Sebastiano Ferraris

Towards a BCH-Free Computation of the Composition of Stationary Velocity Fields

University College London
Medical Physiscs and Biomedical Engineering

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Master of Research

August 4, 2015

Supervisor
Tom Vercauteren

Co-Supervisor
Marc Modat

Clinical Supervisor
Jan Deprest

Acknowledgments

It was not and it is still not easy for me the trail between the study of pure mathematics and the world of medical imaging and biomedical engineering. This route would have never been possible without guides, who have gone before me through the same tortuous paths and who accompanied me in the last year. The main contribution on this side arrived from Marco Lorenzi. With the commitment of a supervisor, he helped me significantly, not only in the development of this thesis, but especially in having introduced me to many of the problems a researcher have to deal with.

I am also grateful to the rowing fellows with whom I shared commitment, challenges and happiness in this first fascinating and demanding year at the GIFT-Surg Project: Francois Chadebecq, Pankaj Daga, Tom Doel, George Dwyer, Michael Ebner, Luis Herrera, Ioannis Kourouklides, Efthymios Maneas, Sacha Noimark, Rosalind Pratt, Marcel Tella, Gustavo Santos, Dzhoshkun Shakir, Guotai Wang and Maria A. Zuluaga.

For the help in the unknown land of infinite dimensional Lie algebra, I have a debt with professor Karl-H. Neeb and dott. Robert Gray.

A non academic, but not less important contribution came on different sides from Andrea Baglione, Filippo Ferraris, Gerardo Ballesio, Giuliano 'er Nuanda' De Rossi, Silvia Porter and Raoul Resta.

The eclectic buildings of UCL would have been unseen for me without Tom Vercauteren, Sebastien Ourselin and Gary Zhang. Their work and their decision, in a warm day of June 2014, to offer me their support - other than a desk, a laptop and a coffee machine - opened the greatest and most important opportunity I've ever had.

In classical music, it is well known that in every concert the two most important tunes are the first one and the last one. Following here the same rule I terminate acknowledging for the great love, effort and patience, Carole Sudre.

Abstract

Image registration is one of the critical tools in medical imaging. It consists in the process of aligning two or more patient images with the aim of determining and quantifying the occurring anatomical correspondences and differences.

To model the anatomical variability from one image to the other the set of diffeomorphisms (bijective differentiable maps with differentiable inverse) appears to be an interesting option.

Comparing two diffeomorphisms, as well as obtaining any meaningful statistics for these elements, is not a straightforward task. Approaching this problem, the *log-Euclidean framework*, proposed to consider the set of diffeomorphisms with a Lie group structure, having a Lie algebra defined as the tangent space at the origin where to compute statistics.

In this local representation the operation of composition of diffeomorphisms is not anymore available. An operation in the tangent space that reflects the composition in the Lie group is therefore required.

Aim of this thesis is to find and compare numerical computations of this operation, called here *log-composition*. A fast numerical methods for its computation would improve the computation of *log-demons* registration algorithm.

This document contains 15xxx words.

Contents

Chapter 1

Introduction and Motivations

*The series is divergent, therefore we may be able
to do something with it.*
- Oliver Heaviside

Medical image registration is a set of tools and techniques oriented to solve the problem of determining correspondences between two or more images acquired from patients scans. Its development is a creative field that has seen the application of a growing number of mathematical theories in the research of customizations and improvements of precision and computational time. It has a wide range of applications: for example it can be used in lungs motion correction [MHSK, MHM$^+$11], in Alzheimer disease diagnoses [PCL$^+$15, FF97, GWRNJ12], and image mosaicing [VPM$^+$06, Sze94].

Determining correspondences between images is often presented as an ill-posed problem: transformations between anatomies are not unique, and the impossibility to recover spatial or temporal evolution of an anatomical transformations, starting only from few images over a long time makes any validation a difficult, if not an impossible task.

The most important feature in image registration algorithms is the *deformation model*: the set of transformations chosen to model the anatomical deformations. This is aimed constrain deformations according to the task that the registration algorithm has to perform and to adapt to the nature of the objects represented by the images (see [ISNC03] for a presentation of the image registration framework and [SDP13] for a recent survey in medical image registration).

1.1 Choosing the Deformations: Diffeomorphisms

If the registration algorithm is meant to model physical transformations that preserve distances, orientations and angles, then the deformation model can be bounded to the group of rigid body transformations [Gal11]. The consequent registration algorithm, called *rigid-registration algorithm*, will be suitable for example to compensate the motion in a rapid sequence of scans, or to investigate small differences that occurs in longitudinal studies.

If the algorithm is meant to model transformations that only preserves topology, then the transformations must allow more freedom than the one chosen for the rigid case. It is in this context that arises the mathematical object of *diffeomorphism* over $\Omega \subseteq \mathbb{R}^d$. This is defined as a bijective differentiable map from Ω to itself, with differentiable inverse, and is particularly well suited to model non-rigid deformation between images. Algorithms involving diffeomorphisms are called *diffeomorphic registration algorithms*.

In both cases, the transformation belongs to a *group structure* (see [Art11]). Here the only operation of composition is defined, and for mathematical reasons it is not possible to define any norm or mean. A strategy to solve this problem is to consider the group with a differentiable manifold structure compatible with the operation of composition, called *Lie group*, and to consider the linear approximation of its element in the tangent space, called *Lie algebra*. In this vector space it possible to define a norm and therefore to compute statistics. This Matster Thesis is strongly based on the mathematical concepts just introduced, but a short formal introduction of each of them would require to overcome at least twice the available number of worlds. The interested reader can refer to the wide literature, starting for example from [Lee12, Arn06, War13, DCDC76].

Considering the group of diffeomorphisms as a Lie group was proposed in medical imaging for the first time in 2006, with the name of *log-Euclidean framework* [AFPA06]. The passage from the structure of the Lie group to the structure of the Lie algebra is performed using two numerical algorithm for the computation of the *Lie exponential* (the map from the Lie group to the Lie algebra) and the *Lie logarithm* (the map from the Lie algebra to the Lie group).

It is important to notice that after a transformation has moved from the Lie group to the Lie algebra, it is not possible anymore to compute the composition with another element of the group unless it is not exponentiated again. The composition can be performed only in the Lie group, and there is no operation available in the Lie algebra that reflects the feature of the composition of the corresponding transformations.

It is in this context that arises the abstract concept of log-composition, whose numerical computations are the main aim of this research.

1.2 Introducing the Log-composition and the BCH formula

Indicating with exp the Lie exponential and with log the Lie logarithm, the log-composition is defined for each couple of vectors $\mathbf{v}_1$, $\mathbf{v}_2$ belonging to the domain of exp as

$$\mathbf{v}_1 \oplus \mathbf{v}_2 := \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2)) \tag{1.1}$$

A schematic perspective of the formula can be visualized in figure 1.1. The Lie group, indicated with $\mathbb{G}$, is represented by the gray surface, while the Lie algebra, indicated with $\mathfrak{g}$, is represented by the tangent plane at the identity e. Starting from the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in $\mathfrak{g}$, their log-composition provides the vector that corresponds to the composition of the corresponding transformation of the initial vectors.

The analytic solution to the log-composition is provided by the BCH formula, when both the Lie exponential and Lie logarithm can be expressed in power series:

$$BCH(\mathbf{u},\mathbf{v}) = \mathbf{u} + \mathbf{v} + \frac{1}{2}[\mathbf{u},\mathbf{v}] + \frac{1}{12}([\mathbf{u},[\mathbf{u},\mathbf{v}]] + [\mathbf{v},[\mathbf{v},\mathbf{u}]]) - \frac{1}{24}[\mathbf{v},[\mathbf{u},[\mathbf{u},\mathbf{v}]]] + ...$$

Each of its term is defined by an infinite series of growing nested Lie bracket.

There are many issues, both theoretical and practical that prevent from the direct use of this formula. The fist and most obvious is that it is an infinite series whose truncations does not possess any asymptotic behaviour. In addition, even considering a large enough number of term, the Lie bracket of two tangent vector of the Lie group of diffeomorphisms involves the computation of the Jacobian matrices: these are not straightforward and raises some issues presented in section 5.2.3.

On the theoretical side, the proof of the BCH is based on the fact that Lie logarithm and exponential have to be expressed in power series. This happen only when the Lie group and its Lie algebra are subsets of a bigger algebra that contains them both. In the case of matrix Lie group (see [Hal15] for its definition) this always happen, but in the case of diffeomorphisms, this happens only thanks to some theoretical arrangements presented in section 3.2.

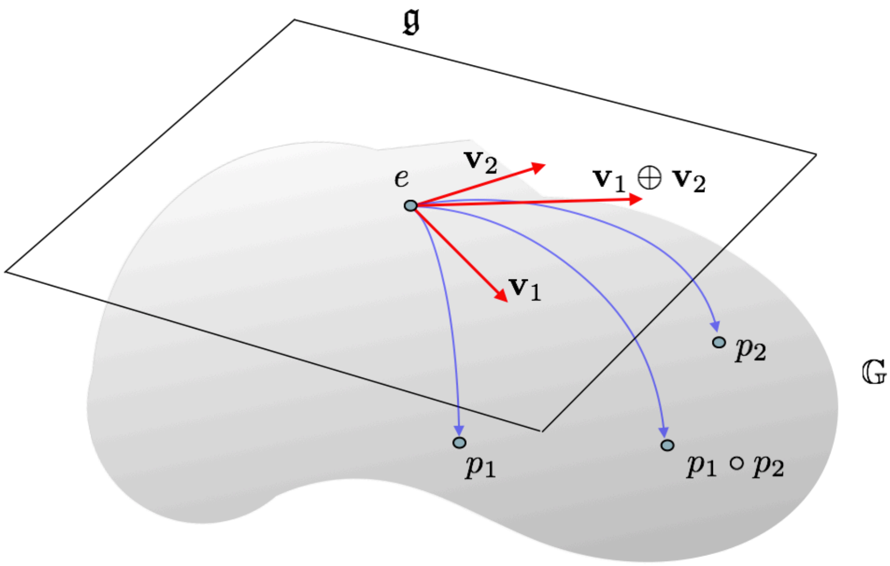

Figure 1.1: graphical visualization of the Lie log-composition $\mathbf{v}_1 \oplus \mathbf{v}_2$. The gray surface represents a Lie group and its tangent plane represents its Lie algebra.

Aim of this research is to present some numerical methods for the computation of the log-composition that avoid the use of the truncated BCH. One of these methods, here presented for the first time, is based on a geometrical construction on the differentiable manifold of the transformation that utilizes the parallel transport (see section 2.3).

Results are not only examined on the Lie group of diffeomorphisms, but also on the matrix Lie group of rigid body transformation of the plane. In this second structure is possible to evaluate the performance on a space where all the closed form are known and the ground truth are known.

1.3 Feasible Applications of the Log-composition in Medical Imaging

One of the reasons why mathematics is considered a powerful tool is consequence of the fact that a concept defined to solve a problem can, at the same time, be used to solve other problem of different origin and nature - the *unreasonable effectiveness of mathematics in the natural sciences*, also known as Wigner principle [Wig60]. This makes every tool developed in the domain of abstraction a versatile tool, but at the same time difficult to understand.

The log-composition has been defined in this chapter from the problem of the computation of the statistics on the group of diffeomorphisms, but its use is not limited to perform only this task.

In medical imaging there are several other situations in which fast and accurate computation of the log-composition can be used:

1. Diffeomorphic demons [VPPA07] and log-demons algorithm [VPPA08]. In particular

in the log-demons, the update at each of the iterative step is computed in the tangent space, using an equivalent formulation of the log-composition.

2. Fast computation of the logarithm computation [BO08]. Details of this algorithm are a part of this research and are discussed in chapter 4.

3. Calculus on diffusion tensor [AFPA06]. The logarithmic multiplication $\odot$ and the logarithmic scalar multiplication $\otimes$ provides the Lie group with a structure of vector field. The log-composition $\oplus$ here proposed provides the Lie algebra with a structure that reflects the composition of the group.

4. Image set classification [HWS$^+$]. As based on the log-euclidean framework, could exploit the property of having the group composition in the tangent space.

5. Computation of the the discrete ladder for the parallel transport [LP14a]. An equivalent of the log-composition is utilized for the computation of the parallel transport. Reversing the procedure, the parallel transport can be utilized for the computation of the log-composition.

Before moving to the next chapter, aimed to present the numerical techniques for its computation, it is important to spend a couple of words about the infinite dimensional Lie groups.

1.4 A Short Remark about the Lie Group of Diffeomorphisms

So far we have talked about the group of diffeomorphisms as a Lie group in a natural way, without underline any particular feature of this concept. Actually, the topic of infinite dimensional Lie gorup is an open field of research whose development has not yet reach a definitive formalization. Aimed to presenting the theoretical problem and difficutlies as well as how we deal with them in this Master Thesis, we retrace the main historical steps and some of the most significant important approaches.

The first attempt to provide some handles to the group of diffeomorphisms for easy manipulation was done for the first time in 1966 by Vladimir Arnold [Arn66] (consider also the equivalent [Arn98], more readable for non-French speakers). To solve differential equation in hydrodynamic, the set of diffeomorphisms *Diff* is considered as a Lie group possessing a Lie algebra. This assumption is not formally followed in accordance to the problem-oriented nature of this paper.

Subsequent steps in the exploration of the set of diffeomorphisms as a Lie group, and in the attempt of finding a formalization can be found in [MA70, EM70, Omo70, Mic80, Les83]. A state of the art of infinite dimensional Lie group in the early eighties can be found in [Mil84a], while more recent results and applications on diffeomorphisms has been published in [OKC92, BHM10, Sch10, BBHM11].

Considering an infinite dimensional group as a differentiable manifold implies the idea of having each of its elements in local correspondence with some generalized “infinite-dimensional Euclidean” space. Attempts to set this correspondence showed that, the transition maps are smooth over the Banach spaces. This led to the idea of Banach Manifolds. It has been shown [KW08] that the group of diffeomorphisms defined as a manifold does not belongs to the category of Banach manifold but requires an even more general space on which the transition maps are smooth: the Frechet space. Here, important theorems from analysis, as the inverse function theorem, the Frobenius theorem, or the main results from the Lie group theory in a finite dimensional settings, as Lie correspondence theorems, do not hold anymore.

These difficulties led some researchers in approaching the set of diffeomorphisms from other perspectives: for example, instead of treating *Diff* as a group equipped with differential structures, it is seen as a quotient of other well behaved groups [Woj94]. In other cases, as in [MA70] first and in [Mil84b] later, Banach spaces are substituted with more general locally convex spaces to underpin the definition of smooth manifolds (an formal introduction to the infinite dimensional linear Lie groups, group of smooth maps and group of diffeomorphisms can be found in [Nee06]).

For the medical imaging purposes, it is not necessary to consider the general theory of infinite dimensional manifolds. Keeping the initial Arnold's problem-oriented perspective, the interest are only about the diffeomorphisms defined on a compact subset Ω of $\mathbb{R}^d$. Without denying the importance of fundamentals and underestimating the doors research for generalized infinite dimensional Lie group may open, on the formal side we will approach the matter in as similar way of what has been done in set theory: we will use a *naive approach* to infinite dimensional Lie group. Here the fundamental definition of infinite dimensional Lie group is a generalization of the finite dimensional case of matrices, and it is left more to the intuition than to a robust formalization.

Chapter 2

Tools from Differential Geometry

Give me six hours to chop down a tree
and I will spend the first four sharpening the axe.
-Abraham Lincoln

2.1 A Lie Group Structure for the Set of Transformations

We consider every group $\mathbb{G}$ as a group of transformations acting on $\mathbb{R}^d$, having in mind the particular case $d = 2, 3$ for 2-dimensional or 3-dimensional images. We will focus our attention to transformations defined by matrices or diffeomorphism. Other than group they also have the structure of Lie group: they are considered with a maximal atlas that makes them differentiable manifold, in which the composition of two transformations and the inverse of each transformation are well defined differentiable maps:

$$\begin{aligned} \mathbb{G} \times \mathbb{G} &\longrightarrow \mathbb{G} \\ (x, y) &\longmapsto xy^{-1} \end{aligned}$$

Differential geometry is, generally speaking, a technique to use the well known calculus features and operators on spaces different from the usual $\mathbb{R}^n$. Adding the differentiable structure to a group of transformations provides new handles to hold and manipulate them; in particular it provides the opportunity to define a tangent space to each point of the group (and so a fiber bundle), a space of vector fields, a set of flows and one parameter subgroup as well as other features that enrich this structure.

Due to space limitations we will refer to [DCDC76] and [Lee12] for the definitions and concepts of differential geometry and [dCV92] for definition and concepts of Riemannian geometry.

2.2 Lie Exponential, Lie logarithm, Lie log-composition and the BCH formula

Let $\mathbf{v}$ be an element in the tangent space $\mathfrak{g}$ of the Lie group $\mathbb{G}$. The *Lie exponential* is defined as

$$\begin{aligned} \exp : \mathfrak{g} &\longrightarrow \mathbb{G} \\ \mathbf{v} &\longmapsto \exp(\mathbf{v}) = \gamma(1) \end{aligned}$$

where $\gamma : [0,1] \to \mathbb{G}$ is the unique one-parameter subgroup of $\mathbb{G}$ having $\mathbf{v}$ as its tangent vector at the identity. The identity is indicated with e for the general case, I for matrices and 1 for diffeomorphisms. The exponential map satisfies the following properties:

1. $\exp(t\mathbf{v}) = \gamma(t)$.
2. $\exp(\mathbf{v}) = e$ if $\mathbf{v} = \mathbf{0}$.
3. $\exp(\mathbf{v}) \circ \exp(-\mathbf{v}) = e$
4. As a direct consequence of the definition here provided, based on the one parameter subgroup, it follows that:

   $$\exp((t+s)\mathbf{v}) = \gamma(t+s) = \gamma(t) \circ \gamma(s) = \exp(t\mathbf{v}) \exp(s\mathbf{v})$$

   This in particular is one of the reasons that justifies the name exponential for a maps between structures in differential geometry.
5. $\exp(\mathbf{v})$ is invertible and $(\exp(\mathbf{v}))^{-1} = \exp(-\mathbf{v})$.
6. $\exp(\mathbf{u} + \mathbf{v}) = \lim_{m\to\infty}(\exp(\frac{\mathbf{v}}{m}) \circ \exp(\frac{\mathbf{v}}{m}))^m$
7. exp is a local isomorphism: it is an isomorphisms between a neighborhood of $\mathbf{0}$ in $\mathfrak{g}$ to a neighborhood of e in $\mathbb{G}$.
8. If $\exp(\mathbf{w}) = \exp(\mathbf{u}) \exp(\mathbf{v})$ then

   $$\exp(-\mathbf{w}) = \exp(-\mathbf{v}) \exp(-\mathbf{u}) \tag{2.1}$$

Proof. We will present only the last statement leaving the others to the literature. The hypothesis $\exp(\mathbf{w}) = \exp(\mathbf{v}) \circ \exp(\mathbf{u})$ follows the subsequent chain of implications (each algebraic passage involves a geometrical construction, not shown here for brevity):

$$\begin{aligned} \exp(\mathbf{w}) &= \exp(\mathbf{v}) \circ \exp(\mathbf{u}) \\ \exp(-\mathbf{w}) \circ \exp(\mathbf{w}) &= \exp(-\mathbf{w}) \circ \exp(\mathbf{v}) \circ \exp(\mathbf{u}) \\ e &= \exp(-\mathbf{w}) \circ \exp(\mathbf{v}) \circ \exp(\mathbf{u}) \\ \exp(-\mathbf{u}) &= \exp(-\mathbf{w}) \circ \exp(\mathbf{v}) \\ \exp(-\mathbf{u}) \circ \exp(-\mathbf{v}) &= \exp(-\mathbf{w}) \end{aligned}$$

□

The neighborhoods of $\mathbb{G}$ and of $\mathfrak{g}$ such that the last property holds, are called *internal cut locus* of $\mathbb{G}$ and $\mathfrak{g}$ respectively. The *cut locus* is the boundary of the internal cut locus.

When we deal with a matrix Lie group of dimension n, the composition in the Lie group consists in the matrix product and we have the following remarkable properties [Hal15], [Kir08]:

1. for all $\mathbf{v}$ in a matrix Lie algebra $\mathfrak{g}$:

$$\exp(\mathbf{v}) = \sum_{k=0}^{\infty} \frac{\mathbf{v}^k}{k!} \tag{2.2}$$

2. If $\mathbf{u}$ and $\mathbf{v}$ are commutative then $\exp(\mathbf{u}+\mathbf{v}) = \exp(\mathbf{u})\exp(\mathbf{v})$.

3. If $\mathbf{c}$ is an invertible matrix then $\exp(\mathbf{cvc}^{-1}) = \mathbf{c}\exp(\mathbf{v})\mathbf{c}^{-1}$.

4. $\det(\exp(\mathbf{v})) = \exp(\mathrm{trace}(\mathbf{v}))$

5. For any norm, $||\exp(\mathbf{v})|| \leq \exp(||\mathbf{v}||)$.

The idea of defining an inverse of the Lie exponential leads to the idea of the Lie logarithm, defined as

$$\begin{aligned} \log : \mathbb{G} &\longrightarrow \mathfrak{g} \\ \varphi &\longmapsto \log(\varphi) = \mathbf{v} \end{aligned}$$

where $\mathbf{v}$ is the tangent vector having φ as it exp.
If $\mathbb{G}$ is a matrix Lie group of dimension n, the following properties hold:

1. for all φ in the matrix Lie group $\mathbb{G}$:

$$\log(\varphi) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\varphi - I)^k}{k} \tag{2.3}$$

   Having indicated the identity matrix with I.

2. For any norm, and for any $n \times n$ matrix $\mathbf{c}$, exists an α such that

$$||\log(I+\mathbf{c}) - \mathbf{c}|| \leq \alpha ||\mathbf{c}||^2 \tag{2.4}$$

3. For any $n \times n$ matrix $\mathbf{c}$ and for any sequence of matrix $\{\mathbf{d}_j\}$ such that $||\mathbf{d}_j|| \leq \alpha/j^2$ it follows:

$$\lim_{k\to\infty} \left(I + \frac{\mathbf{c}}{k} + \mathbf{d}_k\right)^k = \exp(\mathbf{c}) \tag{2.5}$$

The *Lie log-composition* (because based on the Lie logarithm and Lie exponential maps) is defined here as the inner binary operation on the Lie algebra that reflects the composition on the lie group:

$$\begin{aligned} \oplus : \mathfrak{g} \times \mathfrak{g} &\longrightarrow \mathfrak{g} \\ (\mathbf{v}_1, \mathbf{v}_2) &\longmapsto \mathbf{v}_1 \oplus \mathbf{v}_2 = \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2)) \end{aligned}$$

The following properties holds for the Lie log-composition:

1. $\mathfrak{g}$ with the Lie log-composition $\oplus$ is a local topological non-commutative group (local group for short): if $C_{\mathfrak{g}}$ is the internal cut locus of $\mathfrak{g}$ then:

   (a) $(\mathbf{u}_1 \oplus \mathbf{u}_2) \oplus \mathbf{u}_3 = \mathbf{u}_1 \oplus (\mathbf{u}_2 \oplus \mathbf{u}_3)$ for all $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ in $C_{\mathfrak{g}}$.

   (b) $\mathbf{u} \oplus \mathbf{0} = \mathbf{0} \oplus \mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.

   (c) $\mathbf{u} \oplus (-\mathbf{u}) = \mathbf{0}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.

2. For all t, s real, such that $(t+s)\mathbf{u}$ is in $C_{\mathfrak{g}}$,

$$(t\mathbf{u}) \oplus (s\mathbf{u}) = (t+s)\mathbf{u}$$

   And in particular, if the Lie algebra $\mathfrak{g}$ has dimension 1 the local group structure is compatible with the additive group of the vector space $\mathfrak{g}$.

The algebraic structure $(\mathfrak{g}, \oplus)$ is called Lie log-group. Additional observations on this algebraic structure in the particular case of diffeomorphisms, are proposed in the next chapter.

To compute the log-composition there is the Backer-Campbell-Hausdorff formula, or BCH, that provides the exact solution to the log-composition:

$$BCH(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \mathbf{v} + \frac{1}{2}[\mathbf{u}, \mathbf{v}] + \frac{1}{12}([\mathbf{u}, [\mathbf{u}, \mathbf{v}]] + [\mathbf{v}, [\mathbf{v}, \mathbf{u}]]) - \frac{1}{24}[\mathbf{v}, [\mathbf{u}, [\mathbf{u}, \mathbf{v}]]] + ...$$

Ironically, the name of the formula does not refer to Dynkin, who originally developed the proof of the equality in 1947 [Dyn00]. An nice introduction to the particular case of matrices can be found in [Hal15]; while the general case is presented in [KO89], [Ser09], and for application to medical imaging [LP14b], [VPPA08]. This expansion provides the most immediate way to obtain a numerical computation of $\mathbf{u} \oplus \mathbf{v}$, by truncating its terms, but as stated in the introduction, chapter 1 this approximation can be problematic.

2.3 Affine Exponential Affine Logarithm and Parallel Transport: Definitions and Properties

Considering a Lie Group $\mathbb{G}$ with a connection ∇, the vector field $\nabla_U(V)$ associates at each point of the manifold the projection on the tangent plane of the derivative of U in the direction of V.

One of the considerable consequences of the definition of the connection is the possibility of defining *geodesics* and curvature on the manifold without relying on any Riemannian metric. If a Riemannian metric is also defined on the manifold $\mathbb{G}$, then geodesics defined by the metric coincides with the geodesics defined by the connection only for the particular case of Levi-Civita connection (see [dCV92]). A curve $\gamma : [0,1] \to \mathbb{G}$ such that $\gamma(0) = p$ and $\gamma(1) = q$ is a *geodesic* defined by the connection ∇ if

$$\nabla_{\dot{\gamma}} \dot{\gamma} = 0 \tag{2.6}$$

This definition allows a new kind of exponential from the Lie algebra to the Lie group. Given the point p and the tangent vector at this point $\mathbf{v} \in T_p\mathbb{G} \simeq \mathfrak{g}$ we define:

$$\begin{aligned} \exp : \mathbb{G} \times \mathfrak{g} &\longrightarrow \mathbb{G} \\ (p, \mathbf{v}) &\longmapsto \exp_p(\mathbf{v}) = \gamma(1; p, \mathbf{v}) \end{aligned}$$

such that the curve $\gamma(t; p, \mathbf{v}) = \gamma(t)$ on $\mathbb{G}$ is the unique geodesic that satisfies $\gamma(0) = p$ and $\dot{\gamma}(0) = \mathbf{v}$. This second kind of exponential differs from the exponential map previously introduced by the fact that the tangent space that defines the Lie algebra is considered at the generic point p of the Lie group and it is called *affine exponential*.
The inverse of the affine exponential, the *affine logarithm* is defined as:

$$\begin{aligned} \log : \mathbb{G} \times \mathbb{G} &\longrightarrow T_p\mathbb{G} \simeq \mathfrak{g} \\ (p, q) &\longmapsto \log_p(q) = \mathbf{v} \end{aligned}$$

Where $\mathbf{v}$ is the tangent vector in p at the geodesic γ on $\mathbb{G}$ that satisfies $\gamma(0) = p$ and $\gamma(1) = q$. Interestingly the Lie exponential and the Lie logarithm coincide with the affine exponential and the affine logarithm at the identity, if ∇ is a Cartan connection.

For further details and properties we refer to the literature; in this introduction we wish to provide only the intuitive idea that it is possible to move on the fiber bundle of the Lie group, transporting in some sense a tangent vector defined at the identity on another tangent space. Certainly the Lie group possesses a unique Lie algebra, as the tangent space at some point (the group's identity by convention), but two different tangent space (so two times the same isomorphic Lie algebra structure) may not have the basis vectors oriented in the same direction.

2.3.1 An introduction to Parallel Transport: Surfing on the Tangent Bundle

In this section we introduce the concept of parallel transport for the Lie group $\mathbb{G}$. For an introduction to parallel transport in the general case we refer to [MTW73], [Kne51], [KMN00]; for medical imaging applications [LAP11], [PL+11], [LP13] and [LP14b]. On this definition, again borrowed from differential geometry, relies a method for the computation of the log-composition developed in this research for the first time.

Definition 2.3.1. Let $\mathbb{G}$ be a finite dimensional connected Lie group defined with a connection ∇ and V a $\mathcal{C}^\infty$ vector field defined over $\mathbb{G}$. Given $p, q \in \mathbb{G}$ and $\gamma : [0,1] \to \mathbb{G}$ such that $\gamma(0) = p$ and $\gamma(1) = q$, the vector $V_p \in T_p\mathbb{G}$, is *parallel transported along* γ up to $T_q\mathbb{G}$ if V satisfies

$$\forall t \in [0,1] \qquad \nabla_{\dot\gamma} V_{\gamma(t)} = 0$$

The *parallel transport* is the function that maps V_p from $T_p\mathbb{G}$ to $T_q\mathbb{G}$ along γ:

$$\begin{aligned} \Pi(\gamma)_p^q : T_p\mathbb{G} &\longrightarrow T_q\mathbb{G} \\ V_p &\longmapsto \Pi(\gamma)_p^q(V_p) = V_q \end{aligned}$$

Consequence of this definition is that a vector belonging to the tangent space at the identity can be transported on a different tangent space of the manifold, maintaining its direction from the old to the new coordinate reference respect to a chosen curve. Each element of the *fiber bundle* (disjoint union of all of the tangent space), that can be reached by a curve from the origin, become reachable also by any tangent vector at the identity.

Another way of moving vectors between an arbitrary tangent spaces and the tangent space at the identity is expressed in the *change of base formulas* for affine exponential and logarithm [APA06]:

$$\log_p(q) = DL_p(e)\log_e(q) \tag{2.7}$$

$$\exp_p(\mathbf{u}) = p \circ \exp_e(DL_{p^{-1}}(e)\mathbf{u}) \tag{2.8}$$

The left-translation L_p provides a canonical curves for transporting vectors, expressed as the integral curve of the tangent vector field on the manifold of transformations defined by the push forward of L_p, indicated here with DL_p:

Further theoretical developments are beyond the aim of this research, but the reader can refer to the bibliography. In the next properties we explore how did parallel transport and affine exponential behave when expressed as a composition and when there is a change of signs.

Property 2.3.1 (Inversion)**.** $\mathbb{G}$ Lie group, ∇ connection, $p, q \in \mathbb{G}$. Given γ such that $\gamma(0) = p$, $\gamma(1) = q$ and $\mathbf{u} \in T_p\mathbb{G}$, we have:

1. $\Pi(\gamma)_p^q(-\mathbf{u}) = -\Pi(\gamma)_p^q(\mathbf{u})$
2. $q = \exp_p(\mathbf{u}) \iff p = \exp_q(-\Pi(\gamma)_p^q(\mathbf{u}))$

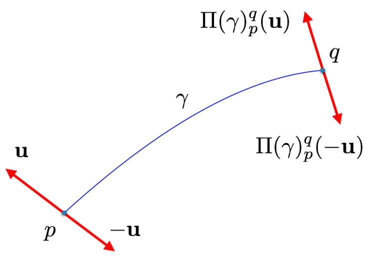

Figure 2.1: First inversion property.

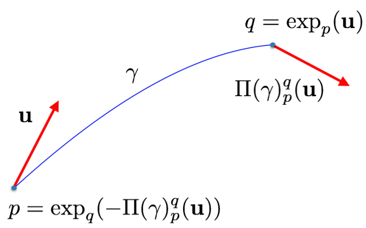

Figure 2.2: Second inversion property.

Proof. The first statement is a consequence of the fact that parallel transport is reversible, conserve the parallelism and it is invariant respect to norm:

i) $\Pi(-\gamma)_q^p\Big(\Pi(\gamma)_p^q(\mathbf{u})\Big) = \mathbf{u}$ where $-\gamma$ corresponds to γ walked in the opposite direction.

ii) $\Pi(\gamma)_p^q(\mathbf{u})$ is parallel in the same tangent space to $\Pi(\gamma)_p^q(\lambda\mathbf{u})$ for any nonzero λ.

iii) $||\Pi(\gamma)_p^q(\mathbf{u})|| = ||\mathbf{u}||$

For the second statement, if $q = \exp_p(\mathbf{u})$ then it exists a curve γ that is a geodesics and connects p with q:

$$\exp_p(\mathbf{u}) = \gamma(1; \mathbf{u}, p) \qquad \nabla_{\dot\gamma}\dot\gamma = 0 \qquad \gamma(0) = p \qquad \gamma(1) = q$$

On the other side, if $p = \exp_q(-\Pi(\gamma)_p^q(\mathbf{u}))$, then it exists a curve β that is a geodesic and connect q with p:

$$\exp_q(-\Pi(\gamma)_p^q(\mathbf{u})) = \beta(1; -\Pi(\gamma)_p^q(\mathbf{u}), q) \qquad \nabla_{\dot\gamma}\dot\beta = 0 \qquad \beta(0) = q \qquad \beta(1) = p$$

Since there is a unique curve that satisfies the condition of being geodesic between two points, we have $\gamma = -\beta$. Therefore, if $q = \exp_p(\mathbf{u})$, then

$$p = \gamma(0; \mathbf{u}, p) = \beta(1; -\Pi(\gamma)_p^q(\mathbf{u}), q)$$

which implies $p = \exp_q(-\Pi(\gamma)_p^q(\mathbf{u}))$. On the other side, if $p = \exp_q(-\Pi(\gamma)_p^q(\mathbf{u}))$, then

$$q = \beta(0; -\Pi(\gamma)_p^q(\mathbf{u}), q) = \gamma(1; \mathbf{u}, p)$$

which implies $q = \exp_p(\mathbf{u})$. □

Property 2.3.2. Let $\mathbb{G}$ be a finite dimensional connected Lie group defined with a Cartan connection ∇ and $\mathbf{u}$ tangent vector in $T_e\mathbb{G}$. Let γ be a geodesic defined on $\mathbb{G}$ such that $\gamma(0) = e$, $\dot\gamma(0) = \mathbf{u}$ and $p = \gamma(1)$, point in the Lie group. Let β be the curve over $\mathbb{G}$ defined as $\beta(t) = p \circ \gamma(t)$, then the two following conditions hold:

1. If ∇ is a Cartan connection then β is a geodesic.

2. For $\mathbf{u}_p := D(L_p)_e(\mathbf{u}) \in T_p\mathbb{G}$, push forward of the left-translation:

$$\exp_p(t\mathbf{u}_p) = p \circ \exp_e(tD(L_{p^{-1}})_p(\mathbf{u}_p)) = p \circ \exp_e(t\mathbf{u}) \tag{2.9}$$

Proof. The first statement belongs to the general theory and it is not proved here: geodesics are left-invariant for a Cartan connection (see [dCV92]). To prove the second statement we consider the properties of β that directly follows from the definition:

$$\beta(0) = p \circ e = p$$
$$\dot{\beta}(0) = DL_p(e)\mathbf{u} \in T_p\mathbb{G}$$

For simplicity $\dot{\beta}(0)$ was indicated with $\mathbf{u}_p$. Considering $\beta(1)$ we have:

$$\beta(1) = p \circ \gamma(1) = p \circ \exp_e(\mathbf{u}) = \exp_p(DL_p(e)\mathbf{u}) = \exp_p(\mathbf{u}_p)$$

where the third equality comes from the change of base formulas for affine exponential 2.7. Following the same deduction and from the linearity of the differential, we have, for any $t \in [0, 1]$:

$$\beta(t) = p \circ \gamma(t) = p \circ \exp_e(t\mathbf{u}) = \exp_p(tDL_p(e)\mathbf{u}) = \exp_p(t\mathbf{u}_p)$$

□

Lemma 2.3.1. Let $\mathbb{G}$ be a finite dimensional connected Lie group, p, q, r points of $\mathbb{G}$ belonging to the cut locus. If exists an ϵ such that

$$||\log(p \circ q) - \log(r)|| < \epsilon$$

then it follows

$$||\log(p) - \log(q^{-1} \circ r)|| < \epsilon$$

Intuitively, the lemma states that if $p \circ q \simeq r$ then $p \simeq q^{-1} \circ r$.

The following theorem is an application of the pole ladder [LAP11] for the computation of the exponential that will underpin one of the numerical methods for the computation of the log-composition.

Theorem 2.3.1. Let $\mathbb{G}$ be a finite dimensional connected Lie group defined with a Cartan connection ∇. Given two vectors $\mathbf{u}, \mathbf{v}$ in the internal cut locus of $\mathfrak{g}$, such that $p = \exp_e(\mathbf{u})$ and $q = \exp_e(\mathbf{v})$, with α integral curve of $\mathbf{u}$,

$$\alpha : [0, 1] \to \mathbb{G} \quad \alpha(0) = e \quad \alpha(1) = p \quad \dot{\alpha}(0) = \mathbf{u}$$

and for $\mathbf{v}_p^{||}$ parallel transport of $\mathbf{v}$

$$\mathbf{v}_p^{||} = \Pi(\alpha)_e^p(\mathbf{v})$$

and $\mathbf{v}_e^{||}$ pull-back of the left translation of the previous vector

$$\mathbf{v}_e^{||} := D(L_{p^{-1}})_e(\mathbf{v}_p^{||})$$

it follows that:

$$||\log_e(\exp_e(\mathbf{v}_e^{||})) - \log_e(\exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big))|| \le ||[\mathbf{u}, \mathbf{v}]||$$

The statement of this fairly intricate theorem involves a construction that can be visualized in figure 2.3.

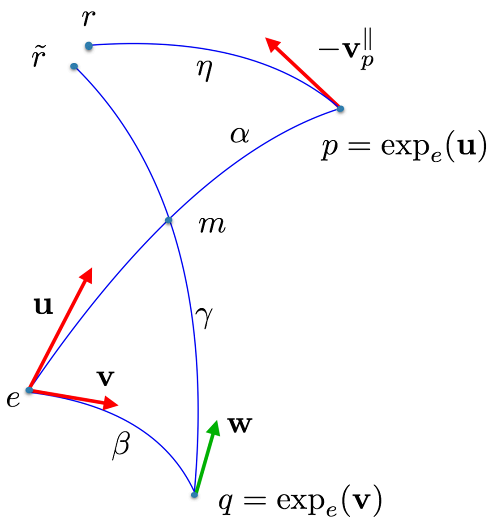

Figure 2.3: Pole ladder applied to parallel transport.

Proof. Let $m \in \mathbb{G}$ be the midpoint of the curve α, $m = \alpha(1/2) = \exp_e(\frac{\mathbf{u}}{2})$ and let γ be the geodesic between $q = \exp_e(\mathbf{v})$ and m:

$$\gamma(0) = q \qquad \gamma(1) = m \qquad \nabla_{\dot{\gamma}}\dot{\gamma} = 0$$

If $\mathbf{w}$ is the tangent vector of γ defined at q such that $\dot{\gamma}(0) = \mathbf{w}$, it follows from the change of base formula 2.7 that

$$\gamma(t) = \exp_q(t\mathbf{w}) = q \circ \exp_e(DL_{p^{-1}}(e)(t\mathbf{w})) = \exp_e(\mathbf{v}) \circ \exp_e(tDL_{p^{-1}}(e)\mathbf{w})$$

And by construction, we can move from the identity e to m directly walking the geodesic α or passing through q. It follows that

$$\begin{aligned}
\exp_q(\mathbf{w}) &= \exp_e(\frac{\mathbf{u}}{2}) \\
q \circ \exp_e(DL_{q^{-1}}(e)(\mathbf{w})) &= \exp_e(\frac{\mathbf{u}}{2}) \\
\exp_e(DL_{q^{-1}}(e)(\mathbf{w})) &= \exp_e(-\mathbf{v}) \circ \exp_e(\frac{\mathbf{u}}{2})
\end{aligned}$$

Let η be the integral curve of the tangent vector $-\mathbf{v}_p^{\|}$ at p. We define two new points, $r := \eta(1)$ and $\tilde{r} := \gamma(2)$ where γ is the integral curve of $2\mathbf{w}$.
On one side we have:

$$\begin{aligned}
\tilde{r} = \gamma(2) = \exp_q(2\mathbf{w}) &= q \circ \exp_e(DL_{q^{-1}}(e)(2\mathbf{w})) \\
&= \exp_e(\mathbf{v}) \circ \exp_e(2DL_{q^{-1}}(e)\mathbf{w}) \\
&= \exp_e(\mathbf{v}) \circ \exp_e(DL_{q^{-1}}(e)\mathbf{w})^2 \\
&= \exp_e(\mathbf{v}) \circ \exp_e(\exp_e(-\mathbf{v}) \circ \exp_e(\frac{\mathbf{u}}{2}))^2 \\
&= \exp_e(\frac{\mathbf{u}}{2}) \circ \exp_e(-\mathbf{v}) \circ \exp_e(\frac{\mathbf{u}}{2})
\end{aligned}$$

On the other side:

$$\begin{aligned} r = \eta(1) &= \exp_p(-\mathbf{v}_p^{\|}) = p \circ \exp_e(DL_{p^{-1}}(e)(-\mathbf{v}_p^{\|})) \\ &= \exp_e(\mathbf{u}) \circ \exp_e(-DL_{p^{-1}}(e)\mathbf{v}_p^{\|}) \\ &= \exp_e(\mathbf{u}) \circ \exp_e(-\mathbf{v}_e^{\|}) \end{aligned}$$

having indicated $DL_{p^{-1}}(e)\mathbf{v}^{\|}$ with $\mathbf{v}_e^{\|}$ for brevity.
By geometrical construction, we have that if the space has no curvature (or equivalently, the Lie group is commutative), $r = \tilde{r}$. Therefore, using the change of signs property 2.1

$$\begin{aligned} \exp_e(\mathbf{u}) \circ \exp_e(-\mathbf{v}_e^{\|}) &= \exp_e(\frac{\mathbf{u}}{2}) \circ \exp_e(-\mathbf{v}) \circ \exp_e(\frac{\mathbf{u}}{2}) \\ \exp_e(\mathbf{v}_e^{\|}) &= \exp_e(\frac{\mathbf{u}}{2}) \circ \exp_e(\mathbf{v}) \circ \exp_e(-\frac{\mathbf{u}}{2}) \end{aligned}$$

When the space is curved, again by construction, it follows that

$$||r - \tilde{r}|| \leq ||[\mathbf{u}, \mathbf{v}]||$$

As a consequence of the previous lemma and observing that if p and q are in the cut locus, than also r and $\tilde{r}$ are in the cut locus, we have finally reach the thesis:

$$||\log_e(\exp_e(\mathbf{v}_e^{\|})) - \log_e(\exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big))|| \leq ||[\mathbf{u}, \mathbf{v}]||$$

□

The previous result can be reformulated as the approximation:

$$\exp_e(\mathbf{v}_e^{\|}) \simeq \exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big) \tag{2.10}$$

that will turn out to be the main tool for the computation of the log-composition using parallel transport.

In the next section we present the numerical methods for the computation of the log composition.

2.4 Numerical Computations of the Log-composition

In this section we provide explicit formulas for the computation of the log composition:

$$\mathbf{v}_1 \oplus \mathbf{v}_2 = \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2)) \tag{2.11}$$

using the tools introduced in the previous sections.

2.4.1 Truncated BCH formula for the Log-composition

As said in the end of section 2.2 the Lie log-composition posses a closed form, the BCH formula, defined as the solution of the equation $\exp(\mathbf{w}) = \exp(\mathbf{u}) \circ \exp(\mathbf{v})$, for $\mathbf{u}$ and $\mathbf{v}$ *analytic* elements in the Lie algebra $\mathfrak{g}$:

$$BCH(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \mathbf{v} + \frac{1}{2}[\mathbf{u}, \mathbf{v}] + \frac{1}{12}([\mathbf{u}, [\mathbf{u}, \mathbf{v}]] + [\mathbf{v}, [\mathbf{v}, \mathbf{u}]]) - \frac{1}{24}[\mathbf{v}, [\mathbf{u}, [\mathbf{u}, \mathbf{v}]]] + ... \tag{2.12}$$

It consists of an infinite series of Lie bracket whose asymptotic behaviour cannot be predicted only from the coefficient of each nested Lie bracket term. In practical applications it can be computed using its *approximation of degree k*, defined as the sum of the BCH terms having no more than k nested Lie bracket. This convention is also coherent with the degree of the BCH expressed as polynomial formal series of adjoint operators (see next section 2.4.2):

$$\begin{aligned}
BCH^0(\mathbf{u},\mathbf{v}) &= \mathbf{u}+\mathbf{v} \\
BCH^1(\mathbf{u},\mathbf{v}) &= \mathbf{u}+\mathbf{v}+\frac{1}{2}[\mathbf{u},\mathbf{v}] \\
BCH^2(\mathbf{u},\mathbf{v}) &= \mathbf{u}+\mathbf{v}+\frac{1}{2}[\mathbf{u},\mathbf{v}]+\frac{1}{12}([\mathbf{u},[\mathbf{u},\mathbf{v}]]+[\mathbf{v},[\mathbf{v},\mathbf{u}]]) \\
BCH^3(\mathbf{u},\mathbf{v}) &= \mathbf{u}+\mathbf{v}+\frac{1}{2}[\mathbf{u},\mathbf{v}]+\frac{1}{12}([\mathbf{u},[\mathbf{u},\mathbf{v}]]+[\mathbf{v},[\mathbf{v},\mathbf{u}]])-\frac{1}{24}[\mathbf{v},[\mathbf{u},[\mathbf{u},\mathbf{v}]]]
\end{aligned}$$

In numerical computations nested Lie brackets can raise several issue, in particular when $\mathbf{u}$ and $\mathbf{v}$ are not close to the origin. Assuming that, as often happens for practical applications in imaging registration $\mathbf{v}$ is smaller than $\mathbf{u}$, we define a intermediate degree for the truncated BCH formula, between 1 and 2:

$$BCH^{3/2}(\mathbf{u},\mathbf{v}) = \mathbf{u}+\mathbf{v}+\frac{1}{2}[\mathbf{u},\mathbf{v}]+\frac{1}{12}[\mathbf{u},[\mathbf{u},\mathbf{v}]]$$

Truncated BCH formulas can be considered as a first step toward the numerical approximations of the log-composition $\mathbf{u}\oplus\mathbf{v}$. They still have some limitations as the fact that they do not provide any information about the error carried by each term, and they work under the assumption that $\mathbf{u}$ and $\mathbf{v}$ are analytic, so when we can expressed locally with a convergent power series, as in the case of tangent vectors to matrix Lie group. Additional limitation can be found when applied to stationary velocity fields. This will be one of the topic of section 3.2.5.

2.4.2 Taylor Expansion Method for the Log-composition

A more sophisticated numerical method to manage the nested Lie brackets for the computation of the log-composition is based on the Taylor expansion.
As shown in the appendix of [KO89] the terms of the BCH can be recollected using the Hausdorff method: each of the therms containing the n-th power of the vector $\mathbf{v}$ are collected together in the formal series A^n. Therefore

$$BCH(\mathbf{u},\mathbf{v}) = \mathbf{u}+A^1\mathbf{v}+A^2\mathbf{v}+A^3\mathbf{v}+\cdots$$

Given the adjoint map:

$$\begin{aligned}
ad_{\mathbf{u}}: \mathfrak{g} &\longrightarrow \mathfrak{g} \\
\mathbf{v} &\longmapsto ad_{\mathbf{u}}\mathbf{v} := [\mathbf{u},\mathbf{v}]
\end{aligned}$$

and the multiple adjoint maps, defined as:

$$\mathrm{ad}_{\mathbf{u}}^n\mathbf{v} := [\underbrace{\mathbf{u},[\mathbf{u},...[\mathbf{u}}_{\text{n-times}},\mathbf{v}]...]]$$

$$\mathrm{ad}_{\mathbf{u}}^{-n}\mathbf{v} := [[...[\mathbf{v},\underbrace{\mathbf{u}]...],\mathbf{u}}_{\text{n-times}}] = (-1)^n\mathrm{ad}_{\mathbf{u}}^n\mathbf{v}$$

it can be demonstrated that then the operator A^1, when applied to $\mathbf{v}$ provides the linear part of $\mathbf{v}$ in the BCH formula and can be written as:

$$A^1 = \frac{\mathrm{ad}_{\mathbf{u}}^{-1}}{\exp(\mathrm{ad}_{\mathbf{u}})-1} = \sum_{n=0}^{\infty}\frac{(-1)^nB_n}{n!}\mathrm{ad}_{\mathbf{u}}^{-n} = \sum_{n=0}^{\infty}\frac{B_n}{n!}\mathrm{ad}_{\mathbf{u}}^{n}$$

where $\{B_n\}_{n=0}^{\infty}$ is the sequence of the second-kind Bernoulli number. If first-kind Bernoulli number are used, then each term of the summation must be multiplied for $(-1)^n$, as did for example in [KO89]. The denominator is defined within the structure of the formal power series ring [MT13].
In conclusion, the log-composition can expressed as:

$$\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \frac{\mathrm{ad}_{\mathbf{u}}^{-1}}{\exp(\mathrm{ad}_{\mathbf{u}}) - 1}\mathbf{v} + \mathcal{O}(\mathbf{v}^2)$$

$$\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_{\mathbf{u}}^n \mathbf{v} + \mathcal{O}(\mathbf{v}^2) \tag{2.13}$$

that will turn out to be an important tool for the computation of the log-composition in the finite dimensional case.

2.4.3 Parallel Transport Method for the Log-composition

To obtain a numerical computation for the log-composition using parallel transport, we have to consider two assumptions:

1. If $\mathbf{v}_e^{\parallel}$ is defined as in theorem 2.3.1, then

$$||\mathbf{u} \oplus \mathbf{v} - (\mathbf{u} + \mathbf{v}_e^{\parallel})|| \leq ||[\mathbf{u}, \mathbf{v}]||$$

2. If the vector $\mathbf{u} \in \mathfrak{g}$ is small enough, then:

$$\exp(\mathbf{u}) \simeq e + \mathbf{u}$$

The first assumption is a consequence of geometrical intuition. On a flat space, or a space with no curvature, the geodesics are straight lines, and $\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \mathbf{v}$ that is equal, again intuitively, to the sum of $\mathbf{u}$ with the parallel transported of $\mathbf{v}$ to the point $\exp_e(\mathbf{u})$, indicated with $\mathbf{v}_p^{\parallel}$ (see figure 2.4). It is not possible to sum two vectors belonging to two different planes, therefore we have to consider the transported $\mathbf{v}_e^{\parallel}$ instead of $\mathbf{v}_p^{\parallel}$. In addition, when the space is not flat, the equalities $\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \mathbf{v}$ do not holds.

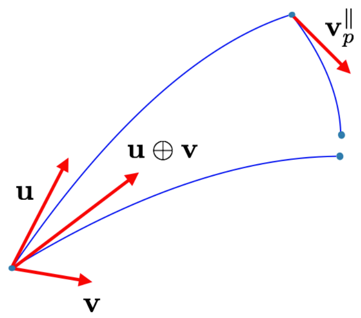

Figure 2.4: Representation of the intuitive idea of the computation of the log-composition using parallel transport.

The validity of the second assumption must be investigated case by case. For example when $\mathbb{G}$ is a matrix Lie group, the formula 2.2 provides $\exp(\mathbf{u}) = I + \mathbf{u} + \mathcal{O}(\mathbf{u}^2)$. In the case of stationary velocity field, we have that the condition hold when $\mathbf{u}$ is small enough (see proposition 8.6 pag. 163 [You10]). More on this will be presented in 3.2.5.

Assuming the validity of these assumptions and from equation 2.10 it follows that

$$\begin{aligned} \mathbf{u} \oplus \mathbf{v} &\simeq \mathbf{u} + \mathbf{v}_e^{\parallel} \\ e + \mathbf{v}_e^{\parallel} &\simeq \exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big) \end{aligned}$$

Therefore

$$\mathbf{u} \oplus \mathbf{v} \simeq \mathbf{u} + \exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big) - e \tag{2.14}$$

When working in the infinite dimensional case, the approximation 2.14 holds under the following additional assumption:

3. Theorem 2.3.1 holds when the Lie group is infinite dimensional.

An eventual confirmation is at the moment not known to the author. We assume it is true in coherence with what has been said in the introduction, section **??** about a naive approach to the infinite dimensional Lie Group theory.

With the truncated BCH and the Taylor expansion, equation 2.14 is the third numerical method for the computation of the log-composition explored in this thesis. The next chapter is devote to introduce two group of transformation - the rigid body transformation and the diffeomorphisms - and to apply the numerical methods presented in this chapter to these cases.

Chapter 3

Spatial Transformations for the Computations of the Log-composition: $SE(2)$ and Diff(Ω)

Every working mathematician knows that if one does not control oneself (best of all by examples), then after some ten pages half of all the signs in formulae will be wrong and twos will find their way from denominators into numerators.
-V.I. Arnold

In the previous chapter we have introduced some essential mathematical tools for the numerical computation of the log-composition. Each of the theoretical elements depends strongly on the transformations considered, and in this chapter we will see how they can be applied for the transformations belonging to $SE(2)$ and Diff(Ω).

3.1 The Lie Group of Rigid Body Transformations

Each element of the group of rigid body transformation (or euclidean group) $SE(2)$ can be computed as the consecutive application of a rotation and a translation applied to any point $(x, y)^T$ of the plane:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = R(\theta) \begin{pmatrix} x \\ y \end{pmatrix} + t = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t^x \\ t^y \end{pmatrix}$$

where the rotation matrix indicated with $R(\theta)$ belongs to the special orthogonal group $SO(2)$ and the translation t is a vector of the plane.
We can represent the elements of $SE(2)$ in two different form: as ternary vector (restricted form)

$$SE(2)^v := \{(\theta, t^x, t^y) \mid \theta \in [0, 2\pi), t^x, t^y \in \mathbf{R}^2\}$$

or with matrices (matrix form)

$$SE(2) := \left\{ \begin{pmatrix} R(\theta) & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & t^x \\ \sin(\theta) & \cos(\theta) & t^y \\ 0 & 0 & 1 \end{pmatrix} \mid \theta \in [0, 2\pi), (t^x, t^y) \in \mathbf{R}^2 \right\}$$

The group $SE(2)$ it is a manifold with a differentiable structure compatible with the operation of composition, whose Lie algebra is given in matrix form by

$$\mathfrak{se}(2) := \left\{ \begin{pmatrix} dR(\theta) & dt \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\theta & dt^x \\ \theta & 0 & dt^y \\ 0 & 0 & 0 \end{pmatrix} \mid \theta \in [0, 2\pi), (dt^x, dt^y) \in \mathbf{R}^2 \right\}$$

and it is indicated with $\mathfrak{se}(2)^v$ in its restricted form.

Given r, element of $SE(2)$ with $\theta \neq 0$, its image with the Lie group logarithm is

$$\begin{aligned} \log(r) &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(r-I)^k}{k} = \begin{pmatrix} dR(\theta) & L(\theta)t \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\theta & \frac{\theta}{2}(\frac{\sin(\theta)}{1-\cos(\theta)} t^x + t^y) \\ \theta & 0 & \frac{\theta}{2}(-t^x + \frac{\sin(\theta)}{1-\cos(\theta)} t^y) \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

where

$$dR(\theta) = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \qquad L(\theta) = \frac{\theta}{2} \begin{pmatrix} \frac{\sin(\theta)}{1-\cos(\theta)} & 1 \\ -1 & \frac{\sin(\theta)}{1-\cos(\theta)} \end{pmatrix}$$

On the way back, the exponential of $dr \in \mathfrak{se}(2)$ is given by:

$$\begin{aligned} \exp(dr) &= \sum_{k=1}^{\infty} \frac{dr^k}{k!} = \begin{pmatrix} R(\theta) & L(\theta)^{-1} dt \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta) & -\sin(\theta) & \frac{1}{\theta}(\sin(\theta) dt^x - (1-\cos(\theta)) dt^y) \\ \sin(\theta) & \cos(\theta) & \frac{1}{\theta}(-(1-\cos(\theta)) dt^x + \sin(\theta) dt^y) \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

where

$$L(\theta)^{-1} = \frac{1}{\theta} \begin{pmatrix} \sin(\theta) & -(1-\cos(\theta)) \\ (1-\cos(\theta)) & \sin(\theta) \end{pmatrix}$$

When θ is zero, $R(\theta)$ and $dR(\theta)$ coincide with the identity, and the transformation results in a translation. For proof and further details see for example [Gal11] [Hal15].

At this point it is important to notice that:

1. The infinite series of matrices do not raises any theoretical issues, since the sum is defined in the group as subset of a bigger algebra that contains both the Lie group and the Lie algebra. It appears to be the natural way to move back and forth from the group to the algebra. A second door to passing from one structure to the other, when the rotation θ is small is provided by the following approximations:

$$\exp(r) \simeq I + r \qquad \log(dr) \simeq dr - I \tag{3.1}$$

   In fact for small θ, $\sin(\theta) \simeq \theta$, $\cos(\theta) \simeq 0$ and $L(\theta) \simeq I$.

2. The map exp is not well defined as bijection over its whole domain $\mathfrak{se}(2)$. Given two elements $(\theta_0, dt_0^x, dt_0^y)$ and $(\theta_1, dt_1^x, dt_1^y)$, they have the same image with exp function if the two following conditions are both satisfied:

   i) Exists an integer k such that $\theta_0 = \theta_1 + 2k\pi$.

   ii) the translation (dt_0^x, dt_0^y) coincides with (dt_1^x, dt_1^y) up to a factor $\frac{\theta_0}{\theta_1}$, where the angles are considered modulo 2π.

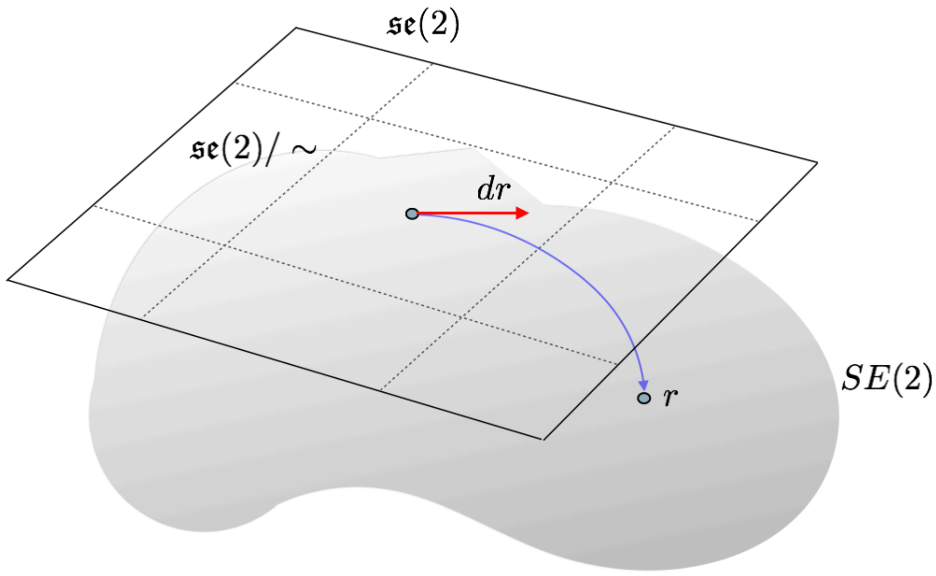

Figure 3.1: The Lie algebra $\mathfrak{se}(2)/\sim$ defined as the quotient of the Lie algebra $\mathfrak{se}(2)$ over the equivalence relation $\sim$ is in bijective correspondence with $SE(2)$.

To have a bijective correspondence the domain of exp has to be restricted to a space where if $\exp(\theta_0, dt_0^x, dt_0^y) = \exp(\theta_1, dt_1^x, dt_1^y)$ implies $(\theta_0, dt_0^x, dt_0^y) = (\theta_1, dt_1^x, dt_1^y)$. It can be easy to prove that the sought space is the quotient of $\mathfrak{se}(2)$ over the equivalence relation $\sim$, defined as

$$\begin{aligned}(\theta_0, dt_0^x, dt_0^y) &\sim (\theta_1, dt_1^x, dt_1^y) \\ &\iff \text{(by definition)} \\ \exists k \in \mathbb{Z} \mid \theta_0 = \theta_1 + 2k\pi \quad \text{and} \quad (dt_0^x, dt_0^y) &= \frac{\theta_0}{\theta_1}(dt_1^x, dt_1^y)\end{aligned}$$

The new algebra defined by the set of equivalence classes of this relation is indicated - with the standard convention, see [Art11] - with $\mathfrak{se}(2)/\sim$. With this restriction of the domain, the function exp is a bijection having log as its inverse. What said so far can be summarize in the following commutative diagram:

$$\begin{tikzcd} \mathfrak{se}(2) \arrow[dr, two heads, "\pi"] & \\ & \mathfrak{se}(2)/\sim \arrow[dl, "\exp"] \\ SE(2) \arrow[uu, "\log"] & \end{tikzcd}$$

and with the schematic figure 3.1.

3.1.1 Computations of Log-composition in $\mathfrak{se}(2)$

The log-composition of two elements $dr_0 = (\theta_0, dt_0^x, dt_0^y)$ ans $dr_1 = (\theta_1, dt_1^x, dt_1^y)$ of $\mathfrak{se}(2)/\sim$ results

$$dr_0 \oplus dr_1 = \log(\exp(dr_0) \circ \exp(dr_1)) \tag{3.2}$$

The approximations of the log-composition using truncated BCH formulas are straightforward:

$$\begin{aligned}
&dr_0 \oplus dr_1 \simeq BCH^0(dr_0, dr_1) := dr_0 + dr_1\\
&dr_0 \oplus dr_1 \simeq BCH^1(dr_0, dr_1) := dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1]\\
&dr_0 \oplus dr_1 \simeq BCH^{3/2}(dr_0, dr_1) := dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1] + \frac{1}{12}[dr_0, [dr_0, dr_1]]\\
&dr_0 \oplus dr_1 \simeq BCH^2(dr_0, dr_1) := dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1] + \frac{1}{12}([dr_0, [dr_0, dr_1]] + [dr_1, [dr_1, dr_0]])
\end{aligned}$$

To compute the approximation with the Taylor method, and so to compute the equation 2.13 for elements in $\mathfrak{se}(2)/\sim$, we observe that the restricted form of the Lie bracket is given by

$$\begin{aligned}
[dr_0, dr_1] &= (0, dR(\theta_0)dt_1 - dR(\theta_1)dt_0)^T\\
&= (0, -\theta_0 dt_1^y + \theta_1 dt_0^y, \theta_0 dt_1^x - \theta_1 dt_0^x)^T
\end{aligned}$$

Therefore, the adjoint operator can be written in matrix form as a dual matrix of dr:

$$\mathrm{ad}_{dr} = \begin{pmatrix} 0 & 0 & 0 \\ dt^y & 0 & -\theta \\ -dt^x & \theta & 0 \end{pmatrix}$$

In fact, when applied to dr_1 it results in the Lie bracket:

$$\mathrm{ad}_{dr_0} dr_1 = \begin{pmatrix} 0 & 0 & 0 \\ dt_0^y & 0 & -\theta_0 \\ -dt_0^x & \theta_0 & 0 \end{pmatrix} \begin{pmatrix} \theta_1 \\ dt_1^x \\ dt_1^y \end{pmatrix} = \begin{pmatrix} 0 \\ -\theta_0 dt_1^y + \theta_1 dt_0^y \\ \theta_0 dt_1^x - \theta_1 dt_0^x \end{pmatrix}$$

To compute the Taylor approximation proposed in equation 2.13 of the log composition, indicating $dt^\star = (dt^y, -dt^x)$ it can be proved easily by induction that

$$\mathrm{ad}_{dr}^n = \begin{pmatrix} 0 & 0 \\ dt^\star & dR(\theta) \end{pmatrix}^n = \begin{pmatrix} 0 & 0 \\ dR(\theta)^{n-1} dt^\star & dR(\theta)^n \end{pmatrix}$$

And so the series involved in the equation 2.13 become

$$\sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_{dr}^n = \sum_{n=0}^{\infty} \frac{B_n}{n!} \begin{pmatrix} 0 & 0 \\ dR(\theta)^{n-1} dt^\star & dR(\theta)^n \end{pmatrix}$$

We can split it in two part, the rotational part $dR(\theta)^n$ and the translational part $dR(\theta)^{n-1}dt^\star$. The rotational part, exploiting the nature of Bernoulli numbers and its generative equation, when $\theta \neq 0$ become

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{B_n}{n!} dR(\theta)^n &= I + \frac{1}{2} dR(\theta) + \sum_{n=1}^{\infty} \frac{B_{2n}}{2n!} dR(\theta)^{2n} \\
&= I + \frac{1}{2} dR(\theta) + (\sum_{n=1}^{\infty} \frac{B_{2n}}{2n!} (i\theta)^{2n}) I \\
&= \frac{1}{2} dR(\theta) + (\sum_{n=0}^{\infty} \frac{B_n}{n!} (i\theta)^n - \frac{1}{2} i\theta) I \\
&= \frac{1}{2} dR(\theta) + (\frac{i\theta e^{i\theta}}{e^{i\theta} - 1} - \frac{1}{2} i\theta) I \\
&= \frac{1}{2} dR(\theta) + \frac{\theta/2}{\tan(\theta/2)} I
\end{aligned}$$

where the equation $dR(\theta)^{2n} = (i\theta)^{2n} I$. For the translational part we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{B_n}{n!} dR(\theta)^{n-1} dt^\star &= dR(\theta)^{-1} \Big(\sum_{n=1}^{\infty} \frac{B_n}{n!} dR(\theta)^n \Big) dt^\star \\
&= dR(\theta)^{-1} \Big(\sum_{n=0}^{\infty} \frac{B_n}{n!} dR(\theta)^n - I \Big) dt^\star \\
&= dR(\theta)^{-1} \Big(\sum_{n=0}^{\infty} \frac{1}{2} dR(\theta) + \frac{\theta/2}{\tan(\theta/2)} I - I \Big) dt^\star \\
&= dR(\theta)^{-1} \Big(\sum_{n=0}^{\infty} \frac{1}{2} dR(\theta) + \frac{\theta/2}{\tan(\theta/2)} I - I \Big) dt^\star \\
&= \Big(\frac{1}{2} I + (\frac{\theta/2}{\tan(\theta/2)} - 1) dR(\theta)^{-1} \Big) dt^\star
\end{aligned}$$

Finally the closed form for the Taylor approximation of the log-composition is [Ver14]:

$$dr_0 \oplus dr_1 = dr_0 + \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}^n_{dr_0} dr_1 + \mathcal{O}(dr_1^2) = dr_0 + \mathbf{J}(dr_0) dr_1 + \mathcal{O}(dr_1^2) \tag{3.3}$$

where

$$\mathbf{J}(dr_0) = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\theta_0/2 - \tan(\theta_0/2)}{\theta_0 \tan(\theta_0/2)} dt_0^x + \frac{1}{2} dt_0^y & \frac{\theta_0/2}{\tan(\theta_0/2)} & -\theta_0/2 \\ -\frac{1}{2} dt_0^x - \frac{\theta_0/2 - \tan(\theta_0/2)}{\theta_0 \tan(\theta_0/2)} dt_0^y & \theta_0/2 & \frac{\theta_0/2}{\tan(\theta_0/2)} \end{pmatrix}$$

therefore the corresponding numerical method indicated with the function Tl as

$$dr_0 \oplus dr_1 \simeq Tl(dr_0, dr_1) := dr_0 + \mathbf{J}(dr_0) dr_1 \tag{3.4}$$

The approximation of the log-composition using parallel transport is a straightforward application of the equation 2.14:

$$dr_0 \oplus dr_1 \simeq pt(dr_0, dr_1) := dr_0 + \exp\big(\frac{dr_0}{2}\big) \exp(dr_1) \exp\big(-\frac{dr_0}{2}\big) - I \tag{3.5}$$

where the composition in the Lie group coincides with the product of matrix in the bigger algebra $GL(3)$ that contains both the Lie group $SE(2)$ and the Lie algebra $\mathfrak{se}(2)$.

3.2 The Lie group of Diffeomorphisms

The passage from the finite to the infinite dimensional case is not free of deceptions. We will investigate in the next two subsections, 3.2.1 and 3.2.2 ,the following facts that are true matrices but not for diffeomorphisms:

1. Lie logarithm and Lie exponential are local isomorphisms.

2. $SE(2)$ and $\mathfrak{se}(2)$ are subset of a bigger algebra, where all of the operations are compatible.

Before moving toward these concepts, we need to clarify some definitions and notations.

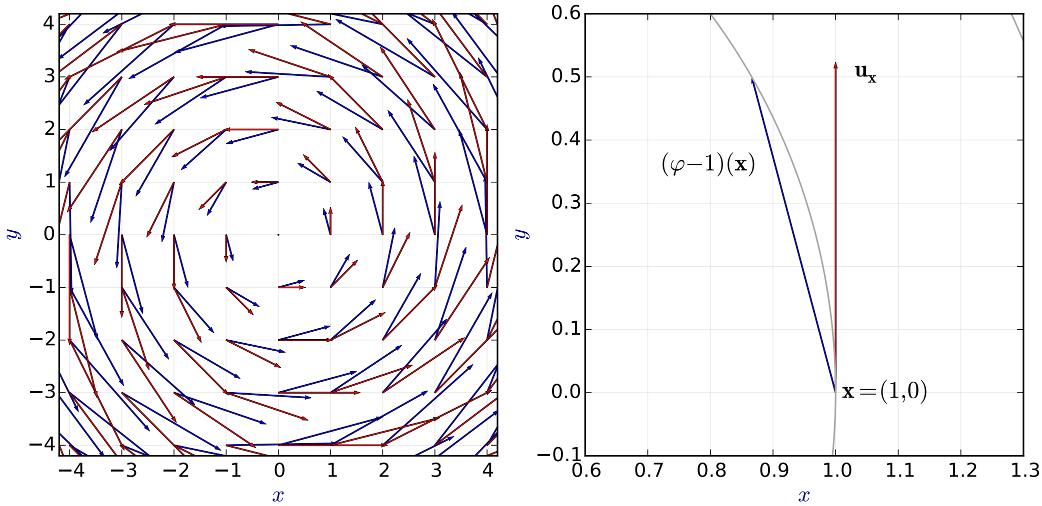

Figure 3.2: the displacement field and the tangent vector field for the transformation φ defined as a rotation of $\pi/6$ around the origin. When φ is subtracted by the identity function, become an element of the algebra of the velocity vector fields $\mathrm{Vect}(\Omega)$.

We define the set of *deformations*, the set of continuous functions from Ω to Ω, compact subset of $\mathbb{R}^d$. If a deformation is invertible with continuous inverse, then it is called *homeomorphism*; the set of homeomorphisms forms a group, indicated with $\mathrm{Hom}(\Omega)$, with the operation of function composition. If an homeomorphism is differentiable and has differentiable inverse then it is called *diffeomorphism*. Again the set of diffeomorphisms forms a group, indicated with $\mathrm{Diff}(\Omega)$.

A *velocity vector field* over Ω is a differentiable function that at each point of Ω, associates a vector of $\mathbb{R}^d$; the set of velocity vector fields, indicated with $\mathrm{Vect}(\Omega)$ forms a vector space, and considering the Lie bracket defined by the directional derivative we obtain that $\mathrm{Vect}(\Omega)$ forms a Lie algebra[1]. If $\{\frac{\partial}{\partial x_i}\}_{i=1}^d$ is a local coordinates system over Ω, $\mathbf{u} = a^i \frac{\partial}{\partial x_i}$ and $\mathbf{v} = b^i \frac{\partial}{\partial x_i}$ are two elements of $\mathrm{Vect}(\Omega)$ written using the Einstein summation convention,

[1] Some books invert the signs of the operation (see the Kirillow's remarks [Kir08] pag. 27); this choice do not have any impact in the study of the algebraic structure, but it does have an impact on the numerical results when the Lie brackets are implemented for numerical computations. At the moment the sign that defines the Lie bracket is chosen on the base of the obtained results.

the Lie bracket can be expressed using the Jacobian:

$$\begin{aligned}
[\mathbf{u},\mathbf{v}] &= a^i \frac{\partial}{\partial x_i}\big(b^j \frac{\partial}{\partial x_j}\big) - b^i \frac{\partial}{\partial x_i}\big(a^j \frac{\partial}{\partial x_j}\big) \\
&= a^i \frac{\partial b_j}{\partial x_i}\frac{\partial}{\partial x_j} + a^i b^j \frac{\partial^2}{\partial x_i x_j} - b^j \frac{\partial a^i}{\partial x^j}\frac{\partial}{\partial x_i} - b^j a^i \frac{\partial^2}{\partial x_j x_i} \\
&= a^i \frac{\partial b_j}{\partial x_i}\frac{\partial}{\partial x_j} - b^j \frac{\partial a^i}{\partial x^j}\frac{\partial}{\partial x_i} = J_{\mathbf{v}}\mathbf{u} - J_{\mathbf{u}}\mathbf{v}
\end{aligned}$$

It was proved that the Lie algebra of vector fields have, as its Lie group, the group of diffeomorphisms ([Mil82], [OKC92]).

A single vector in the Lie algebra Vect(Ω) (represented by one red arrow in figure 1.1) is a vector field defined over Ω (represented by the set of red arrows in figure 3.2). We would expect that, vice versa, to a single diffeomorphism in the Lie group corresponds a single vector of the Lie algebra. This is not the case, since Lie logarithm and Lie exponential are not local isomorphisms on the whole domain.

3.2.1 Local isomorphisms for a subset of Diffeomorphisms: one-parameter subgroup and stationary velocity fields

In the case of matrices, the exponential map is a local isomorphisms: it is always possible to find an open neighbor of $\mathbf{0}$ in the Lie algebra and an open neighbor of the identity element in the Lie group (in the same topology induced by the metric inherited by the bigger algebra), such that the exponential map is defined and invertible.

In the infinite dimensional case there are diffeomorphisms arbitrarily close to the identity that are not embedded to any one-parameter subgroups, therefore the exponential map is not a local isomorphism (see the counterexample in [Mil84b], pag. 1017 or the definition of Koppel-diffeomorphisms [Gra88] pag. 115).

Since for medical image registration we are interested only in the diffeomorphisms that can be parametrized by tangent vector fields, this feature is worthed to be investigated, but it requires some definitions.

If φ is a one-parameter subgroup on the manifold *Diff* (Ω), then its derivative satisfies the *stationary* (or homogeneous) ordinary differential equation:

$$\frac{d\varphi(t)}{dt} = V_{\varphi(t)} \tag{3.6}$$

Where the stationary vector field $V_{\varphi(t)}$ defined over Ω is an element of the Lie algebra of $\mathcal{V}(\Omega)$ called *stationary velocity field* or SVF. In fact

$$\frac{d\varphi(t)}{dt} = \lim_{\epsilon\to 0}\frac{\varphi(t+\epsilon)}{\epsilon} = \lim_{\epsilon\to 0}\frac{\varphi(\epsilon)\varphi(t)}{\epsilon} = V_{\varphi(t)}$$

Vice versa, given an SVF, thanks to Cauchy theorem exists always a unique solution φ to the ODE 3.6, given the initial condition $\varphi(0) = 1$, that satisfies the property of one-parameter subgroup.

We indicate with *Diff* $^1(\Omega)$ the *set of diffeomorphisms embedded in a one parameter subgroup*, i.e. the solutions of 3.6. We notice that *Diff* $^1(\Omega)$ does not form a group. In fact if φ and ψ are in *Diff* $^1(\Omega)$ and satisfy respectively $\frac{d\varphi_1(t)}{dt} = U_{\varphi_1(t)}$ and $\frac{d\varphi_2(t)}{dt} = V_{\varphi_2(t)}$, then their composition $\varphi_1 \circ \varphi_2$ does not satisfy any stationary ordinary differential equation. To have closure for the composition of one parameter subgroup, we have to extend our attention to non stationary (or non homogeneous) ordinary differential equation of the form:

$$\frac{d\psi(t)}{dt} = W_{(t,\psi(t))} \tag{3.7}$$

Where $W_{(t,\psi(t))}$ is a non-stationary vector field, called here time varying vector field, or TVVF. If compared with to the SVF, it does not depends only on the spatial position $\mathbf{x}$ but there is also a temporal dependency.

Think for example to a satellite orbiting around the globe: it is subject to the earth's vector field in respect to which it is constant for a fixed position, and to the lunar vector field that it is not fixed but varies in respect to the time. Conventionally the temporal domain T contains the origin and formally we can write:

$$\begin{aligned} W : T \times \Omega &\longrightarrow \mathbb{R}^d \\ t, \psi(t) &\longmapsto W_{(t,\psi(t))} \end{aligned}$$

for ψ diffeomorphism (or in the previous example, position of the satellite at time t) that when applied to a point of Ω is indicated with $\varphi(t, \mathbf{x})$ or $\psi^{(t)}(\mathbf{x})$.

A crucial observation for our purpose is that non-autonomous ODE are particular cases of autonomous one. Writing the diffeomorphism $\psi(t)$ applied to $\mathbf{x}$ in local coordinates as

$$\psi^{(t)}(\mathbf{x}) = (\psi_1^{(t)}(\mathbf{x}), \psi_2^{(t)}(\mathbf{x}), \dots, \psi_d^{(t)}(\mathbf{x})) \in \mathbb{R}^d$$

Defining a new function $\psi_0^{(t)}(\mathbf{x}) = t_0 + t$ for all $\mathbf{x} \in \Omega$, we can obtain then the new diffeomorphism $\tilde{\psi}^{(t)}$ that in local coordinates is expressed as

$$\tilde{\psi}^{(t)}(\mathbf{x}) = (\psi_0^{(t)}(\mathbf{x}), \psi_1^{(t)}(\mathbf{x}), \psi_2^{(t)}(\mathbf{x}), \dots, \psi_d^{(t)}(\mathbf{x})) \in T \times \mathbb{R}^d$$

that reduces the ODE 3.7 to an ODE of the form 3.6. In the example of satellite, is like considering the temporal dimension as an additional dimension of the space. The vector that influence the satellite is an SVF for every point in the domain of space-time.

It follows that tationary ODE and non-stationary ODE have solutions that belong to $\mathit{Diff}^1(\Omega)$ and $\mathit{Diff}^1(T \times \Omega)$ respectively. For each instant of time the solution of non-stationary ODE, are embedded in the set of one-parameter subgroup of $\mathit{Diff}(\Omega)$, but for two different instant of time, the solution can belongs to two different one parameter subgroups.

In conclusion, we have that there in the case of diffeomorphisms exp is not a local isomorphism, unless we do not restrict the group of diffeomorphisms to the one embedded in a one parameter subgroup $\mathit{Diff}^1(T \times \Omega)$. In addition the set of diffeomoerphisms restricted to the one that solves the equation 3.6 does not form any group with the composition. This happen only if we extend to the solution of the non stationary ODE 3.7, and therefore to TVVF. In addition, indicating with SVF the set of stationary velocity fields and with TVVF the set of time varying velocity fields, we have that

$$\mathit{Diff}^1(\Omega) = \exp(\mathrm{SVF}) = \exp(\mathrm{TVVF})$$

but to a given SVF exists only one one-parameter subgroup φ that satisfies the ODE 3.6. The same thing does not necessarily happens for the TVVF.

In the LDDMM framework (briefly mentioned in section **??**) TVVFs are initially considered, while with the paper of Arsigny [ACPA06], and in subsequent works, the attention has been restricted to SVF, in order to be able to use the scaling and squaring and the inverse scaling and squaring algorithms for the numerical computation of the Lie exponential and the Lie logarithm. In fact the scaling and squaring method, as every numerical method based on the phase flow [YC06], works only under the assumption that the transformation belongs to the same one-parameter subgroup.

3.2.2 A bigger algebra for the group of Diffeomorphisms

As well as for any matrix Lie group, both the group $SE(2)$ and the algebra $\mathfrak{se}(2)$ are subset of the same bigger algebra of matrices in the general linear group $GL(3, \mathbb{R})$. The product of the algebra coincides with the composition of the group and thanks to the linearity, scalar product is compatible both with the product and the composition.

The existence of a bigger algebra is not important only in the research of an elegant structure: the power series expansions of the exponential (2.2) and the logarithm (2.3) as well as expressions like (2.4) and (2.5) would be meaningless without the possibility of expressing the sum of two elements of a multiplicative group. Moreover, if the bigger algebra that contains both Lie group and Lie algebra exists, a unique norm in this space can be defined and utilized to compare elements in the both subspaces.

In the case of diffeomorphism of the compact subset Ω of $\mathbb{R}^d$, we can identify a bigger vector space that contains both Lie group and Lie algebra, but it is less straightforward than in the case of matrices, and for this aim it is necessarily to have some definitions at hand.

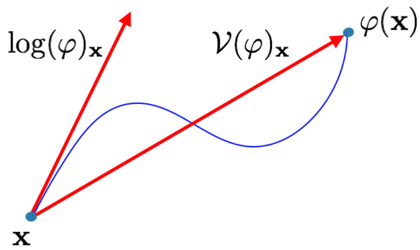

Figure 3.3: for small deformations, the displacement field $\log(\varphi)$ and the tangent field $\mathcal{V}(\varphi)$, computed at the point $\mathbf{x}$ of Ω, are close to each others.

There are two ways to associate a diffeomorphisms φ to a velocity vector field. The first one is elementary but fundamental in this context: it consists in subtracting to φ the identity function 1. If $\mathbf{x}$ is in Ω and $\varphi(\mathbf{x})$ is the new point after the transformation, then the associated velocity vector field, called here *displacement field of* φ, is the function that at the point $\mathbf{x}$ associate the vector defined as the difference $\varphi(\mathbf{x}) - \mathbf{x}$. To recover the deformation from a velocity field $\mathbf{u}$ is enough to add the identity; in this case we have the *deformation of* $\mathbf{u}$. We indicate this operation of adding and subtracting the identity with the function $\mathcal{V}$:

$$\mathcal{V}(\varphi) = \varphi - 1 \qquad\qquad \mathcal{V}^{-1}(\mathbf{u}) = \mathbf{u} + 1$$

We can see that displacement fields of diffeomorphisms are elements of $\mathcal{V}(\Omega)$, that is the analogous of the bigger algebra that contains Lie group and Lie algebra in the case of matrices. We can observe that this operation of subtracting the identity to the deformation has already been used implicitly in the power series expansion of the Lie logarithm for matrices, see equation 2.3.

The second way to associate a velocity vector field to φ is with the Lie logarithm defined in the chapter 2. It is interesting to notice that when $\mathbf{u}$ is small then $\mathcal{V}^{-1}$ and exp are closed to each other and $\mathcal{V}^{-1}$ can be considered a good approximation of exp (see figure 3.3). The very same happens for matrices, as noticed in equations (3.1).

At this point it is important to notice that, while a displacement field of φ can always be defined, the exponential map it is not defined for any diffeomorphism. This is the second remarkably difference between the matrix Lie group and the Lie group of diffeomorphisms that will be investigated in the next section.

3.2.3 Norm for elements in the one-parameter subgroup

A metric between tangent vector fields of Ω can be defined as

$$d(\mathbf{u},\mathbf{v}) = \Big(\int_{\Omega} ||\mathbf{u}-\mathbf{v}||^2_{L^2} d\mathbf{x} \Big)^{1/2} \tag{3.8}$$

and induces a metric to compute the distance between stationary velocity fields.

The Lie group $Diff^1(\Omega)$ do not possess any norm, but the corresponding displacement fields defined by $\mathcal{V}$, as tangent vector fields does. Given two diffeomorphisms φ_0, φ_1 we have

$$d^1(\varphi_0,\varphi_1) = \Big(\int_{\Omega} ||\mathcal{V}(\varphi_0) - \mathcal{V}(\varphi_1)||^2_{L^2} d\mathbf{x} \Big)^{1/2} \tag{3.9}$$

Despite the limitation that Lie algebra and Lie group of diffeomorphisms are not subset of the same bigger algebra, we can nevertheless consider a function that measure the best approximation of metric we can have for $Diff^1(\Omega)$ and SVF:

$$d^m(\mathbf{u},\varphi) = \Big(\int_{\Omega} ||\mathbf{u} - \mathcal{V}(\varphi)||^2_{L^2} d\mathbf{x} \Big)^{1/2} \tag{3.10}$$

The next section is about the parametrization of SVF in the applications, and it is followed by the one that presents the numerical methods for the log-composition when applied to SVF.

3.2.4 Parametrization of SVF: Grids and Discretized Vector Fields

Even if images are discrete elements, the underpinning model of the transformations is based on the continuous. There are several motivations that led to this choice: as underlined by [Sze94], the most important is that images are discrete measurement of the continuous property of an object. Therefore it is reasonable have a model as close as possible to the continuous object rather than to a set of discrete measurements. Certainly it is important to keep in mind the fact that the continuous approximation is obtained - in a non unique way - from the discretized image with an interpolation scheme. This imply that, for example if the distance between two separate objects is less than the size of a voxel, in continuous approximation based on the discretized image the two object will be not anymore separated.

Also transformations between images are discretized vector fields, where each vector is applied to an element of a grid. These transformations can only be considered as a model of the group of diffeomorphisms (a model of a model, in image registration!) and reflects only partially the continuous property of the original transformation. On the other side the possibility of working with discretized elements means working with something that can be managed by computers.

As in many implementation, the data structure utilized to store images, as well as displacement fields are 5-dimensional matrices

$$M = M(x_i, y_j, z_k, t, d) \qquad (i,j,k) \in L, \quad t \in T \quad d = 1,2,3 \tag{3.11}$$

where (x_i, y_j, z_k) are discrete position of a lattice L in the domain of the images, t is the time parameter in a discretized domain T and d is index of the coordinate axis. So, the discretized *tangent vector* $\mathbf{v}_\tau(x_i, y_j, z_k)$ at time t, has coordinates defined by

$$\mathbf{v}_t(x_i, y_j, z_k) = (M(x_i, y_j, z_k, t, 1), M(x_i, y_j, z_k, t, 2), M(x_i, y_j, z_k, t, 3))$$

3.2.5 Computations of Log-composition for SVF

A closed-form for the Taylor Expansion method 2.4.2 to compute the log-composition with elements in $Diff^1(\Omega)$ is not known. We will therefore compare the truncated BCH formula with the parallel transport method 2.3.1. The Lie bracket that appears of SVF in the truncated BCH of degree $0, 1, 1.5$ and 2, are computed using the Jacobian matrix J:

$$[\mathbf{u}, \mathbf{v}] := J_u\mathbf{v} - J_u\mathbf{v} \qquad \forall \mathbf{u}, \mathbf{v} \in \mathfrak{g} \tag{3.12}$$

as a consequence of its definition (see [Lee12]). It has been shown that this definition is uniquely defined as action on the space of $\mathbb{C}^\infty$ function on the same domain and it satisfies the axioms of Lie bracket of a Lie algebra.

Therefore the truncated approximation of the BCH formula presented in the equation 2.12 become:

$$\begin{aligned}
BCH^0(\mathbf{u}, \mathbf{v}) &= \mathbf{u} + \mathbf{v} \\
BCH^1(\mathbf{u}, \mathbf{v}) &= \mathbf{u} + \mathbf{v} + \frac{1}{2}(J_\mathbf{u}\mathbf{v} - J_\mathbf{u}\mathbf{v}) \\
BCH^{3/2}(\mathbf{u}, \mathbf{v}) &= \mathbf{u} + \mathbf{v} + \frac{1}{2}(J_u\mathbf{v} - J_u\mathbf{v}) + \frac{1}{12}\big(2J_\mathbf{u}J_\mathbf{u}\mathbf{v} + 2J_\mathbf{u}J_\mathbf{v}\mathbf{u} - J_{(J_\mathbf{u}\mathbf{v} - J_\mathbf{u}\mathbf{v})}\mathbf{u}\big) \\
BCH^2(\mathbf{u}, \mathbf{v}) &= \mathbf{u} + \mathbf{v} + \frac{1}{2}(J_u\mathbf{v} - J_u\mathbf{v}) \\
&\quad + \frac{1}{12}\big(2J_\mathbf{u}J_\mathbf{u}\mathbf{v} + 2J_\mathbf{u}J_\mathbf{v}\mathbf{u} - J_{(J_\mathbf{u}\mathbf{v} - J_\mathbf{u}\mathbf{v})}\mathbf{u} + 2J_\mathbf{v}J_\mathbf{v}\mathbf{u} + 2J_\mathbf{v}J_\mathbf{u}\mathbf{v} - J_{(J_\mathbf{v}\mathbf{u} - J_\mathbf{v}\mathbf{u})}\mathbf{v}\big)
\end{aligned}$$

Lie brackets of SVF can become extremely small, in particular, as we will see in the last chapter, when the standard deviation of the Gaussian filter that generates the fields is small. Whether it is not known how to apply Taylor method presented in 2.4.2 for the SVF, the parallel transport method for the computation of the log-composition follows directly from equation 2.14:

$$\mathbf{u}_0 \oplus \mathbf{u}_1 \simeq \mathbf{u}_0 + \exp_e\big(\frac{\mathbf{u}_0}{2}\big) \circ \exp_e(\mathbf{u}_1) \circ \exp_e\big(-\frac{\mathbf{u}_0}{2}\big) - e$$

Here the exponential function can be computed with several algorithms (scaling and squaring, forward Euler, composition method, Taylor expansion, see [BZO08] for a comparison of their performances). Following the original setting of the Log-euclidean metric proposed in [ACPA06] we use the scaling and squaring, keeping in mind that this choice impact on the results.

Chapter 4

Log-composition to Compute the Lie Logarithm

I think you might do something better with the time
than wasting it in asking riddles that have no answers.
-*Alice in Wonderland.*

The *logarithm computation problem* can be stated as follows:

Given p in a Lie group $\mathbb{G}$,
what is the element $\mathbf{u}$ *in its Lie algebra* $\mathfrak{g}$
such that $\exp(\mathbf{u}) = p$ *?*

There are several numerical methods to compute the approximation of the problem's solution. Arsigny, who first pointed the applications of the Lie logarithm in medical image registration in [AFPA06] and [APA06], proposed the Inverse scaling and squaring (see also [YC06]). In this chapter we investigate other numerical iterative algorithms for the computation of the Lie logarithm, called here *logarithm computation algorithm*; they modifications of the algorithm presented in [BO08] that is based on the BCH formula, and so on the log-composition. Each of the numerical method to compute the log-composition become naturally a numerical method for the computation of the logarithm computation algorithm.
The first step toward this direction is to introduce the space of the approximations of a Lie algebra and a the Lie group.

4.1 Spaces of Approximations

As seen in section 3.1 and 3.2, if the element $\mathbf{u}$ of $\mathfrak{se}(2)$ or SVF is small enough we can approximate $\exp(\mathbf{u})$ with $e + \mathbf{u}$. Aim of this section is to investigate these approximations aimed to compute the logarithm.
Let $C_{\mathfrak{g}}$ and $C_{\mathbb{G}}$ the internal cut locus of a Lie algebra and a Lie group $\mathfrak{g}$ and $\mathbb{G}$ (the subset where the exponential map is well defined). We define two approximating functions:

$$\begin{aligned} \mathrm{app} : C_{\mathfrak{g}} &\longrightarrow C_{\mathfrak{g}}^{\sim} \\ \mathbf{u} &\longmapsto \exp(\mathbf{u}) - e \end{aligned}$$

$$\begin{aligned} \mathrm{App} : C_{\mathbb{G}} &\longrightarrow C_{\mathbb{G}}^{\sim} \\ \exp(\mathbf{u}) &\longmapsto e + \mathbf{u} \end{aligned}$$

Where $C_{\mathfrak{g}}^{\sim}$ is the space of approximations of elements of $C_{\mathfrak{g}}$, and $C_{\mathbb{G}}^{\sim}$ is the space of approximations of elements in $C_{\mathbb{G}}$, defined as

$$\begin{aligned} C_{\mathfrak{g}}^{\sim} &:= \{\exp(\mathbf{u}) - e \mid \mathbf{u} \in C_{\mathfrak{g}}\} \cup C_{\mathfrak{g}} \\ C_{\mathbb{G}}^{\sim} &:= \{e + \mathbf{u} \mid \mathbf{u} \in C_{\mathbb{G}}\} \cup C_{\mathbb{G}} \end{aligned}$$

In general $C_{\mathfrak{g}}^{\sim} \neq C_{\mathfrak{g}}$ and $C_{\mathbb{G}}^{\sim} \neq C_{\mathbb{G}}$, but in the considered cases of $\mathfrak{se}(2)$ and SVF, when $\mathbf{u}$ is *small enough* it follows that $\exp(\mathbf{u}) - e \in C_{\mathfrak{g}}$ and $e + \mathbf{u} \in C_{\mathbb{G}}$. Therefore the elements of $C_{\mathfrak{g}}^{\sim}$ are compatible with all of the operations of the internal cut locus of the Lie algebra $C_{\mathfrak{g}}$ and the elements of $C_{\mathbb{G}}^{\sim}$ are compatible with all of the operations of Lie group $\mathbb{G}$.
Lets examine what does *small enough* means in these two cases:

$\mathfrak{se}(2)$ - Since $\mathfrak{se}(2)$ and $SE(2)$ are subset of the bigger algebra $SE(2)$ then exp and log can be defined as infinite series. From

$$\exp(\mathbf{u}) = I + \mathbf{u} + O(\mathbf{u}^2)$$

It follows that $\text{app}(\mathbf{u}) - \mathbf{u} = O(\mathbf{u}^2)$. Thus for all $\mathbf{u}$ in the internal cut locus smaller than δ for any norm, exists $M(\delta)$ such that

$$||\text{app}(\mathbf{u}) - \mathbf{u}|| < M(\delta)||\mathbf{u}^2||$$

SVF - In case of SVF we do not have any Taylor series and big-O notation available but, according to the proposition 8.6 at page 163 of [You10], if $\mathbf{u}$ is, for any norm, smaller than $\epsilon < 1/C$, where C is the Lipschitz constant in the same norm, then $1 + \mathbf{u}$ is a diffeomorphism. With this condition holds that $C_{\text{SVF}}^{\sim} = C_{\text{SVF}}$.

Therefore, for each small enough $\mathbf{u}$ in $\mathfrak{se}(2)$ or SVF, and for the definition of the log-composition (equation 2.11) the following properties holds:

1. The approximations $\mathbf{u} \simeq \text{app}(\mathbf{u})$, $\exp(\mathbf{u}) \simeq \text{App}(\exp(\mathbf{u}))$ are bounded.
2. $\mathbf{u} = \mathbf{v} \oplus (-\mathbf{v} \oplus \mathbf{u})$
3. $\text{app}(\mathbf{v} \oplus \mathbf{u}) = \exp(\mathbf{v}) \circ \exp(\mathbf{u}) - 1 \in C_{\mathfrak{g}}^{\sim}$

With this machinery, we can finally reformulate the algorithm presented in [BO08] for the numerical computation of the Lie logarithm map using the log-composition.

4.2 The Logarithm Computation Algorithm using Log-composition

If the goal is to find $\mathbf{u}$ when its exponential is known, we can consider the sequence transformations $\{\mathbf{u}_j\}_{j=0}^{\infty}$ that approximate $\mathbf{u}$ as consequence of

$$\mathbf{u} = \mathbf{u}_j \oplus (-\mathbf{u}_j \oplus \mathbf{u}) \Longrightarrow \mathbf{u} \simeq \mathbf{u}_j \oplus \text{app}(-\mathbf{u}_j \oplus \mathbf{u})$$

This suggest that a reasonable approximation for the $(j+1)$-th element of the series can be defined by

$$\mathbf{u}_{j+1} := \mathbf{u}_j \oplus \text{app}(-\mathbf{u}_j \oplus \mathbf{u})$$

If we chose the initial value $\mathbf{u}_0$ to be zero, then the algorithm presented in [BO08] become:

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j \oplus \text{app}(-\mathbf{u}_j \oplus \mathbf{u}) \end{cases} \tag{4.1}$$

Making explicit the log-computaiton and the approximation, follows:

$$\mathbf{u}_{j+1} = \mathbf{u}_j \oplus \big(\exp(-\mathbf{u}_j) \circ \exp(\mathbf{u}) - e\big) \tag{4.2}$$

$$= \log\Big(\exp(\mathbf{u}_j) \circ \exp\big(\exp(-\mathbf{u}_j) \circ \varphi - e\big)\Big) \tag{4.3}$$

where $\exp(\mathbf{u}) = \varphi$ is given by the problem, and $\mathbf{u}_j$ by the previous step. The BCH provides the exact solution of the second member, while strategy that we have examined to compute the log-composition, become a numerical method for the computation of the logarithm.

4.2.1 Truncated BCH Strategy

At each step, we compute the approximation $\mathbf{v}_{j+1}$ with the k-th truncation of the BCH formula. The compact form of the algorithm is given by:

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathrm{BCH}^k(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) \end{cases} \tag{4.4}$$

For $k = 0$, the approximation $\mathbf{u}_{j+1}$ results simply the sum of the two vectors $\mathbf{u}_j$ and $\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})$:

$$\begin{aligned} \mathrm{BCH}^0(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) &= \mathbf{u}_j + \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) \\ &= \mathbf{u}_j + \exp(-\mathbf{u}_j) \circ \varphi - e \end{aligned}$$

When $k = 1$, it results

$$\begin{aligned} \mathrm{BCH}^1(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) &= \mathbf{u}_j + \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) + \frac{1}{2}[\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})] \\ &= \mathbf{u}_j + \exp(-\mathbf{u}_j) \circ \varphi - e + \\ &+ \frac{1}{2}(\mathbf{u}_j \cdot (\exp(-\mathbf{u}_j) \circ \varphi - e) - (\exp(-\mathbf{u}_j) \circ \varphi - e) \cdot \mathbf{u}_j) \end{aligned}$$

And for $k = 2$ it become

$$\begin{aligned} \mathrm{BCH}^2(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) &= \mathbf{u}_j + \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) + \frac{1}{2}[\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})] + \\ &+ \frac{1}{12}\Big([\mathbf{u}_j, [\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})]] + \\ &+ [\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}), [\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}), \mathbf{u}_j]]\Big) \\ &= \mathbf{u}_j + \exp(-\mathbf{u}_j) \circ \varphi - e + \frac{1}{2}[\mathbf{u}_j, \exp(-\mathbf{u}_j) \circ \varphi - e] + \\ &+ \frac{1}{12}\Big([\mathbf{u}_j, [\mathbf{u}_j, \exp(-\mathbf{u}_j) \circ \varphi - e]] + \\ &+ [\exp(-\mathbf{u}_j) \circ \varphi - e, [\exp(-\mathbf{u}_j) \circ \varphi - e, \mathbf{u}_j]]\Big) \end{aligned}$$

When considering $k = \infty$ and so, the theoretical BCH formula, the following theorem, presented in [BO08], provides an error bound:

Theorem 4.2.1 (Bossa)**.** The iterative algorithm

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j \oplus \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) \end{cases} \tag{4.5}$$

converges to $\mathbf{v}$ with error $\delta_n \in \mathbb{G}$, where

$$\delta_n := \log(\exp(\mathbf{v}) \circ \exp(-\mathbf{v}_n)) \in O(||p - e||^{2^n})$$

We observe that this upper limit can be computed only when a closed-form for the log-composition is available, as for example $\mathfrak{se}(2)$.

4.2.2 Parallel Transport Strategy

If we apply the parallel transport method for the computation of the log-composition, we obtain another version of the logarithm computation algorithm:

$$\begin{cases} \mathbf{u}_0 = \mathbf{0} \\ \mathbf{u}_{j+1} = \mathbf{u}_j + \exp(\frac{\mathbf{u}_j}{2}) \circ \exp\Big(\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})\Big) \circ \exp(-\frac{\mathbf{u}_j}{2}) - e \end{cases} \tag{4.6}$$

That is computed as:

$$\mathbf{u}_{j+1} = \mathbf{u}_j + \exp(\frac{\mathbf{u}_j}{2}) \circ \exp\Big(\exp(-\mathbf{u}_j) \circ \varphi - e\Big) \circ \exp(-\frac{\mathbf{u}_j}{2}) - e$$

We notice that mixing the operation of composition, sum and scalar product makes sense when the involved vectors are *small enough*, as stated in 4.1. Analytical computation of an upper bound error is not straightforward in this case. See section 5.6 for further details and other possible researches.

4.2.3 Symmetrization Strategy

The algorithm 4.1 could have been reformulated alternatively as $\mathbf{u}_{j+1} = \mathrm{app}(\mathbf{u} \oplus -\mathbf{u}_j) \oplus \mathbf{u}_j$. The log-composition is not symmetric therefore the two version in some cases may not return the same value. In an attempt to move toward the solution of this issue we consider

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j \oplus \frac{1}{2}\big(\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) + \mathrm{app}(\mathbf{u} \oplus -\mathbf{u}_j)\big) \end{cases} \tag{4.7}$$

Writing directly the approximations and using the BCH approximation of degree 1 it become:

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j + \frac{1}{2}\big(\exp(-\mathbf{u}_j) \circ \varphi) - e + \varphi \circ \exp(-\mathbf{u}_j) - e\big) \end{cases} \tag{4.8}$$

Experimental results of the methods presented in this section are presented in the next chapter.

Chapter 5

Experimental Results

"A victory is twice itself when the achiever brings home full numbers."
Much ado about nothing*, Leonato, scene 1.*

This chapter is devoted to show most relevant results of the numerical methods investigated for the computation of the log-composition.
Computations are performed with a software written in Python (repository available on the UCL CMIC gitlab https://cmiclab.cs.ucl.ac.uk), based on the following libraries - numpy, matplotlib [Hun07], math, scipy [JOP+], nibabel, timeit, random - as well as on the library NiftyBit, implemented by Pancaj Daga. Real data manipulation has been performed with NiftyReg [MRT+10] and the dataset are part of the ADNI (Alzheimer Disease Neuroimaging Initiative) [JBF+08].

5.1 Log-composition for $\mathfrak{se}(2)$

There are several norms in the space of 3×3 squared matrices that can be inherited by the group $SE(2)$ and the Lie algebra $\mathfrak{se}(2)$ when represented by matrices. For our tests we considered the tangent space $\mathfrak{se}(2)$ with the inherited Frobenius norm:

$$||(\theta, dt^x, dt^y)||_{\text{fro}} = \sqrt{2\theta^2 + (dt^x)^2 + (dt^y)^2} \qquad (\theta, dt^x, dt^y) \in \mathfrak{se}(2)$$

Numerical tests show that for the studied cases, no qualitative differences are detected if choosing instead the L^2 norm.

5.1.1 Methods and Results

To compare the errors the computation of the log-composition for the methods here presented, two sets of 3000 transformations of elements in $\mathfrak{se}(2)$ are randomly sampled with increasing norms in the interval $[0.1, 2.0]$. This interval is divided into 6 segments delimited by $I = \text{linspace}([0.1, 2.0], 7)$ and for each couple of subintervals $[I(n_0), I(n_0 + 1)]$, $[I(n_1), I(n_1 + 1)]$ two sets of 500 transformations $\{dr_0^{(j)}\}_{j=1}^{500}$, $\{dr_1^{(j)}\}_{j=1}^{500}$ having norms belonging to the respective intervals are sampled:

$$\begin{aligned} j = 1, ..., 500 \qquad n_0, n_1 = 0, ..., 5 \\ ||dr_0^{(j)}||_{\text{fro}} \in [I(n_0), I(n_0 + 1)] \\ ||dr_1^{(j)}||_{\text{fro}} \in [I(n_1), I(n_1 + 1)] \end{aligned}$$

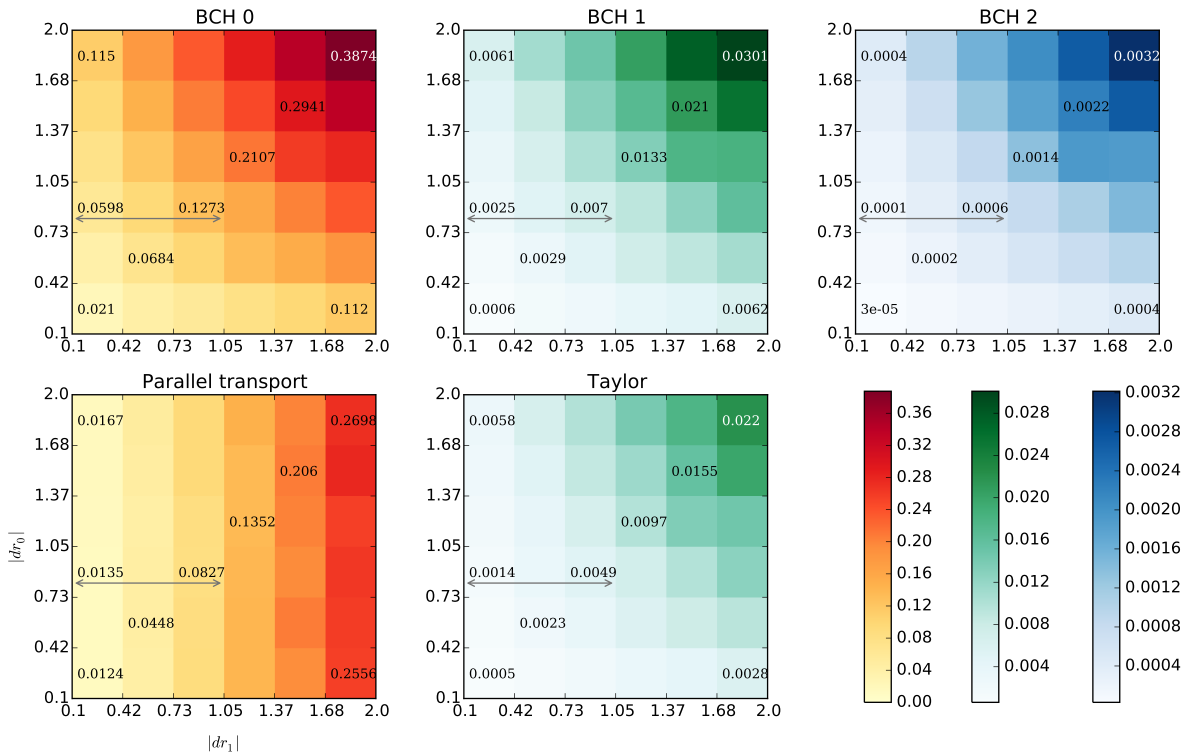

Figure 5.1: Comparison of the errors for each numerical method to compute the Log-composition $dr_0 \oplus dr_1$ in $\mathfrak{se}(2)$. Truncated BCH of degrees 0,1,2, parallel transport method and Taylor method are considered for different values of the norm of dr_1 (x-axes) and norm of dr_0 (y-axes). The value of each square corresponds to the average error of 500 random samples in each of the 6 sub-intervals between 0.1 and 2.0. Errors with BCH 0 and parallel transport method are comparable, but the parallel transport method is not symmetric and has better performance when dr_1 is small. BCH 1 and Taylor are comparable as well and they are both symmetric, but the best performance in terms of approximation is the BCH 2. Values of the sub-square under the *gray arrows* are shown in the boxplot 5.1 where variance, quartiles and outliers are visualized.

If N is one of the numerical methods presented in section 3.1 for the computation of the log-composition - $\mathrm{BCH}^0, \mathrm{BCH}^1, \mathrm{BCH}^2, \mathrm{Tl}, \mathrm{pt}$ - then the error between the ground truth and the approximation provided by one of these numerical methods is given by

$$\mathrm{Error}(dr_0, dr_1, N) := ||dr_0^{(j_0)} \oplus dr_1^{(j_1)} - N(dr_0, dr_1)||_{\mathrm{fro}}$$

In figure 5.1, each of the figure corresponds to a different method and each of the grade scale is the value computed with the function:

$$f(n_0, n_1, N) = \mathbb{E}\Big(\{\mathrm{Error}(dr_0^{(j)}, dr_1^{(j)}, N)\}_{j=1}^{500}\Big)$$

Where the norm of $dr_0^{(j)}$ belongs to the interval $[I(n_0), I(n_0+1)]$ and the norm of $dr_1^{(j)}$ belongs to $[I(n_1), I(n_1+1)]$, and where $\mathbb{E}$ is the mean value.
The data indicated by the gray arrows in each plot corresponds are showed in the box-plot 5.2

From these results in $\mathfrak{se}(2)$ we can see that the second truncation error of the BCH formula provides the best result (the unit of measure is the same as the measure chosen for the translation or the rotation: it can be inches, cm, pixel, ...).

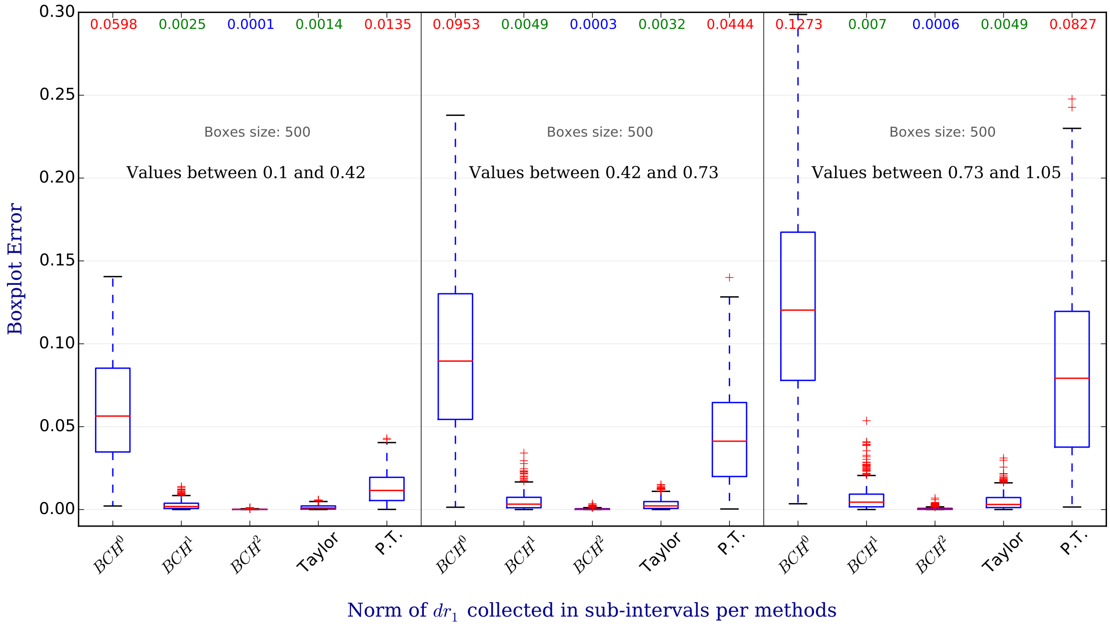

Figure 5.2: Errors of the numerical methods for the computation of the Log-composition of $dr_0 \oplus dr_1$ in $\mathfrak{se}(2)$. Norm of dr_0 is in the interval $[0.37, 1.05]$, norm of dr_1 in the interval $[0.1, 1.05]$ divided in 3 segments. Mean values of each box are shown in the first row in different colors. Shown data corresponds to a section of the image scale 5.1, indicated by a gray arrow. As expected all of the error means increase with the of norm of dr_1, but the rate of the growth is different for each method.

Method based on the BCH^0, that is utilized for example in the additive demons, do not involves any Lie bracket. Its results show that the bigger is the norm of the transformation involved, the bigger is its Lie bracket and its nested Lie bracket as appears in the BCH^1 and BCH^2. Do not take into account Lie brackets means do not take into account the curvature of the space [MTW73], whose significance is given by the experimental results. Parallel transport method tries to compensate the curvature using a geometrical approach considering different tangent spaces to the manifold of the transformation than the one at the origin. As expected from the formula is not symmetric. It provides better results than the BCH^0, and when the norm of dr_1 is small, results are close to the one obtained with BCH^1 when norms of dr_0 and dr_1 are below 1.3.

Log-composition based on Taylor method has slightly better results than the BCH^1, but do not reach BCH^2, which provides the best results. This may be due to the fact that the Taylor belongs to $\mathcal{O}(dr_1^2)$ while the BCH^2 involves the Lie bracket $[dr_0, [dr_0, dr_1]] + [dr_1, [dr_1, dr_0]]$. Even if the truncated BCH does not have a known asymptotic error (or big-O notation), this last observation provides that BCH^2 have a bigger asymptotic order of converges than $\mathcal{O}(dr_1^2)$, in $\mathfrak{se}(2)$.

5.2 Log-composition for SVF

Before getting into the results for the log-composition of SVF it is important to spend some words about how random SVF are created and how to compare the norm of the approximation of $\mathbf{u}_0 \oplus \mathbf{u}_1$ with the ground truth when this is not available.

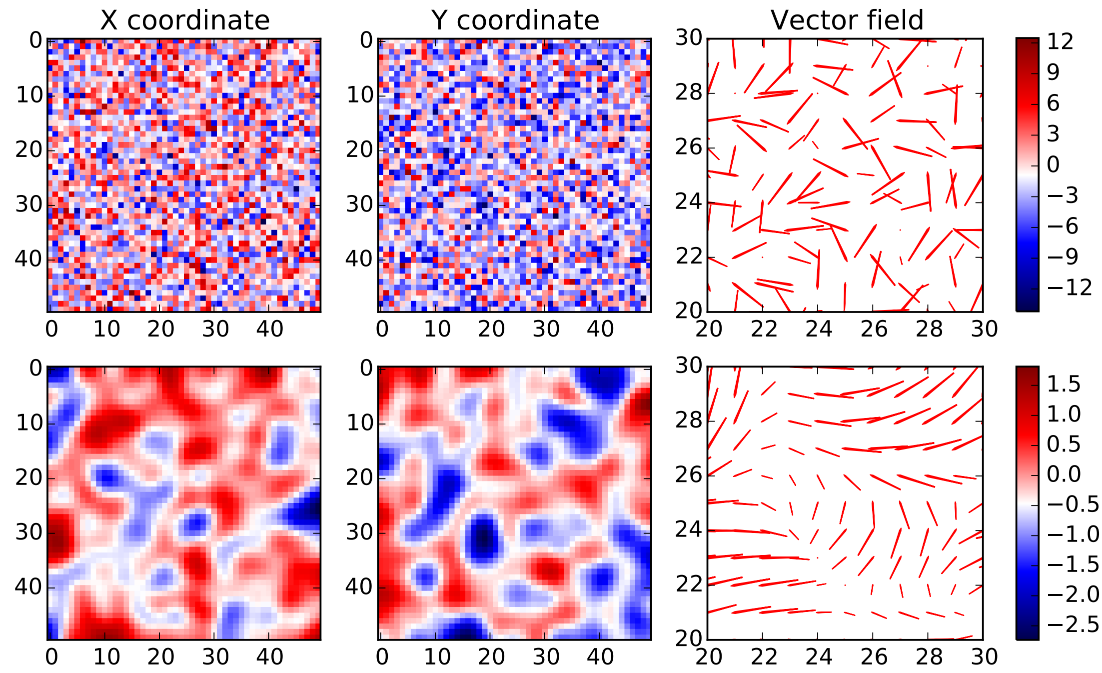

Figure 5.3: Random generated vector field before and after the Gaussian smoother: in the first row a random generated vector field of dimension $50 \times 50 \times 2$ where the value at each pixel are sampled from a random variable with normal distribution of mean 0 and sigma 4. The second row shows the same random vector field after a Gaussian smoothing of sigma 2 (the code is based on the scipy library ndimage.filters.gaussian_filter). In the last column shows the quiver of the vector field in the squared subregion of size 10×10 at the point $(20, 20)$. From the colorscale it is also possible to see that the values distribution of the filtered image is not anymore symmetric.

5.2.1 Methods: random generated SVF.

A random generated SVF is a 5-dimensional matrix with the structure presented in the equation 3.11. The values of the dimensions are fixed while the values of the vectors are built in two phases. First of all, the value for each axis are randomly sampled from a normal distribution of mean zero and standard deviation σ_{init}. To introduce a correlation among each value, a Gaussian filter with standard deviation σ_{gf} is then applied to regularize the values. In figure 5.3 it is possible to see the effects of the two phases on a 50×50 image.

The norm of a discretized SVF $\mathbf{u}$ is the one associated to the metric 3.8 in the discretized case: l^2 instead of L^2 and on $\Delta\Omega$ discretization of the domain:

$$||\mathbf{u}|| = \Big(\sum_{\mathbf{x} \in \Delta\Omega} ||\mathbf{u}(\mathbf{x})||^2_{l^2} \Big)^{1/2}$$

and it coincides with the Frobenius Norm of the 5-dimensional matrices $\mathbf{u}$. When the SVF $\mathbf{u}$ is exponentiated in the Lie algebra $\exp(\mathbf{u}) = \varphi$, we have to rely on the fact that $\varphi - 1$ is a vector field whose norm can be computed with the discretization of 3.9:

$$||\varphi|| = \Big(\sum_{\mathbf{x} \in \Delta\Omega} ||\mathcal{V}(\varphi)(\mathbf{x})||^2_{l^2} \Big)^{1/2} = \Big(\sum_{\mathbf{x} \in \Delta\Omega} ||\varphi(\mathbf{x}) - \mathbf{x}||^2_{l^2} \Big)^{1/2}$$

Where $\Delta\Omega$ is the discretized domain according to a grid.

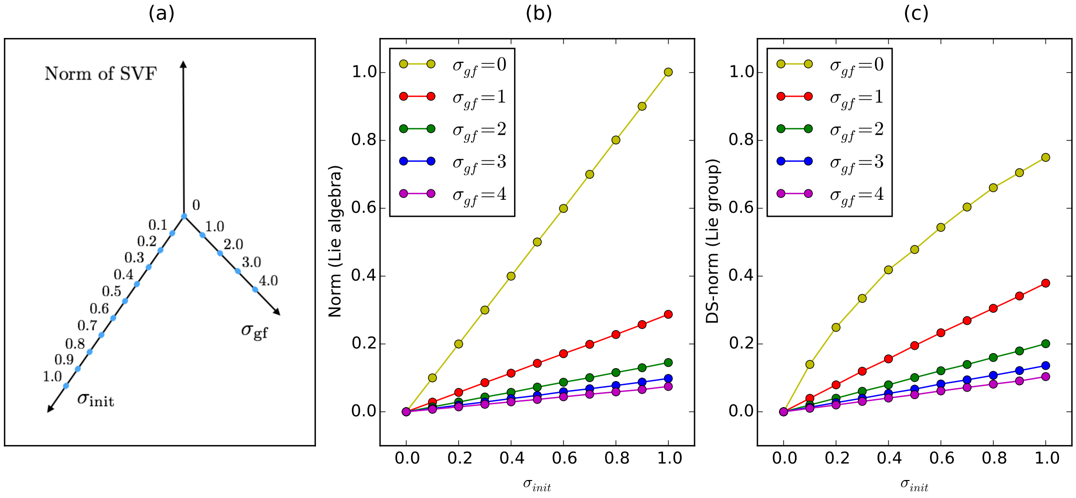

Figure 5.4: Relationship between the initial standard deviation $\sigma_{\rm init}$ that defines the random SVF (stationary velocity field), the standard deviation of the Gaussian filter $\sigma_{\rm gf}$ utilized to regularize the SVF and its norm. Figure (a) represents schematically the two factors that define the norm of an SVF and with the blue dots we emphasized the values that has been chosen for the numerical computations proposed in (b) and (c). Figure (b) shows the mean of the norm of 10 random generated SVF, as element of a Lie group, with initial standard deviation $\sigma_{\rm init}$ (on the x-axis) and Gaussian filter with standard deviation $\sigma_{\rm gf}$ (in different colors). Figure (c) shows the norm of the same element after exponentiating and so after having them in the Lie algebra. It is important to remark that it is not possible in general define a norm on a group. Nevertheless for matrices and for SVF it is possible to extend the norm from the Lie algebra to the Lie group, as proposed in chapter 3 with the definition of displacement field norm (DS-norm).

To distinguish from the previous one, we called it DS-norm (displacement norm); as before, it coincides with the Frobenius norm of the 5-dimensional matrices $\varphi - I$.

Figure 5.4 shows how the initial standard deviations $\sigma_{\rm init}$ that defines the SVF are related with their norm before and after the exponentiation for 5 different choices of the value of the standard deviation of the Gaussian filter $\sigma_{\rm gf}$. Except for the extreme case in which $\sigma_{\rm gf} = 0$, that does not represent any SVF, we can see that an element in the Lie algebra $\mathbf{u}$ has a bigger norm of the correspondent in the Lie group. Moreover in both algebraic structure the norm shows a linear increasing trend of the norm with the increase of $\sigma_{\rm init}$. An increase in $\sigma_{\rm gf}$ implies a decrease in the slope of the trend. Also in this case the slope decreases regularly with an exponential model.

The linear regression of the model for each $\sigma_{\rm gf}$ are given, in Cartesian coordinate by:

$$y = m_{\rm alg}(\sigma_{\rm gf})x \qquad \sigma_{\rm gf} \geq 0 \qquad\qquad y = m_{\rm grp}(\sigma_{\rm alg})x \qquad \sigma_{\rm alg} \geq 0 \tag{5.1}$$

Where we indicated with $m_{\rm alg}$ and $m_{\rm grp}$ angular coefficients for the results obtained in the Lie group and in the Lie algebra respectively. They follow an exponential model, given by

$$m_{\rm alg}(\sigma_{\rm gf}) = \alpha_0 e^{-\beta_0 \sigma_{\rm gf}} + \gamma_0 \qquad m_{\rm grp}(\sigma_{\rm gf}) = \alpha_1 e^{-\beta_1 \sigma_{\rm gf}} + \gamma_1 \tag{5.2}$$

Where the parameters $\alpha_i, \beta_i, \gamma_i$ for $i = 1, 2$ can be computed numerically using an exponential regression algorithm.

α_0	β_0	γ_0	α_1	β_1	γ_1
0.91422836	1.48548466	0.08393943	0.67302265	0.82680977	0.07765811

These values will be useful when we will need to compute σ_{init} for a given σ_{gf}, when expecting a certain value of the norm of the SVF.

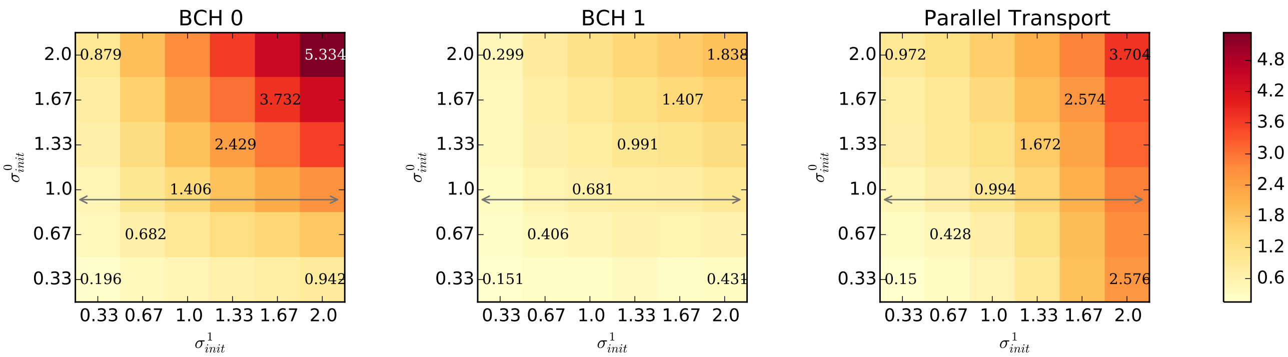

Figure 5.5: Mean errors for the numerical computation of the Log-composition of randomly generated stationary velocity fields (SVF). Initial standard deviations of the SVF σ^0_{init} and σ^1_{init} are given by the values on the axis for each sampling of 15 elements. The error is computed without a ground truth according to the formula 5.4. When the norm of $\mathbf{u}_1$ is small (see figure 5.4 to infer the norm from the standard deviations), parallel transport method and truncated BCH of degree 1 have comparable results, but parallel transport, as expected from the formula, is not symmetric respect to the size of the input vectors. Results of another sampling with the value of σ^0_{init} and σ^1_{init} are shown in figure 5.6.

After showing how a random SVF is generated by the parameters σ_{init} and σ_{gf}, and what is the relationship between the parameters and the resulting norm in both Lie algebra and Lie group, we can move toward the results of the numerical method for the log-composition obtained with these objects.

5.2.2 Log-composition for synthetic SVF

In figure 5.5 we present the results of the numerical computation of the log-composition $\mathbf{u}_0 \oplus \mathbf{u}_1 = \log(\exp(\mathbf{u}_0) \circ \exp(\mathbf{u}_1))$. Despite the lack of a ground truth for the SVF, given $\mathbf{u}_0$ and $\mathbf{u}_1$ in the Lie algebra we can compare the numerical approximation of $\mathbf{u}_0 \oplus \mathbf{u}_1$ after exponentiating, with $\exp(\mathbf{u}_0)\circ\exp(\mathbf{u}_1)$. The norm utilized is the one proposed in the equation 3.10:

$$\text{Error}_{\oplus}(\mathbf{u}_0, \mathbf{u}_1) = \Big(\int_{\Omega} ||\mathcal{V}(\exp(\mathbf{u}_0 \oplus \mathbf{u}_1)) - \mathcal{V}(\exp(\mathbf{u}_0) \circ \exp(\mathbf{u}_1))||^2_{L^2} d\mathbf{x} \Big)^{1/2} \tag{5.3}$$

that, when discretized become

$$\text{Error}_{\oplus}(\mathbf{u}_0, \mathbf{u}_1) = \Big(\sum_{\mathbf{x} \in \Delta\Omega} ||\exp(\mathbf{u}_0 \oplus \mathbf{u}_1)(\mathbf{x}) - \big(\exp(\mathbf{u}_0) \circ \exp(\mathbf{u}_1)\big)(\mathbf{x})||^2_{l^2} \Big)^{1/2} \tag{5.4}$$

For the unknown analytical value of $\mathbf{u}_0 \oplus \mathbf{u}_1$ the error of the above equation is 0. When we use one of the introduced numerical method we obtain the results presented in figure 5.5 and 5.6. The limitation of this strategy to compute the error of the numerical method without a ground truth is that it is based on the numerical algorithm utilized for the computation of the Lie exponential. In this case we utilized the scaling and squaring proposed in [ACPA06]. This limitation may eventually not bias the results, since the scaling and squaring is utilized it the both sides of the difference in the computation of the error, therefore we expect that this numerical approximation do not biasing the results unbalancing one side over the other.

In figure 5.5, we can see the difference between the numerical methods based on BCH^0, BCH^1 and parallel transport. To each square correspond the mean of 15 log-compositions

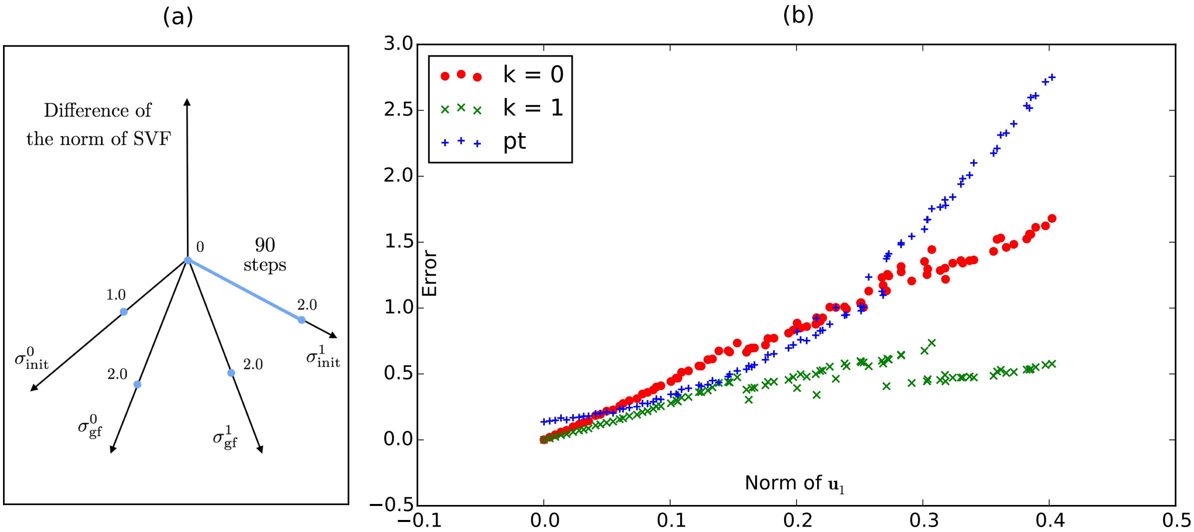

Figure 5.6: Comparisons of the errors of numerical computation of $\mathbf{u}_0 \oplus \mathbf{u}_1$ with the method of truncated BCH of degree 0,1 and parallel transport. Parameters' values of the random generated SVF are schematically represented in figure (a). A first set of 100 SVF, that appears as first element in the log-composition, are generated with fixed parameters $\sigma^0_{\text{gf}} = 2.0$ and $\sigma^0_{\text{init}} = 1.0$; a second set of 100 SVF, that appears as second element in the log-composition, are generated with the parameters $\sigma^1_{\text{gf}} = 2.0$ and σ^1_{init} uniformly scattered in the the interval $(0.0, 2.0)$. On the x-axis fo figure (b) is shown the value of the resulting norm of $\mathbf{u}_1$ for the chosen parameters. On the y-axes the errors for the numerical computation of the Log-composition

between the SVF $\mathbf{u}_0$ and $\mathbf{u}_1$ randomly generated with σ^0_{init} and σ^1_{init} equals to one of the value in the array $(0.33, 0.67, 1, 1.33, 1.37, 2.0)$. The standard deviation of the Gaussian filter σ^0_{gf} and σ^1_{gf} are constant and equal to 2.0. As previously noticed for matrices, the method based on the truncated BCH are symmetric while the same does not happen for the parallel transport.

A more detailed analysis on the subinterval indicated by the gray arrow on the figure 5.5, is shown in 5.6. Here we can see that for SVF the numerical method based on parallel transport is never better than truncated BCH of order 1 (green x) and provides worst results than the simple sum in the truncated BCH of order 0 (red circle) when the norm is greater than 0.35 (with a much higher computational cost). It is also true that for small value the performance of the parallel transport method is below the line draw by the red circles and it is tangential to the green one.

In the results previously showed we can notice some notable absents: non of the method for the numerical computation of the log-composition based on truncated BCH of order greater than 1 has been proposed. The reason of this is explained in the next section.

5.2.3 Truncated BCH formula: The problem of the Jacobian matrix.

The use of the truncated BCH for the numerical approximation fo the log-composition is problematic for SVF. As shown in figure 5.1, i the finite dimensional case the truncated BCH of degree 2 provides the best results over the Taylor method and the parallel transport method. In the infinite dimensional case the truncated BCH involves the Jacobian matrix, that is problematic by a numerical point ov view.

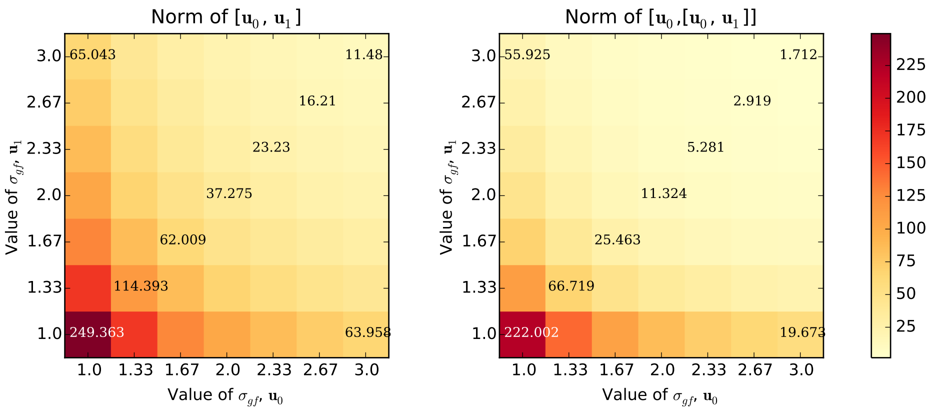

Figure 5.7: Relationship between the standard deviation of the Gaussian smoother that generates the SVF and the norm of the Lie bracket. Each square contains the means of 10 Lie bracket (left) or nested Lie bracket (right) generated with initial standard deviation equals to 2 and standard deviation of the gaussian smoother σ_{gf} indicated on the axes.

On one side every time the differentiation of a vector is required (usually computed with finite difference method -reference-), the results is unstable and sensitive to noise. On the other side, in figure 5.7 we can see that the smoother are the SVFs involved in the log composition, the smaller is the norm of the resulting Lie brackets. Therefore, for a couple of very smooth stationary velocity field, the higher term of the BCH carry little or not information.

For both of these reasons it follows that an increase in the degree of the truncated BCH does not necessarily implies a better approximation or an increase in the robustness of the method. Looking at the boxplot in figure 5.8 we can see why in the previous section we did not compare the truncated BCH of order higher than 1 with the parallel transport method. The truncated BCH of higher degrees does not have better results than the degree 1. It is also interesting to notice that in some cases the truncated BCH formula of degree 2 provides worst results than the truncated BCH formula of degree 1, in particular when the involved SVF have been generated with a small Gaussian smoother.

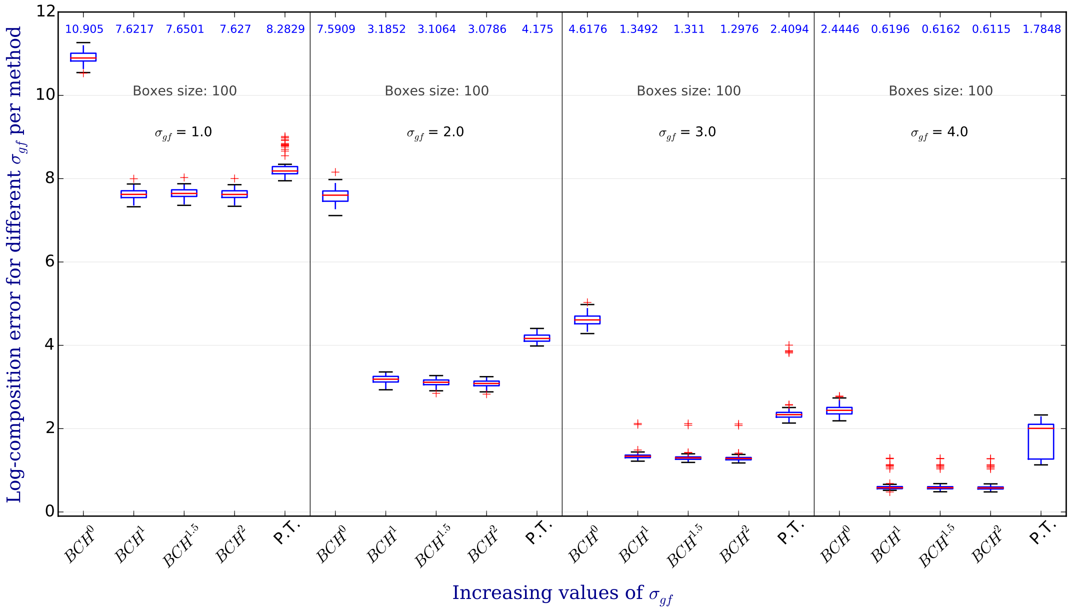

Figure 5.8: Boxplot to compare the error between truncated BCH methods of degree 0, 1, 1.5 and 2. The size of each box is 100 and the approximation of $\mathbf{u}_0 \oplus \mathbf{u}_1$ is performed with $||\mathbf{u}_0|| = 1.0$ and $||\mathbf{u}_0|| = 0.1$. The standard deviation of the Gaussian filter σ_{gf} belongs to the set $(1.0, 2.0, 3.0, 4.0)$ and the initial standard deviation is computed such that $\sigma_{\text{init}} = ||\mathbf{u}||/m_{\text{alg}}(\sigma_{\text{gf}})$ according to the formula 5.1. With this strategy we have been able to compare vector of constant norm generated with increasing values for σ_{gf}. The numbers written in blue above each box represents the mean value of the errors. For small σ_{gf}, an increase in the order of the approximation does not always corresponds to a decrease in the error, and in general there no great improvements can be registered when the degree is greater than 1.

5.3 A Problem for Three Brains

The experiments performed on synthetic data provides important informations to validate and compare the methods, but gives little or not information at all about what may happen in the real cases.

To obtain a validation for the real cases, one of the possibility would be to embed the method in a diffeomorphic demon registration algorithm and compare its results with the different versions of the log-composition implemented for the computation of the update.

But, since the log-composition is only a small component of the registration algorithm, it may be difficult to understand to what extent the change of the numerical method for the log-composition would impact the results and what is due to other components of the software that play a more important role (as the optimization strategy) in the registration algorithm.

5.3.1 Design of Experiment

For our purposes we design a simple experiment that involves three T1 MRI longitudinal scans of the same patient in three points in time A, B and C. A is the baseline, B is the first follow-up after 7 months and C is the third follow-up, after 12 months. Dataset has been collected from ADNI (Alzheimer Disease Neuroimaging Initiative) [JBF$^+$08].

Using NiftyReg is it possible to obtain the SVF of the transformations obtained by the

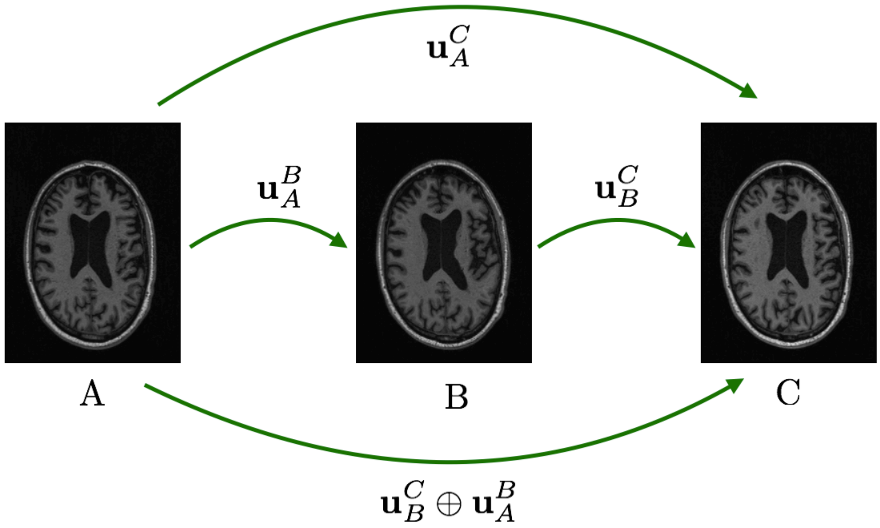

Figure 5.9: design of the experiment with SVF obtained with real data. The...

non-rigid registration algorithm based on cubic B-spline. If φ_A^B is the transformation from A to B that aligns the image B with A, and $\mathbf{u}_A^B$ the corresponding SVF, then for an ideal registration algorithm (that fully satisfies the transitivity property) we have

$$\varphi_B^C \circ \varphi_A^B = \varphi_A^C \tag{5.5}$$

From this assumption, it follows:

$$\begin{aligned} \log(\varphi_B^C \circ \varphi_A^B) &= \log(\varphi_A^C) \\ \log(\exp(\mathbf{u}_B^C) \circ \exp(\mathbf{u}_A^B)) &= \log(\exp(\mathbf{u}_A^C)) \\ \log(\exp(\mathbf{u}_B^C) \circ \exp(\mathbf{u}_A^B)) &= \mathbf{u}_A^C \end{aligned}$$

The exact solution of the log-composition would provide then:

$$\int_\Omega ||\mathbf{u}_B^C) \oplus \exp(\mathbf{u}_A^B - \mathbf{u}_A^C||_{L^2}^2 d\mathbf{x} = 0$$

While, when $\mathbf{u}_B^C) \oplus \exp(\mathbf{u}_A^B$ is approximated with some numerical method M, the previous integral provides its error.

$$\int_\Omega ||M(\mathbf{u}_B^C) \oplus \exp(\mathbf{u}_A^B) - \mathbf{u}_A^C||_{L^2}^2 d\mathbf{x} = \text{Error}_\oplus \tag{5.6}$$

Certainly, the equation 5.5 is correct only for the ideal registration algorithms, and the distance between the two members, computed using for example the metric d^1 defined in equation 3.9 can be utilized to validate the registration algorithm:

$$d^1(\varphi_B^C \circ \varphi_A^B, \varphi_A^B)) = \text{Error}_{\text{reg}}$$

The error $\text{Error}_{\text{reg}}$ is embedded in the the computations of the equation 5.6, but it has the same effect for all of the numerical approximations of the log-composition. Therefore, even if the equation 5.6 does not measures exclusively the numerical error of the log-composition, they are consistent.

5.3.2 Results

In figure ... mean error boxplot and divided by patient graph!

5.4 Log-Algorithm for SVF

To compare the different numerical methods for the the logarithm computation algorithm we remain where a ground truth is available, i.e. in the finite dimensional case. We consider a data set of 200 random matrices in the Lie group $SE(2)$ with their respective logarithm in $\mathfrak{se}(2)$ computed using the closed form presented in chapter 3. For each of the method considered and for each of the random matrix we have convergence up to a precision of 10^{-12} before the 50^{th} iteration.

In figure 5.10 we can see the average number of iterations required to reach the solution with a precision of 10^{-4} for each of the numerical method considered. Each box represents 200 random matrices in $SE(2)$ with Frobenius norm uniformly selected between 1 and 3.

As obtained in the log-composition the results obtained with parallel transport is comparable with the one obtained with the truncated BCH of degree 1. Remarkably, in this case, an increase in the degree of the truncated BCH does not ensure a faster convergence and the Taylor method, provides the slower algorithm. The reason for these facts are not clear are the moment and are worthed to be investigated in a future work.

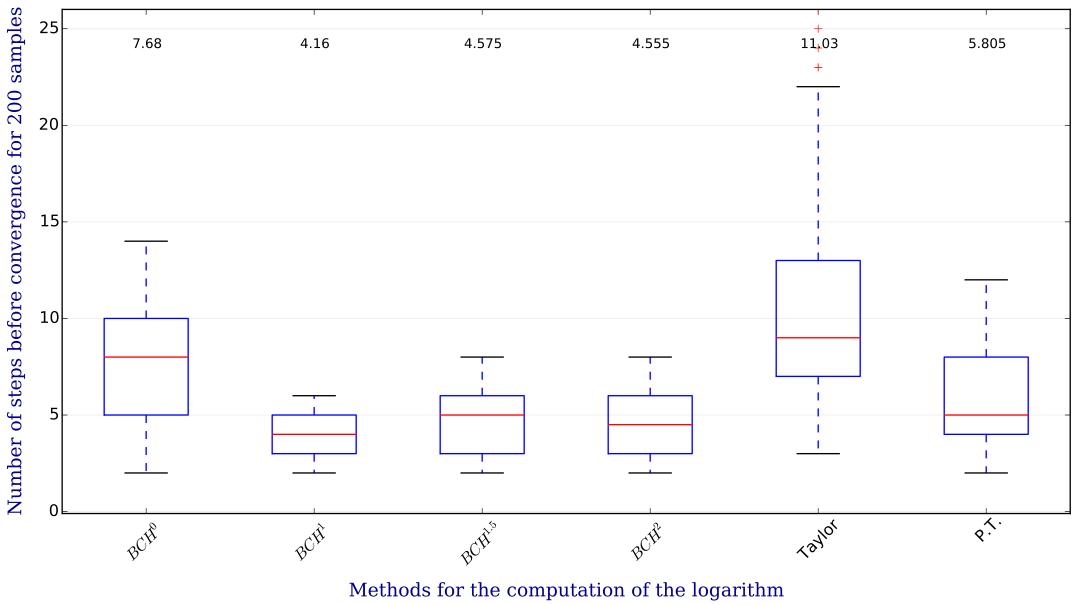

Figure 5.10: Number of steps required to obtain convergence in for different methods utilized in the logarithm computation algorithm. The data set contains 200 random matrices in the Lie group $SE(2)$, with Frobenius norm between 1 and 3. On the top it is possible to visualize the mean number of step to reach the convergence for each method.

5.5 Empirical Evaluations of the Computational Time

For the finite dimensional case, a dataset of 36000 random generated matrix with norm between 0.0 and 2.0 has been utilized to measure the computational time of each of the numerical method for the computation of the log-composition here presented. Sum of the computational time for the whole data set is give in seconds, in the following table:

Ground	BCH^0	BCH^1	$BCH^{1.5}$	BCH^2	Taylor	p.t.
1.07402015	0.18845153	0.57751322	1.24413943	1.78752184	0.77354765	2.26586294

The first column provides the computation of the ground truth, i.e. the closed for the computation of $dr_0 \oplus dr_1$. We can see that the computation of the BCH^0 is the fastest, since it consists only in a sum, while the computation of the parallel transport is four time slower than the BCH^1.

When dealing with SVF, the computational time strongly depends on the size of the vector field involved. In figure 5.11 we can see the relation between the size of the vector field and the increase in the computational time for a data set of 20 random generated SVF.

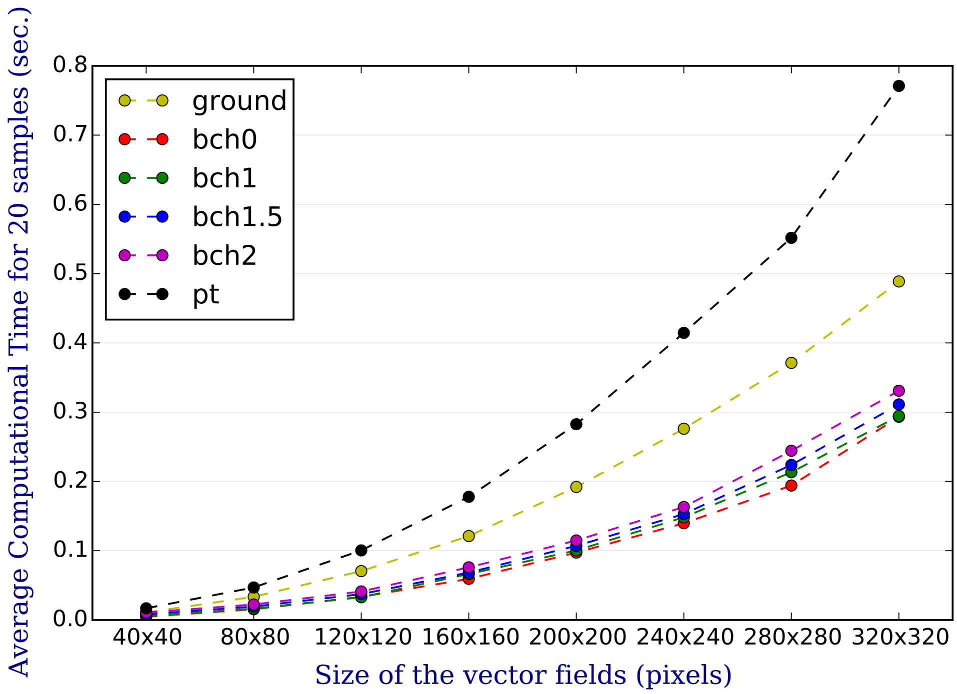

Figure 5.11: Relationship between the size of the figure (x-axes) and the computational time for a data set of 20 random generated SVF. The yellow line labeled with ground represents the time of the computation of $\exp \mathbf{u}_0 \circ \exp \mathbf{u}_1$, while the other represents the exponentiation of the numerical method for the computation of the log-composition.

5.6 Conclusions

A seen from results not

Considering only the results, this one-year research can be considered much ado about nothing, but...
Computational time...!

5.7 Further Researches

The BCH is proved only when the exp and log can be expressed in power series, so when the Lie group and the Lie algebra involved belongs to the same bigger group. This is not the case of the infinite dimensional Lie group of diffeomorphisms,

Starting from the definition of Lie log-group of diffeomorpshisms $(\mathfrak{g}, \oplus)$, to have an algebraic definition of this approximation, we can consider its quotient over the ideal generated by $(\mathrm{ad}^m_{\mathbf{u}}, \mathrm{ad}^n_{\mathbf{u}})$, which provides the group $({}^{\mathfrak{g}}\big/_{(\mathrm{ad}^m_{\mathbf{u}}, \mathrm{ad}^n_{\mathbf{u}})}, \oplus)$. Further investigations in this direction is not prosecuted.

Bibliography

[ACPA06] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A log-euclidean framework for statistics on diffeomorphisms. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 924–931. Springer, 2006.

[AFPA06] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, August 2006.

[APA06] Vincent Arsigny, Xavier Pennec, and Nicholas Ayache. Bi-invariant means in lie groups. application to left-invariant polyaffine transformations. 2006.

[Arn66] Vladimir Arnold. Sur la géométrie différentielle des groupes de lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. In *Annales de l'institut Fourier*, volume 16, pages 319–361. Institut Fourier, 1966.

[Arn98] Vladimir Arnold. *Topological methods in hydrodynamics*, volume 125. Springer Science & Business Media, 1998.

[Arn06] Vladimir Arnold. Ordinary differential equations. translated from the russian by roger cooke. second printing of the 1992 edition. universitext, 2006.

[Art11] Michael Artin. Algebra. 2nd, 2011.

[BBHM11] MARTIN BAUER, MARTINS BRUVERIS, PHILIPP HARMS, and PETER W MICHOR. Geodesic distance for right invariant sobolev metrics of fractional order on the diffeomorphism group. 2011.

[BHM10] Martin Bauer, Philipp Harms, and Peter W Michor. Sobolev metrics on shape space of surfaces in n-space. *Arxiv preprint arXiv:1009.3616*, 2010.

[BO08] Matias Bossa and Salvador Olmos. A New Algorithm for the Computation of the Group Logarithm of Diffeomorphisms. In Xavier Pennec and Sarang Joshi, editors, *Second International Workshop on Mathematical Foundations of Computational Anatomy - Geometrical and Statistical Methods for Modelling Biological Shape Variability*, New York, USA, 2008.

[BZO08] Matias Bossa, Ernesto Zacur, and Salvador Olmos. Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008.

[DCDC76] Manfredo Perdigao Do Carmo and Manfredo Perdigao Do Carmo. *Differential geometry of curves and surfaces*, volume 2. Prentice-hall Englewood Cliffs, 1976.

[dCV92] Manfredo Perdigao do Carmo Valero. *Riemannian geometry*. 1992.

[Dyn00] EB Dynkin. Calculation of the coefficients in the campbell–hausdorff formula. *Selected Papers of EB Dynkin with Commentary. Originally in Institute of Mathematics Moscow State University, Acad a.n. Kormologorow, submitted 1947*, 14:31, 2000.

[EM70] David G Ebin and Jerrold Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Mathematics*, pages 102–163, 1970.

[FF97] Nick C Fox and Peter A Freeborough. Brain atrophy progression measured from registered serial mri: validation and application to alzheimer's disease. *Journal of Magnetic Resonance Imaging*, 7(6):1069–1075, 1997.

[Gal11] Jean Gallier. *Geometric methods and applications: for computer science and engineering*, volume 38. Springer Science & Business Media, 2011.

[Gra88] Janusz Grabowski. Free subgroups of diffeomorphism groups. *Fundamenta Mathematicae*, 2(131):103–121, 1988.

[GWRNJ12] Serge Gauthier, Liyong Wu, Pedro Rosa-Neto, and Jianping Jia. Prevention strategies for alzheimer's disease. *Translational neurodegeneration*, 1(1):1–4, 2012.

[Hal15] Brian Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2015.

[Hun07] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95, 2007.

[HWS$^+$] Zhiwu Huang, Ruiping Wang, Shiguang Shan, Xianqiu Li, and Xilin Chen. Log-euclidean metric learning on symmetric positive definite manifold with application to image set classification.

[ISNC03] Luis Ibanez, William Schroeder, Lydia Ng, and Josh Cates. The itk software guide. 2003.

[JBF$^+$08] Clifford R Jack, Matt A Bernstein, Nick C Fox, Paul Thompson, Gene Alexander, Danielle Harvey, Bret Borowski, Paula J Britson, Jennifer L Whitwell, Chadwick Ward, et al. The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging*, 27(4):685–691, 2008.

[JOP$^+$] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed 2015-08-03].

[Kir08] Alexander Kirillov. *An introduction to Lie groups and Lie algebras*, volume 113. Cambridge University Press Cambridge, 2008.

[KMN00] Arkady Kheyfets, Warner A Miller, and Gregory A Newton. Schild's ladder parallel transport procedure for an arbitrary connection. *International Journal of Theoretical Physics*, 39(12):2891–2898, 2000.

[Kne51] MS Knebelman. Spaces of relative parallelism. *Annals of Mathematics*, pages 387–399, 1951.

[KO89] S Klarsfeld and JA Oteo. The baker-campbell-hausdorff formula and the convergence of the magnus expansion. *Journal of physics A: mathematical and general*, 22(21):4565, 1989.

[KW08] Boris Khesin and Robert Wendt. *The geometry of infinite-dimensional groups*, volume 51. Springer Science & Business Media, 2008.

[LAP11] Marco Lorenzi, Nicholas Ayache, and Xavier Pennec. Schild's ladder for the parallel transport of deformations in time series of images. In *Information Processing in Medical Imaging*, pages 463–474. Springer, 2011.

[Lee12] John Lee. *Introduction to smooth manifolds*, volume 218. Springer Science & Business Media, 2012.

[Les83] JA Leslie. A lie group structure for the group of analytic diffeomorphisms. *Boll. Un. Mat. Ital. A (6)*, 2:29–37, 1983.

[LP13] Marco Lorenzi and Xavier Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International journal of computer vision*, 105(2):111–127, 2013.

[LP14a] Marco Lorenzi and Xavier Pennec. Discrete Ladders for Parallel Transport in Transformation Groups with an Affine Connection Structure. In Frank Nielsen, editor, *Geometric Theory of Information*, Signals and Communication Technology, pages 243–271. Springer, 2014.

[LP14b] Marco Lorenzi and Xavier Pennec. Efficient parallel transport of deformations in time series of images: from schild's to pole ladder. *Journal of Mathematical Imaging and Vision*, 50(1-2):5–17, 2014.

[MA70] J Marsden and R Abraham. Hamiltonian mechanics on lie groups and hydrodynamics. *Global Analysis,(eds. SS Chern and S. Smale), Proc. Sympos. Pure Math*, 16:237–244, 1970.

[MHM+11] JR McClelland, S Hughes, M Modat, A Qureshi, S Ahmad, DB Landau, S Ourselin, and DJ Hawkes. Inter-fraction variations in respiratory motion models. *Physics in medicine and biology*, 56(1):251, 2011.

[MHSK] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King. Respiratory motion models: A review. *Medical Image Analysis*, 17(1):19–42, 2015/04/01.

[Mic80] P Michor. Manifolds of smooth maps, ii: the lie group of diffeomorphisms of a non-compact smooth manifold. *Cahiers de Topologie et Géométrie Différentielle Catégoriques*, 21(1):63–86, 1980.

[Mil82] John Milnor. On infinite dimensional lie groups. *Preprint, Institute for Advanced Study, Princeton*, 1982.

[Mil84a] J Milnor. Remarks on infinite-dimensional lie groups, in 'relativity, groups and topology, ii'(les houches, 1983), 1007–1057, 1984.

[Mil84b] John Milnor. Remarks on infinite-dimensional lie groups. In *Relativity, groups and topology. 2*. 1984.

[MRT+10] Marc Modat, Gerard R Ridgway, Zeike A Taylor, Manja Lehmann, Josephine Barnes, David J Hawkes, Nick C Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010.

[MT13] Carlo Mariconda and Alberto Tonolo. *Calcolo discreto: Metodi per contare*. Apogeo Editore, 2013.

[MTW73] Charles W Misner, Kip S Thorne, and John Archibald Wheeler. *Gravitation*. Macmillan, 1973.

[Nee06] KH Neeb. Infinite dimensional lie groups, 2005 monastir summer school lectures. *Lecture Notes January*, 2006.

[OKC92] V Yu Ovsienko, BA Khesin, and Yu V Chekanov. Integrals of the euler equations of multidimensional hydrodynamics and superconductivity. *Journal of Soviet Mathematics*, 59(5):1096–1101, 1992.

[Omo70] Hideki Omori. On the group of diffeomorphisms on a compact manifold. In *Proc. Symp. Pure Appl. Math., XV, Amer. Math. Soc*, pages 167–183, 1970.

[PCL+15] Ferran Prados, Manuel Jorge Cardoso, Kelvin K Leung, David M Cash, Marc Modat, Nick C Fox, Claudia AM Wheeler-Kingshott, Sebastien Ourselin, Alzheimer's Disease Neuroimaging Initiative, et al. Measuring brain atrophy with a generalized formulation of the boundary shift integral. *Neurobiology of aging*, 36:S81–S90, 2015.

[PL+11] Xavier Pennec, Marco Lorenzi, et al. Which parallel transport for the statistical analysis of longitudinal deformations. In *Proc. Colloque GRETSI*. Citeseer, 2011.

[Sch10] Rudolf Schmid. Infinite-dimensional lie groups and algebras in mathematical physics. *Advances in Mathematical Physics*, 2010, 2010.

[SDP13] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, July 2013.

[Ser09] Jean-Pierre Serre. *Lie algebras and Lie groups: 1964 lectures given at Harvard University*. Springer, 2009.

[Sze94] Richard Szeliski. Image mosaicing for tele-reality applications. In *Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on*, pages 44–53. IEEE, 1994.

[Ver14] Tom Vercauteren. Technical memo. Pre-print, 2014.

[VPM+06] Tom Vercauteren, Aymeric Perchant, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. Robust mosaicing with correction of motion distortions and tissue deformations for in vivo fibered microscopy. *Medical image analysis*, 10(5):673–692, 2006.

[VPPA07] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Non-parametric diffeomorphic image registration with the demons algorithm. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 319–326. Springer, 2007.

[VPPA08] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Symmetric log-domain diffeomorphic registration: A demons-based approach. In Dimitris N. Metaxas, Leon Axel, Gabor Fichtinger, and Gábor Székely, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008, 11th International Conference, New York, NY, USA, September 6-10, 2008, Proceedings, Part I*, volume 5241 of *Lecture Notes in Computer Science*, pages 754–761. Springer, 2008.

[War13] Frank W Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94. Springer Science & Business Media, 2013.

[Wig60] Eugene P Wigner. The unreasonable effectiveness of mathematics in the natural sciences. richard courant lecture in mathematical sciences delivered at new york university, may 11, 1959. *Communications on pure and applied mathematics*, 13(1):1–14, 1960.

[Woj94] Wojciech Wojtyński. One-parameter subgroups and the bch formula. *Studia Mathematica*, 111(2):163–185, 1994.

[YC06] Lexing Ying and Emmanuel J Candès. The phase flow method. *Journal of Computational Physics*, 220(1):184–215, 2006.

[You10] Laurent Younes. Shapes and diffeomorphisms, volume 171 of applied mathematical sciences, 2010.