5 then
        phi=1-(theta^2)/12;
        w=1/12+theta^2/720;
    elseif abs((theta-%pi))<1e-5 then
        phi=theta*(%pi-theta)/4;
        w=(1-phi)/theta^2;
    else phi=(theta/2)/(tan(theta/2));
        w=(1-phi)/theta^2;
        end,
Jl=phi*eye(3,3)+w*r*r'-Skew(r)/2;
endfunction

//Differential of left and right translations for SE(3) in the principal chart

function [Jl]=J_L(f)
    f=reg(f);
    Jl(1:3,1:3)=Jrot_L(f(1:3));
    Jl(4:6,4:6)=RotMat(f(1:3));
endfunction

function [Jr]=J_R(f)
    Jr(1:3,1:3)=Jrot_R(f(1:3));
```

```
    Jr(4:6,1:3)=-Skew(f(4:6));
    Jr(4:6,4:6)=eye(3,3);
endfunction

//Differential of inversion for SE(3)

function [J]=Jinv(f)
    f=reg(f);
    r=f(1:3);
    R=RotMat(r);
    Sr=Skew(r);
    theta=norm(r);
    if theta<1e-5 then
        betA=1/2-theta^2/24;
        etA=1/6-theta^2/120;
    else
    betA=(1-cos(theta))/theta^2;
    etA=1/theta^2*(1-sin(theta)/theta);
end,

    U=etA*Sr^2-betA*Sr+eye(3,3);
    J(1:3,1:3)=-eye(3,3);
    J(4:6,4:6)=-R';
    J(4:6,1:3)=-Skew(R'*t)*U;
endfunction

//Group composition for SE(3), new operator: *.
function [u]=ajout3zeros(v,n)
    u=zeros(6,1);
    if n==0 then u(4:6,1)=v; end
    if n==1 then u(1:3,1)=v; end
endfunction


deff('x=%s_u_s(a,b)','x=ajout3zeros(RotVect(RotMat(a(1:3))*RotMat(b(1:3))),1)+
    ajout3zeros(RotMat(a(1:3))*b(4:6)+a(4:6),0)');

//Group inversion for SE(3)
function [in]=INV(a)
    a=reg(a);
    r=a(1:3);
    theta=norm(r);
    t=a(4:6);
    in=ajout3zeros(-r,1)+ajout3zeros((RotMat(-r))*(-t),0);
```

```
endfunction
//deff('x=%s_5(a)','x=ajout3zeros(RotVect(RotMat(a(1:3))^(-1)),1)+ajout3zeros
    (-(RotMat(a(1:3))^(-1))*a(4:6),0)');


//Left- and right- invariant inner product in the principal chart (propagation
    of Frobenius inner product)
function [g]=Q_L(a,f)
    g0=zeros(6,6);
    g0(1:3,1:3)=eye(3,3);
    g0(4:6,4:6)=a*eye(3,3);
    g=J_L(f)'^(-1)*g0*J_L(f)^(-1);
endfunction

function [g]=Q_R(a,f)
    g0=zeros(6,6);
    g0(1:3,1:3)=eye(3,3);
    g0(4:6,4:6)=a*eye(3,3);
    g=J_R(f)'^(-1)*g0*J_R(f)^(-1);
endfunction

//Associated left- and right- invariant metrics in the principal chart
function [n]=norma2_L(a,f,v)
    v=reg(v);
    n=v'*Q_L(a,f)*v;
endfunction

function [n]=norma2_R(a,f,v)
    v=reg(v);
    n=v'*Q_R(a,f)*v;
endfunction

//Group exponential and logarithm from Id

function [f]=ExpId(v)
    v=reg(v);
    r=v(1:3);
    dt=v(4:6);
    theta=norm(r);
    f(1:3)=r;
    Sr=Skew(r);
    if theta==0 then
        f(4:6)=dt;
```

```
    elseif theta<1e-5 then f(4:6)=dt+(1/6-theta^3/120)*Sr^2*dt+(1/2-theta^2/24)
    *Sr*dt;
    else
        f(4:6)=dt+theta^(-2)*(1-sin(theta)/theta)*Sr^2*dt+theta^(-2)*(1-cos(
    theta))*Sr*dt;
    end
endfunction


function [v]=LogId(f)
    f=reg(f);
    r=f(1:3);
    t=f(4:6);
    theta=norm(r);
    v(1:3)=r;
    Sr=Skew(r);
    if theta==0 then v(4:6)=t;
    elseif theta<1e-5 then v(4:6)=t-1/2*Sr*t+(1/2-theta^2/90)*Sr^2*t;
    else
        fact=0.5*theta*sin(theta)/(1-cos(theta));
        v(4:6)=t-1/2*Sr*t+(1-fact)/theta^2*Sr^2*t;
    end
endfunction

//Group exponential and logarithm from any point f (first for SO(3))
//-> Group geodesics are left- and right- invariant

function [rr]=Exprot_L(r,a)
    rr=RotVect(RotMat(r)*RotMat(inv(Jrot_L(r))*a));// R*Exp(DL(R^-1)*a)
endfunction

function [a]=Logrot_L(r,rr)
    a=Jrot_L(r)*RotVect((RotMat(-r)*RotMat(rr)));
endfunction

function [ff]=Exp_L(f,v)
    v=reg(v);
    f=reg(f);
    ff=f*. ExpId(inv(J_L(f))*v);
endfunction

function [v]=Log_L(f,ff)
    ff=reg(ff);
    f=reg(f);
```

```
    v=J_L(f)*LogId(INV(f)*. ff);
endfunction


//Riemannian exponential and logarithm from Id (for left- and right-invariant
    metric)
function [f]=Riem_L_ExpId(a,v)
    v=reg(v);
    f(1:3)=v(1:3);
    f(4:6)=a*v(4:6);
endfunction

function [f]=Riem_R_ExpId(a,v)
    f=INV(Riem_L_ExpId(a,-v));
endfunction

function [v]=Riem_L_LogId(a,f)
    v(1:3)=f(1:3);
    if a<>0 then v(4:6)=1/a*f(4:6);
    else disp("impossible: alpha=0");
    end
endfunction

function [v]=Riem_R_LogId(a,f)
    r=f(1:3);
    v(1:3)=r;
    v(4:6)=RotMat(-r)*f(4:6);
endfunction


//Riemannian exponential and logarithm from any point f0 (for left- and right-
    invariant metric)
function [f]=Riem_L_Exp(a,f0,v)
    f=f0 *. Riem_L_ExpId(a,inv(J_L(f0))*v);//left-invariance of the left
    Riemannian geodesics
endfunction

function [f]=Riem_R_Exp(a,f0,v)
    f=Riem_R_ExpId(a,inv(J_R(f0))*v)*. f0;//right-invariance of the right
    Riemannian geodesics
endfunction


function [v]=Riem_L_Log(a,f0,f)
```

```
    v=J_L(f0)*Riem_L_LogId(a,INV(f0)*. f);
endfunction

function [v]=Riem_R_Log(a,f0,f)
    v=J_R(f0)*Riem_R_LogId(a,f*. INV(f0));
endfunction


//########## Statistical setting on SE(3) ############

//Generation of a random element of SE(3)

function [f]= UnifRnd()
    M=(rand(3,3)-0.5*ones(3,3))*2;
    t=(rand(3,1)-0.5)*2;
    r=regrot(RotVect(RotMat(M)));
    f(1:3)=r;
    f(4:6)=t;
endfunction

//Computations of the variance in the 3 geometric frameworks: Group, L and R

function [s]=sigma2(m,tabr,tabw)
    siz=size(tabr(1,:),"*");
    s=0;
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    for (i=1:siz)
        s=s+tabw(i)*norm(Logrot_L(m,tabr(:,i)))^2;
    end
endfunction

function [s]=sigma2_L(a,m,tabf,tabw)
    siz=size(tabf(1,:),"*");
    s=0;
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    for (i=1:siz)
        s=s+tabw(i)*norma2_L(a,m,Riem_L_Log(a,m,tabf(:,i)));
    end
endfunction

function [s]=sigma2_R(a,m,tabf,tabw)
    siz=size(tabf(1,:),"*");
```

```
    s=0;
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    for (i=1:siz)
        s=s+tabw(i)*norma2_R(a,m,Riem_R_Log(a,m,tabf(:,i)));
    end
endfunction

//Preliminaries: computation of the mean for rotations

function [m]=RotMean(tabr,tabw)
    siz=size(tabr(1,:),"*");
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    m=tabr(:,1);//mbis=mt et m=m(t+1)
    aux=zeros(6,1);
    first=0;
    while first==0 | (norm(Logrot_L(mbis,m))^2>1e-5*sigma2(mbis,tabr,tabw))
        mbis=m;
        aux=zeros(3,1);
        for (i=1:siz)
            aux=aux+tabw(i)*Logrot_L(mbis,tabr(:,i));
        end
        m=Exprot_L(mbis,aux);
        first=1;
    end
endfunction

//Group exponential barycenter on SE(3):
// - the previous mean for the rotation part,
// - closed form from article of Xavier Pennec "Exponential Barycenters of
    ...",(2012)

//Intermediate function
function [M]=MatdeExp(v)
    v=regrot(v);
    r=v(1:3);
    theta=norm(r);
    Sr=Skew(r);
    if theta==0 then
        M=eye(3,3);
    elseif theta<1e-5 then M=eye(3,3)+(1/6-theta^3/120)*Sr^2+(1/2-theta^2/24)*
    Sr;
```

```
    else M=eye(3,3)+theta^(-2)*(1-sin(theta)/theta)*Sr^2+theta^(-2)*(1-cos(
    theta))*Sr;
    end
endfunction

function [m]=expbar(tabf,tabw)
    siz=size(tabf(1,:),"*");
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    for (i=1:siz)
        f=tabf(:,i);
        tabr(1:3,i)=f(1:3);
        tabt(1:3,i)=f(4:6);
    end
    rmean=RotMean(tabr,tabw);
    //Partie translation, p34 de expbar
    M=zeros(3,3);
    t=zeros(3,1);
    for (i=1:siz)
        Maux=inv(MatdeExp(RotVect(RotMat(rmean)*RotMat(-tabr(:,i)))));
        M=M+tabw(i)*Maux;
        t=t+tabw(i)*Maux*RotMat(-tabr(:,i))*tabt(:,i);
    end
    m(1:3,1)=rmean;
    m(4:6,1)=inv(M)*t;
endfunction

//Riemannian center of mass for the left-invariant metric
function [m]=Frechet_L(a,tabf,tabw)
    siz=size(tabf(1,:),"*");
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    m=tabf(:,1);
    aux=zeros(6,1);
    first=0;
    while first==0 | norma2_L(a,mbis,Riem_L_Log(a,mbis,m))>1e-5*sigma2_L(a,mbis
    ,tabf,tabw)//convergence test
        mbis=m;
        disp("mbisL=");disp(mbis);
        aux=zeros(6,1);
        for (i=1:siz)
            aux=aux+tabw(i)*Riem_L_Log(a,mbis,tabf(:,i));
        end
        m=Riem_L_Exp(a,mbis,aux);
```

```
        first=1;
    end
endfunction

//RIemannian center of mass for the right-invariant metric
function [m]=Frechet_R(a,tabf,tabw)
    siz=size(tabf(1,:),"*");
    if siz <> size(tabw,"*") then disp("longueurs de tableaux incompatibles");
    end
    m=tabf(:,1);
    aux=zeros(6,1);
    first=0;
    while first==0 | norma2_R(a,mbis,Riem_R_Log(a,mbis,m))>1e-5*sigma2_R(a,mbis
    ,tabf,tabw);//convergence test
        mbis=m;
        disp("mbisR=");disp(mbis);
        aux=zeros(6,1);
        for (i=1:siz)
            aux=aux+tabw(i)*Riem_R_Log(a,mbis,tabf(:,i));
        end
        m=Riem_R_Exp(a,mbis,aux);
        first=1;
    end
endfunction

//########## Displays

//Plot a 3d vector between p0 and p1
function vectarrow(p0,p1,col)
      if max(size(p0))==3
      if max(size(p1))==3
          x0 = p0(1);
          y0 = p0(2);
          z0 = p0(3);
          x1 = p1(1);
          y1 = p1(2);
          z1 = p1(3);
          param3d1([x0;x1],[y0;y1],list([z0;z1],col));   // Draw a line between
     p0 and p1

          p = p1-p0;
          alpha = 0.1;  // Size of arrow head relative to the length of the
    vector
```

```
        bet = 0.1;  // Width of the base of the arrow head relative to the
    length

        hu = [x1-alpha*(p(1)+bet*(p(2)+%eps)); x1; x1-alpha*(p(1)-bet*(p(2)+
    %eps))];
        hv = [y1-alpha*(p(2)-bet*(p(1)+%eps)); y1; y1-alpha*(p(2)+bet*(p(1)+
    %eps))];
        hw = [z1-alpha*p(3);z1;z1-alpha*p(3)];

        set(gca(),"auto_clear","off");//hold on: retains the current graphe
    and add something to it
        plot3d(hu(:),hv(:),list(hw(:),col)) ; // Plot arrow head
        set(gca(),"grid",[1 1]);
        xlabel('x');
        ylabel('y');
        zlabel('z');
        set(gca(),"auto_clear","on");
      else
        error('p0 and p1 must have the same dimension')
      end
  else
      error('this function only accepts 3D vector')
  end
endfunction

//Plot the trihedron representing f of SE(3)
function plot_trihedron(f,col)
    a=f(4:6); //origin of the frame
    R=Rot(RotMat(f(1:3)));
    e1=R*([1 0 0]');//basis vectors of the frame
    e2=R*([0 1 0]');
    e3=R*([0 0 1]');
    vectarrow(a,a+e1,col);//translation of vector a
    set(gca(),"auto_clear","off");
    vectarrow(a,a+e2,col);
    set(gca(),"auto_clear","off");
    vectarrow(a,a+e3,col);
    set(gca(),"auto_clear","on");
endfunction

//Plot group, L and R geodesics starting at f with initial tangent vector v
function plot_comparison(a,f,v,fin,pas)
        for s=0:pas:fin
        fs1=Exp_L(f,s*v);
```

```
        plot_trihedron(fs1,5);//red
        set(gca(),"auto_clear","off");
        fs2=Riem_L_Exp(a,f,s*v);
        plot_trihedron(fs2,3);//green
        set(gca(),"auto_clear","off");
        fs3=Riem_R_Exp(a,f,s*v);
        plot_trihedron(fs3,2);//bleu
        set(gca(),"auto_clear","off");
        end
endfunction

//Plot group, L and R geodesics linking points f1 and f2
function plot_linkinggeodesics(a,f1,f2)
    u=Log_L(f1,f2);
    uL=Riem_L_Log(1,f1,f2);
    uR=Riem_R_Log(1,f1,f2);
    for s=0:0.1:1
        fs=Exp_L(f1,s*u);
        plot_trihedron(fs,5);
        set(gca(),"auto_clear","off");
        fsL=Riem_L_Exp(1,f1,s*uL);
        plot_trihedron(fsL,3);
        set(gca(),"auto_clear","off");
        fsR=Riem_R_Exp(1,f1,s*uR);
        plot_trihedron(fsR,2);
        set(gca(),"auto_clear","off");
    end
endfunction

//Plot the group exponential barycenter and the L and R Riemannian centers of
    mass
function plot_comparison_mean(tabf,tabw)
    wsum=sum(tabw);
    if wsum<>1.0 then
        disp("weights dont sum to 1");
    end
    siz=size(tabf(1,:),"*");
    for i=1:siz
        plot_trihedron(tabf(:,i),1);
        set(gca(),"auto_clear","off");
    end
        m1=expbar(tabf,tabw);
        plot_trihedron(m1,5);
        set(gca(),"auto_clear","off");
```

```
        m2=Frechet_L(1,tabf,tabw);
        plot_trihedron(m2,3);
        set(gca(),"auto_clear","off");
        m3=Frechet_R(1,tabf,tabw);
        plot_trihedron(m3,2);
        set(gca(),"auto_clear","off");
endfunction
```

Bibliography

[1] D.A.W. Thompson and J.T. Bonner. *On Growth and Form*. Canto (Cambridge University Press). Cambridge University Press, 1992. ISBN 9780521437769. URL http://books.google.co.uk/books?id=7_F40UJmLFcC.

[2] U. Grenander. *General pattern theory: a mathematical study of regular structures*. Oxford mathematical monographs. Clarendon, 1993. URL http://books.google.co.uk/books?id=Z-8YAQAAIAAJ.

[3] M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41:61–84, 2001.

[4] L. Younes. *Shapes and Diffeomorphisms*. Applied Mathematical Sciences. Springer London, Limited, 2012. ISBN 9783642263484. URL http://books.google.fr/books?id=WEFRMAEACAAJ.

[5] X. Pennec. *L'INCERTITUDE DANS LES PROBLEMES DE RECONNAISSANCE ET DE RECALAGE. APPLICATION EN IMAGERIE MEDICALE ET BIOLOGIE MOLECULAIRE*. INRIA, 1996. ISBN 9782726110690. URL http://books.google.fr/books?id=65bDOQAACAAJ.

[6] Xavier Pennec. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006. ISSN 0924-9907. doi: 10.1007/s10851-006-6228-4. URL http://dx.doi.org/10.1007/s10851-006-6228-4.

[7] Xavier Pennec and Vincent Arsigny. Exponential Barycenters of the Canonical Cartan Connection and Invariant Means on Lie Groups. In Frederic Barbaresco, Amit Mishra, and Frank Nielsen, editors, *Matrix Information Geometry*, pages 123–168. Springer, May 2012. ISBN 978-3-642-30231-2 (Print) / 978-3-642-30232-9 (Online). doi: 10.1007/978-3-642-30232-9. URL http://hal.inria.fr/hal-00699361.

[8] Xavier Pennec. *Statistical Computing on Manifolds for Computational Anatomy*. Habilitation à diriger des recherches, Université Nice Sophia-Antipolis, December 2006. URL http://www.inria.fr/sophia/asclepios/Publications/Xavier.Pennec/Pennec_HDR.pdf.

[9] M. Nakahara. *Geometry, Topology, and Physics*. Graduate student series in physics. Institute of Physics Publishing, 2003. ISBN 9780750306065. URL http://books.google.co.uk/books?id=cH-XQB0Ex5wC.

[10] M. Postnikov. *Geometry VI: Riemannian Geometry*. Encyclopaedia of Mathematical Sciences. Springer, 2001. ISBN 9783540411086. URL http://books.google.co.uk/books?id=1kyUImXf8U0C.

[11] P. Colmez. *Eléments d'analyse et d'algèbre (et de théorie des nombres)*. Ecole Polytechnique, 2009. ISBN 9782730215633. URL http://books.google.fr/books?id=hMUiEhsa9WUC.

[12] Michael Wüstner. A connected lie group equals the square of the exponential image. *Journal of Lie Theory*, 13(1):307–309, 2003. URL http://eudml.org/doc/122929.

[13] C. Chevalley. *Théorie des groupes de Lie: II-III*. Actualités scientifiques et industrielles. Hermann, 1951. URL http://books.google.fr/books?id=mfc_AQAAIAAJ.

[14] Vincent Arsigny. *Processing Data in Lie Groups: An Algebraic Approach. Application to Non-Linear Registration and Diffusion Tensor MRI*. Thèse de sciences (phd thesis), École polytechnique, November 2006. URL http://www.inria.fr/sophia/asclepios/Publications/Arsigny/arsigny_phd.pdf.

[15] M. Fréchet. *L'intégrale abstraite d'une fonction abstraite d'une variable abstraite et son application a la moyenne d'un élément aléatoire de nature quelconque*. La Revue Scientifique, 1944. URL http://books.google.co.uk/books?id=9DjhHAAACAAJ.

[16] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré*, 10(4):215–310, 1948. URL http://eudml.org/doc/79021.

[17] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977. ISSN 1097-0312. doi: 10.1002/cpa.3160300502. URL http://dx.doi.org/10.1002/cpa.3160300502.

[18] Wilfrid S. Kendall. Probability, convexity, and harmonic maps with small image i: Uniqueness and fine existence. *Proceedings of the London Mathematical Society*, s3-61(2):371–406, 1990. doi: 10.1112/plms/s3-61.2.371. URL http://plms.oxfordjournals.org/content/s3-61/2/371.abstract.

[19] M. Émery and G. Mokobodzki. Sur le barycentre d'une probabilité dans une variété. *Séminaire de probabilités de Strasbourg*, 25:220–233, 1991. URL http://eudml.org/doc/113758.

[20] Boris Kolev. Groupes de lie et mécanique, 2007.

[21] Huiling Le. Locating Fréchet means with application to shape spaces. *Adv. Appl. Probab.*, 33(2):324–338, 2001. ISSN 0001-8678. doi: 10.1239/aap/999188316.

[22] Huiling Le. Estimation of Riemannian barycenters. *LMS J. Comput. Math.*, 2004.

[23] Shlomo Sternberg. Lectures on differential geometry, 1964. URL http://opac.inria.fr/record=b1078755. Autre tirage : 1965.

[24] N. Jacobson. *Lie Algebras*. Dover Books on Mathematics Series. Dover, 1979. ISBN 9780486638324. URL http://books.google.fr/books?id=hPE1Mmm7SFMC.

[25] E. Levi. Sulla struttura dei gruppi finiti e continui. *Torino Atti*, 40:551–565, 1905.

[26] S. Bromberg and A. Medina. Geometrie des groupes oscillateurs et varietes localement symetriques. *ArXiv Mathematics e-prints*, April 2002.

[27] A. Medina. Groupes de Lie munis de pseudo-métriques de Riemann bi-invariantes. Sémin. géométrie différentielle 1981-1982, Montepellier 1982, Exp. No.6, 37 p. (1982)., 1982.

[28] A. Medina and P. Revoy. Algèbres de lie et produit scalaire invariant. *Annales scientifiques de l'École Normale Supérieure*, 18(3):553–561, 1985. URL http://eudml.org/doc/82165.

[29] M.P. do Carmo. *Riemannian Geometry*. Mathematics (Boston, Mass.). Birkhauser Verlag AG, 1992. ISBN 9783764334901. URL http://books.google.fr/books?id=uXJQQgAACAAJ.

[30] Willi Jager and Helmut Kaul. Uniqueness and stability of harmonic maps and their jacobi fields. *manuscripta mathematica*, 28(1-3):269–291, 1979. ISSN 0025-2611. doi: 10.1007/BF01647975. URL http://dx.doi.org/10.1007/BF01647975.

[31] Marc Arnaudon and Xue-Mei Li. Barycenters of measures transported by stochastic flows. *Ann. Probab.*, 33(4):1509–1543, 2005. ISSN 0091-1798. doi: 10.1214/009117905000000071.

[32] M. Émery and W.A. Zheng. *Fonctions convexes et semimartingales dans une variété*. Institut de Recherche Mathématique Avancée Strasbourg: Publication de l'. Institut de Recherche Mathematique Avancée, 1983. URL http://books.google.fr/books?id=BvEvcgAACAAJ.