Sebastiano Ferraris

The Log-composition of Stationary Velocity Fields in Diffeomorphic Image Registration

University College London
Medical Physiscs and Biomedical Engineering

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

July 27, 2015

Supervisor
Tom Vercauteren

Co-Supervisor
Marc Modat

Clinical Supervisor
Ian Deprest

Acknowledgments

It was not and it is still not easy for me the trail between the study of pure mathematics to 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, and in solving various problems and doubts, but especially in having - laboriously - introduced me to the craft of biomedical research.

I am also grateful to the rowing fellow with whom I shared commitment, challenges and happiness of 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 the 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 important opportunity I've ever had.

In classic 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 patient, Carole Sudre.

Abstract

Image registration is one of the critical tool in medical imaging. It consists in the process of aligning of two or more patient images with the aim of determining and quantifying the occurring anatomical correspondences and differences. It is widely used in both academic studies and applications, and it continuously challenges researchers to enhance accuracy, improve reliability and reduce computational time.

Among many approaches to the problem, the introduction of the *Lie group of diffeomorphisms* from differential geometry provides an interesting set of deformation to model the organs' deformations. Comparing two diffeomorphisms, as well as obtaining any meaningful statistics for these elements is not a straightforward task, due to the lack of a norm or a metric. The *log-Euclidean framework* introduced in 2006, proposes to map the diffeomorphisms in the tangent vector space at the identity, so to have a local representation of a diffeomorphisms in a vector space where norms can be defined. The maps from the Lie group of diffeomorphisms to its tangent space is known in Lie group theory as *Lie exponential* and its inverse is the *Lie logarithm.*

These two transformations allow in particular the computation of the composition of diffeomorphisms in the tangent space, operation called in this thesis *log-composition.* The necessity of finding fast numerical methods for its computation arises for example in the *log-demons* and in the *symmetric log-demon* registration algorithm.

In this research we analyze existing numerical methods for the computation of the log-composition, based on the BCH formula and we compare them with two new methods developed in this research, one based on the Taylor expansion and the other on the geometrical concept of parallel transport.

This document contains 15xxx words.

Contents

Chapter 1

Introduction: Diffeomorphisms in Medical Imaging Registration

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

Aim of this chapter is to introduce the log-composition. It arises in medical image registration, when diffeomorphisms are utilized to model the transformation of anatomies between images. Before getting into it formalization, it is important to spend some few words about the context and the reasons that led to this definition.

1.1 Toward an ill-posed Problem

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.

Involved difficulties and opportunities are a consequence of the fact that dealing with image registration problem means dealing with an ill-posed problem. Transformations between anatomies are not unique, and the impossibility to recover spatial or temporal evolution of an anatomical transformation from temporally isolated images, makes any validation a difficult, if not an impossible task. In addition each situation inevitably brings some prior knowledge within the initial data, that may imply some modifications in the problems' setting and may imply some additional constraints. This, of course, impacts dramatically the range of possible choices in searching for a solution and in the consequent results.

Certainly it is the practical situation that provides the hint in choosing the optimal constraints, but it almost never provides enough information to reduce the large amount of options involved. A wide range of variants in methodologies and approaches to solve the registration problem has been thus proposed in the last decades: a quick glance to Google scholar reveals about 1200000 papers in *medical image registration* (55% of the whole *image registration* resources).

1.1.1 Some Examples of Medical Image Registration

One of the most studied application of image registration is in the domain of brain imaging: there this tool can be used to examine differences between subjects and distinguish their biological features - cross-sectional studies - or to compare different acquisition of the same subject after a fixed period of time or before and after a surgical operation - longitudinal studies.

In both cross-sectional and longitudinal studies an accurate comparison between images and the parameters of the transformation involved may result in a quantification of anatomical variability and in a better understanding of the patients' features. For example, brain atrophy is considered a biomarker to diagnose Alzheimer disease and to analyze its evolution; most of the algorithms and techniques involved in the atrophy measurement require longitudinal or cross-sectional scans to be aligned, and so are directly affected by the solution of the registration algorithm [PCL+15], [FF97], [GWRNJ12].

Also when dealing with motion correction, if a sequence of images is affected by the motion of cardiac pulses or respiratory cycles, registration algorithms are often used for the realignment. For example, in lungs radiotherapy, the correspondence between the lungs' deformation and the respiratory signal defines a model to direct the X-ray or electrons beam on the cancer, avoiding as much as possible the sane tissue. Lungs deformation is obtained using a registration algorithms that provides the direction of the motion of each voxel in each phase of breathing [MHSK], [MHM+11].

Another application of image registration is the operation of gluing together several pictures with partially coincident regions, with the aim of obtaining a bigger image of the whole scene. This procedure, called *mosaicing*, exploit registration algorithms to aligns images using information obtained form the overlapping regions [VPM+06], [Sze94].

The next section moves toward some details of one iterative framework most commonly utilized by image registration algorithms.

1.1.2 Image Registration Problem

A *d-dimensional image* is a continuous function from a subset Ω of the coordinate space $\mathbb{R}^d$ (having in mind particular cases $d = 2, 3$) to the set of real numbers $\mathbb{R}$. Given two of them, $F : \Omega_F \to \mathbb{R}$ and $M : \Omega_M \to \mathbb{R}$, called respectively *fixed image* and *moving image*, the *image registration problem* consists in finding the transformation function

$$\begin{aligned} \varphi : \mathbb{R}^d \supseteq \Omega_F &\longrightarrow \Omega_M \subseteq \mathbb{R}^d \\ \mathbf{x} &\longmapsto \varphi(\mathbf{x}) \end{aligned}$$

such that for each point $\mathbf{x} \in \Omega_F$ the element $M(\varphi(\mathbf{x}))$ and $F(\mathbf{x})$ are as closed as possible according to a chosen measure of similarity. Other than obtaining φ, also the investigation of its features and parameters are a part of the problem.

The definition of image registration problem proposed here can be represented by the following diagram, where φ is the solution that, in the ideal case, makes f the function that match the correspondences in the intensities when images are aligned:

$$\begin{tikzcd} \Omega_F \arrow[r, "\varphi"] \arrow[d, "F"] & \Omega_M \arrow[d, "M"] \\ \mathbb{R} \arrow[r, dashed, "\mathrm{f}"] & \mathbb{R} \end{tikzcd}$$

The composition of the moving image after the transformation, $M \circ \varphi$, is called *warped image*, and if $\Omega_F \neq \Omega_M$, it is always possible to apply an homeomorphism to transform them into a common domain Ω, called *background space*, on which both of the images are defined.

Initially, this setting leaves two main degrees of freedom in searching for a solution: the transformation's domain to which φ belongs (also called *deformation model*), and the metric to measure the similarity between images. Once these are chosen, they are the main constituent of the *image registration framework*: an iterative process that provides at each step a new function φ that approach one of the possible solution to the registration problem.

1.1.3 Iterative Registration Framework

The definition of registration problem and the iterative framework described above raise several issues. For example there are no reasons to believe that the correspondence that models the deformation between images is unique. In addition the condition $M(\varphi(\mathbf{x})) = F(\mathbf{x})$ for each point $\mathbf{x} \in \Omega_F$ can be satisfied by functions that do not represents any reasonable biological transformation between anatomies.

One way to deal with these issue is to add some constraints on the transformation φ, such that it is bound to model realistic changes that can occur in biological tissues. The kind and quality of the constraints are one of the features that distinguish one registration algorithm to the other, and they can be mathematically expressed by the definition of a deformation model and an *energy function* (or objective function). This last measures the similarity between the fixed image and the warped image, and it is indicated here with Sim. Moreover a regularization term, here indicated with Reg, is added to the similarity measure, to add further constraints on the transformation, penalizing its measured irregularities. The objective function, therefore has in general the form:

$$\mathcal{E}(F, M, \varphi) = \mathrm{Sim}(F, M, \varphi) + \mathrm{Reg}(\varphi) \tag{1.1}$$

In the registration framework, an optimization algorithm is utilized to minimize the equation 1.1 and to provide the sought transformation bonded to a chosen domain.

Finally, since images are modeled by continuous functions but are represented as discrete structure, a resampling technique has to be chosen among several options (see for example [Gon]). Its choice has a relatively small impact on the image registration algorithm, nevertheless it implies another range of possibility in the definition of the registration framework.

According to the registration framework here presented, there are 5 parameters that determine the consequent image registration algorithm, each with its range:

$$\begin{aligned}
\{\varphi\} &\in \{\text{Transformations}\} \\
\mathrm{Sim} &\in \{\text{Similarity measures}\} \\
\mathrm{Reg} &\in \{\text{Regularization Terms}\} \\
\mathrm{Opt} &\in \{\text{Optimization techniques}\} \\
\mathrm{Res} &\in \{\text{Resampling techniques}\}
\end{aligned}$$

They provide the constraints imposed to the image registration algorithm to solve the image registration problem, and their choice is made in consequence of the specific situation. The flowchart of the framework is shown in figure 1.1: given a fixed image F, a moving image M and an initial transformation φ_0, the warped image $M \circ \varphi_0$ is computed, and the energy function 1.1 is optimized at each step by the algorithm. The resulting new transformation φ takes the place of φ_0 for the subsequent steps.

Many of the available registration algorithms roughly follow this scheme, and the specific choice of the 5 parameters involved provides a preliminary classification of the algorithm. For further details see for example the recent surveys [SDP13] and the less recent [ZF03].

1.2 Diffeomorphisms in Medical Imaging Registration

In any image registration algorithm, one of the most relevant feature is the choice of the family to whom the transformation belongs. This is as an important constraint that change

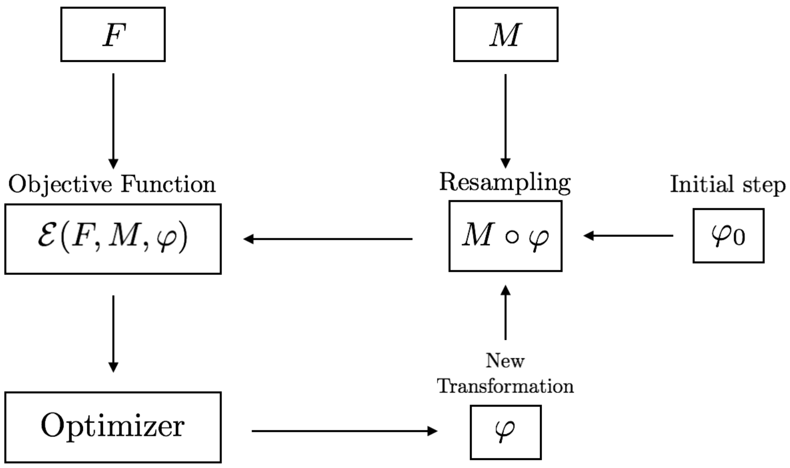

Figure 1.1: Flowchart of the image registration framework.

according to the aim of the registration and to the nature of the objects represented by the images.

If the algorithm is meant to model physical transformations that preserve distances, orientations and angles, then the set of transformations can be bonded to the group of rigid rotations and translations in three dimensions, $SE(3)$. 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 scans.

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 the mathematical objects of *diffeomorphisms* are taken into account. These are defined as bijective differentiable maps with differentiable inverse, and are particularly well suited to model non-rigid deformations between images (for a general introduction on diffeomorphisms see for example [Lee12], [Arn06]). Consequent algorithms are called *diffeomorphic registration algorithms*.

These algorithms, thanks to the property of invertibility and topology-preserving of the transformations involved, appears to be a natural choice to model organs' deformations and in many cases are the ideal candidate for the set of transformation in the registration framework. This is due to the fact that in longitudinal studies, anatomies are involved in a smooth process of modification over time that do not presume any breaking of topology. Also in most of the cross-sectional studies variations in the topology of the same organ in different patients are not expected.

It is importance to notice that, when implemented in image registration algorithms, diffeomorphisms do not have only radically different results than the rigid body transformations: they also possess a different mathematical structure.

1.2.1 Utility and Liability of Diffeomorphisms

On the algebraic side, the set of diffeomorphisms appears particularly interesting for their group structure and within their differentiable nature (see for example [Mic80] and [Lem97]). They have also an infinite-dimensional structure of vector space, and their mathematical formalization as Lie group (so as differentiable manifold within a group structure, see [War13], [Lee12]) is an open field of research whose development has not yet reach a definitive formalization.

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 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 element in local correspondence with some generalized "infinite-dimensional Euclidean" space. Attempt 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 holds 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 group [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 is only toward the diffeomorphisms defined on a compact subset 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.

1.2.2 State of the Art

In the development of diffeomorphic image registration, we can broadly identify some steps that led to the diffeomorphic demons and to the consequent concept of log-composition presented in this research:

1981-1996 ▷ The use of diffeomorphisms in medical image registration began from the research of a solution to a class partial differential equations: deformations are modeled as the consequent effect of two balancing forces applied to an elastic body [Bro81] or to conserve the energy momentum [CRM96]. In this early stage, diffeomorphisms are the domain of the solution of a set of differential equation, and are not considered with their Lie group structure.

1998-2004 ▷ Based on the concept of attraction, the demons algorithm [Thi98], [PCA99] proposes the computation of the transformation between images in an iterative framework, where

the update of the transformation at each step is parametrized with a discrete vector field of independent vectors (or demons) that is optimized at each step. Each voxel of the moving image is considered within a vector that transforms it into a new position, according to the positions of the voxel of the same intensity in the fixed image.
Here diffeomorphisms are not directly involved and the vectors at each voxel are independent elements. In the same year of [Thi98], the utilization of diffeomorphism was taken into account in image matching and computational anatomy, not only as the set of solutions of some family of differential equations, but with its tangent space [DGM98, Tro98, GM98].

2005-2006 ▷ In this period it has been proposed the Beg's version of Large Deformation Diffeomorphic Metric Mapping (called in this research Beg-LDDMM, to distinguish from others LDDMM versions) [BMTY05] for diffeomorphic image registration and the log-Euclidean framework [ACPA06b, AFPA06] as an investigation of the tangent space to the Lie group of diffeomorphisms as a space where to perform statistics. The Beg-LDDMM utilizes in practice all of the opportunities provided by differential geometry in considering tangent vectors to the space of transformation in a framework for the computation of image registration. In this setting, the tangent vector field comes from the solution of the ODE that models the transformations and it consists of the set of the non-stationary vector field (also time varying vector field or TVVF). After the log-Euclidean framework aimed at the computation of statistics of diffeomorphisms, only the subset of the group of diffeomorphisms that are parametrized by stationary vector fields (also stationary velocity field or SFV) is taken into account for practical computations.

2007-2013 ▷ The restriction to SVF was subsequently considered in some further improvements of Beg-LDDMM, as DARTEL [Ash07] and Stationary-LDDMM [HBO07]. Log-Euclidean framework brought new life also to the demons algorithm, that, in 2007, become the diffeomorphic demons [VPPA07]. Subsequent approaches involving the symmetrization of the energy function and the use of a different measures of similarity (local correlation coefficient instead of L^2) are proposed in symmetric log-demons [VPPA08] and LCC-demons [LAF$^+$13] respectively.

In the next section we will focus our attention on the diffeomorphic demons algorithm, as the starting point of the operation of log-composition, main subject of the following chapters.

1.3 Demons Algorithms: From Classic to Diffeomorphic

The first demons-based algorithm in image registration was proposed by [Thi98] in analogy with the Maxwell's demon in thermodynamics. This early version - often called *classic demons* - does not involves diffeomorphisms: the deformation is not bonded to any particular set of transformations and its smoothness is obtained with a Gaussian filter.

All the vectors applied to each voxel in the moving image are mutually independent, and are attracted by all of the voxels of the fixed image with similar intensity. The force of attraction are inspired by the optical flow equations [HS81], and the algorithm works under the hypothesis that the intensity of a moving object is constant over time and it is therefore not robust to noise.

The final deformation, solution of the registration problem is obtained composing at each step the previous transformation with an update. Indicating with φ_k the deformation obtained at the beginning of the k-th iteration and with $\delta\varphi_k$ the update computed at the same step, they can be expressed as the addition between the identity and a displacement field V or δV:

$$\varphi_k(\mathbf{x}) = \mathbf{x} + V_k(\mathbf{x})$$
$$\delta\varphi_k(\mathbf{x}) = \mathbf{x} + \delta V_k(\mathbf{x})$$

And with this notation the $k+1$-th deformation is computed by composition as:

$$\begin{aligned}\varphi_{k+1}(\mathbf{x}) :&= (\delta\varphi_k \circ \varphi_k)(\mathbf{x}) \\ &= \mathbf{x} + \delta V_k(\mathbf{x}) + V_k(\mathbf{x} + \delta V_k(\mathbf{x}))\end{aligned}$$

Since the third addend is close to $V_k(\mathbf{x})$, many implementations - as for example the open-source Insight Segmentation and Registration Toolkit (ITK) [YAL+02] - consider by default only the sum between V_{k+1} and V_k in the computation of the update:

$$\begin{aligned}\varphi_{k+1}(\mathbf{x}) :&= (\delta\varphi_k + \varphi_k)(\mathbf{x}) \\ &= \mathbf{x} + V_k(\mathbf{x}) + \delta V_k(\mathbf{x})\end{aligned}$$

Demons algorithms with this implementation of the update are called *additive demons*.

In [CBD+03] authors presents the PASHA demons as an extension of the classic demons, where a global energy function is considered and optimized according to an alternating minimization scheme. It is important to notice that again the PASHA algorithm does not involve any diffeomorphism, but it utilizes the framework presented in the previous section within maintaining the application of a Gaussian filter G to smooth the transformations:

$$\varphi_{k+1}(\mathbf{x}) := G_1(\varphi_k(\mathbf{x}) + G_2(\delta\varphi_k(\mathbf{x}))$$

In general, if G_1 is the identity the model is sometime called *fluid*, while if G_2 is the identity is called *elastic*.

Diffeomorphisms were introduced later within the demons algorithm (*diffeomorphic demons* [VPM+06]) after the presentation of the log-Euclidean framework [AFPA06]. To each stationary velocity field $V \in \mathcal{V}(\Omega)$ is associated a diffeomorphisms φ by the ODE $d\varphi/dt = V_{\varphi(t)}$, with the initial condition $\varphi(0) = \mathbf{x}$.
Using Lie theory, SVF are considered elements in the *Lie algebra* - vector space defined by the differentiable vector field over Ω, denoted with $\mathcal{V}(\Omega)$ or $\mathfrak{g}$ in Lie theory - while the set of diffeomorphisms defines a *Lie group* - denoted with Diff(Ω) or with $\mathbb{G}$ -.

Roughly speaking, the Lie algebra $\mathcal{V}(\Omega)$ is the tangent space (as local linear approximation) to the Lie group Diff(Ω), and these two spaces are in local correspondence thanks to two functions: the *Lie exponential* and the *Lie logarithm*. *Lie exponential* maps vector fields on the corresponding Lie group elements, while the *Lie logarithm* - inverse of the Lie exponential under some condition, see [DCDC76] or [Lee12] - maps each diffeomorphisms in the correspondent tangent vector field:

$$\varphi = \exp(V) \quad V = \log(\varphi) \qquad\qquad \varphi \in \mathbb{G} \quad V \in \mathfrak{g}$$

In this settings, the update can not be computed simply with a sum of vector fields, since it must reflect the composition of the corresponding diffeomorphisms in the Lie group.

Several approaches have been presented to face the problem of the computation of the update. Diffeomorphic demons compute the transformation at each step of the iterative algorithm as the composition between the diffeomorphism φ_k obtained at the previous step with the update δV_k, obtained with the optimization algorithm:

$$\varphi_{k+1} := \varphi_k \circ \exp(\delta V_k)$$

In a subsequent version, the log-demons [VPPA08], the composition is performed in the tangent space toward exponential and logarithm functions

$$V_{k+1} := \log(\exp(V_k) \circ \exp(\delta V_k)) \tag{1.2}$$

For this last computation, another theoretical element from the theory of Lie group has been utilized: the BCH formula. It provides the solution for Z of the equation

$$\exp(Z) = \exp(X) \circ \exp(Y)$$

As we will see in the subsequent sections, its solution involves an infinite series of nested Lie bracket that do not make its computation straightforward. To face the problem of its numerical approximation, whose solutions are utilized to solve 1.2, we define in this thesis an inner binary operation called log-composition:

$$X \oplus Y := \log(\exp(X) \circ \exp(Y)) \qquad \forall\ X, Y \in \mathfrak{g}$$

That in the seminal paper about the computation of the coefficients of the BCH formula [Dyn00] appears indicated with Φ.

The main aim of this research is to present a comparison between numerical methods for its computation. Before presenting some details of the mathematical theory that underpins the numerical methods it is important to notice that the practical applications of the Log-composition do not impact only the update's composition in the log-demons.

1.3.1 Possible applications of the Log-composition

In medical imaging there are several situations in which numerical methods and approximations passes through concepts equivalent to the log-composition. Its fast and accurate computation may therefore have an impact in the following 5 situations:

1. Symmetric diffeomorphic demons [VPPA08] - as introduced in equation 1.2.
2. Fast computation of the logarithm [BO08] - as discussed in chapter 4.
3. Calculus on diffusion tensor [AFPA06] - the log-composition appears as the dual operation of $\odot$ of the logarithmic multiplication for tensor defined at page 413. An approach to symmetric positive definite matrices that starts from the tangent space (where a metric can be directly computed without the application of the logarithm) may benefit of an accurate log-composition.
4. Image set classification [HWS$^+$] - as based on the log-euclidean metric on the group of symmetric positive definite matrices.
5. Computation of the the discrete ladder for the parallel transport - in equation (2) of page 11 fo the paper [LP14a], an equivalent of the log-composition is utilized to the computation of parallel transport. Reversing the procedure, parallel transport can be used for the computation of log-composition (as presented in 2.3.1). Therefore any other improvement of the computation of the log-composition can be applied in this context and it provides immediate results to compute the parallel transport.

The next chapter is aimed to the formal definition of the log-composition, underpinned with the tools from differential geometry theory, and to present two new numerical technique for its computation.

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 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 fo the Lie group $\mathbb{G}$ indicated with $\mathfrak{g}$. 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; for matrices will be indicated with I, for diffeomorphisms with 1. 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:
$$\exp((t+s)\mathbf{v}) = \gamma(t+s) = \gamma(t) \circ \gamma(s) = \exp(t\mathbf{v}) \exp(s\mathbf{v})$$
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: which means that 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 showed here for brevity):

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

$\square$

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 $\exp(\mathbf{c}\mathbf{v}\mathbf{c}^{-1}) = \mathbf{c} \exp(\mathbf{v})\mathbf{c}^{-1}$.

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:

$$\oplus : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g} \tag{2.6}$$
$$(\mathbf{v}_1, \mathbf{v}_2) \longmapsto \mathbf{v}_1 \oplus \mathbf{v}_2 = \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2)) \tag{2.7}$$

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 $C_{\mathfrak{g}}$.

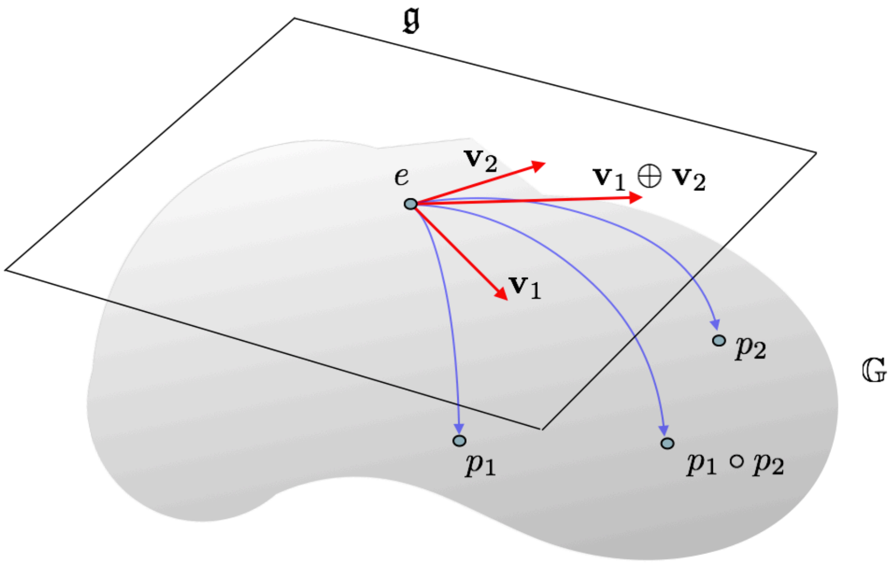

Figure 2.1: graphical visualization of the Lie log-composition $\mathbf{v}_1 \oplus \mathbf{v}_2$. The gray surface represents a Lie group and its tangent plane represents its Lie algebra.

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]. A nice introduction for the particular case of matrices can be found in [Hal15]; for the general case [KO89], [Ser09], and for application to medical imaging [VPPA08]. For our purposes, this expansion provides the most immediate way to obtain a numerical computation of $\mathbf{u} \oplus \mathbf{v}$, by truncating its terms. This approximation is problematic because it does not provides any error bound.

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

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 0 \tag{2.8}$$

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]:

$$\log_p(q) = DL_p(e)\log_e(q) \tag{2.9}$$

$$\exp_p(\mathbf{u}) = p \circ \exp_e(DL_{p^{-1}}(e)\mathbf{u}) \tag{2.10}$$

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

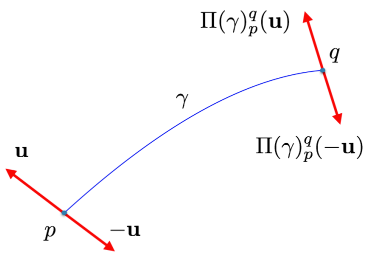

Figure 2.2: First inversion property.

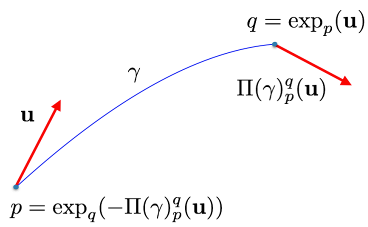

Figure 2.3: Second inversion property.

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) $\Pi(-\gamma)_q^p\left(\Pi(\gamma)_p^q(\mathbf{u})\right) = \mathbf{u}$ 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:

$$\exp_p(t\mathbf{u}_p) = p \circ \exp_e(tD(L_{p^{-1}})_p(\mathbf{u}_p)) = p \circ \exp_e(t\mathbf{u}) \tag{2.11}$$

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:

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

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 2.9. 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 $\mathbf{v}_p^{\parallel}$ parallel transport of $\mathbf{v}$

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

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

$$\mathbf{v}_e^{\parallel} := D(L_{p^{-1}})_e(-\Pi(\alpha)_p^e(\mathbf{v}_p^{\parallel}))$$

it follows that:

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

The statement of this fairly intricate theorem involves a construction that can be visualized in figure 2.4.

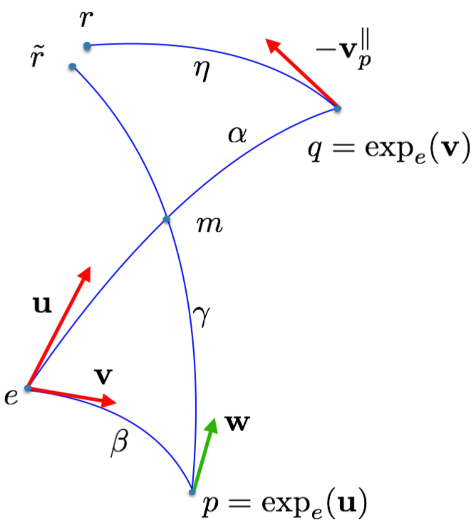

Figure 2.4: Pole ladder applied to parallel transport.

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 2.9 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:

$$\exp_e(\mathbf{v}_e^{\|}) \simeq \exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big) \tag{2.12}$$

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:

$$\mathbf{v}_1 \oplus \mathbf{v}_2 = \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2)) \tag{2.13}$$

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 $\mathfrak{g}$:

$$BCH(\mathbf{u},\mathbf{v}) = \mathbf{u} + \mathbf{v} + \frac{1}{2}[\mathbf{u},\mathbf{v}] + \frac{1}{12}([\mathbf{u},[\mathbf{u},\mathbf{v}]] + [\mathbf{v},[\mathbf{v},\mathbf{u}]]) - \frac{1}{24}[\mathbf{v},[\mathbf{u},[\mathbf{u},\mathbf{v}]]] + ... \tag{2.14}$$

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:

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

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:

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

$$\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_{\mathbf{u}}^{n}\mathbf{v} + \mathcal{O}(\mathbf{v}^2) \tag{2.15}$$

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 2.5). 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.

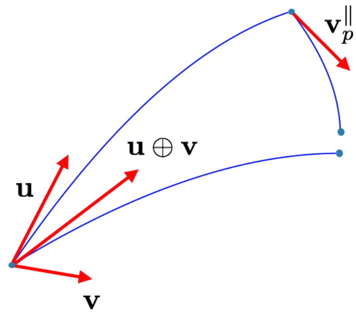

Figure 2.5: Representation of the intuitive idea of the computation of the log-composition using parallel transport.

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 2.12 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

$$\mathbf{u} \oplus \mathbf{v} \simeq \mathbf{u} + \exp_e\big(\frac{\mathbf{u}}{2}\big) \circ \exp_e(\mathbf{v}) \circ \exp_e\big(-\frac{\mathbf{u}}{2}\big) - e \tag{2.16}$$

When working in the infinite dimensional case, the approximation 2.16 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 1.2.1 about a naive approach to the infinite dimensional Lie Group theory.

With the truncated BCH and the Taylor expansion, equation 2.16 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

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 cases of $SE(2)$ and the for the Lie group of diffeomorphisms parametrized with stationary velocity fields:

$SE(2)$ - The group of rigid body transformation of the plane (any combination of bi-dimensional rotations and translations) is a good playground to test the numerical methods introduced so far, since results can be compared with a closed form. A representation of this Lie group as a subgroup of the general linear group $GL(2)$, with corresponding Lie algebra will be provided, with closed form for the log-computation.

Diff(Ω) - The group of diffeomorphisms over Ω, indcated with Diff(Ω) is defined over the wide set of all of the differentiable and invertible functions. For our applications we will restrict the set to the diffeomorphisms that can be parametrized by stationary velocity fields or SVF. This set, indicated with Γ has remarkable property thanks to the one-parameter subgroup, and its elements are in correspondence with SVF that belongs to the Lie algebra of Diff(Ω). The infinite dimensional vector space of stationary velocity field, is the second group utilized to test the numerical methods here presented for the computation of the log-composition. In this case we do not know any closed form, but considering an "improper norm" in the space of deformations it is possible to compare SVF and assess the quality of the results.

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 defined by θ belongs to the special orthogonal group $SO(2)$. 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 (see [Hal15, Gal11] for an introduction).

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

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 therefore

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

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 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 $\frac{\theta_0 \mod 2\pi}{\theta_1}$.

   To have a bijective correspondence we have to restrict the domain of exp 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

$$(\theta_0, dt_0^x, dt_0^y) \sim (\theta_1, dt_1^x, dt_1^y)$$
$$\overset{\text{def}}{\iff}$$
$$\exists k \in \mathbb{Z} \mid \theta_0 = \theta_1 + 2k\pi \ \text{ and } \ (dt_0^x, dt_0^y) = \frac{\theta_0 \mod 2\pi}{\theta_1}(dt_1^x, dt_1^y)$$

   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 domainexp 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

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

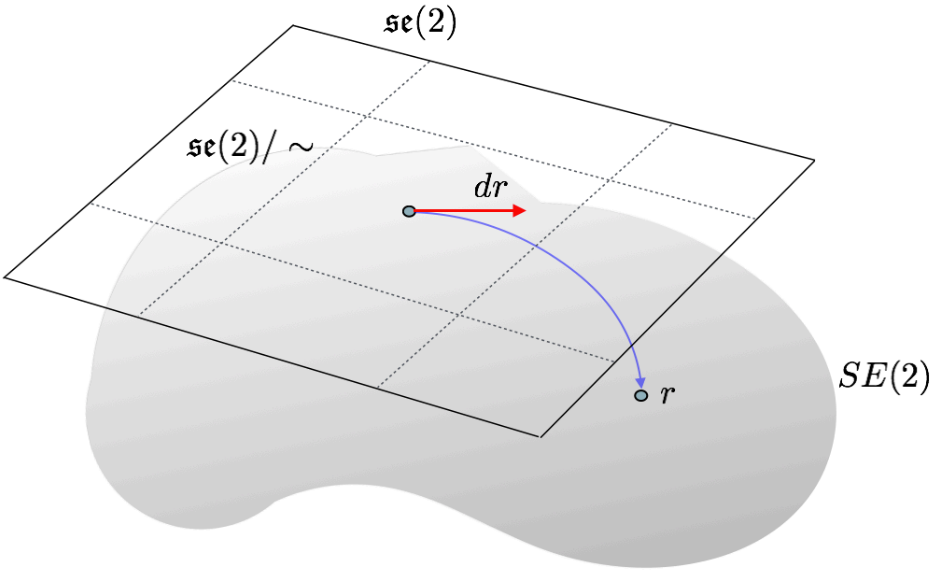

Figure 3.1: The Lie algebra $\mathfrak{se}(2)/\sim$ defined as the quotient of the Lie algebra $\mathfrak{se}(2)$ over the equivalence relation $\sim$ is in bijective correspondence with $SE(2)$.

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

$$dr_0 \oplus dr_1 \simeq BCH^0(dr_0, dr_1) = dr_0 + dr_1$$

$$dr_0 \oplus dr_1 \simeq BCH^1(dr_0, dr_1) = dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1]$$

$$dr_0 \oplus dr_1 \simeq BCH^{3/2}(dr_0, dr_1) = dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1] + \frac{1}{12}[dr_0, [dr_0, dr_1]]$$

$$dr_0 \oplus dr_1 \simeq BCH^2(dr_0, dr_1) = dr_0 + dr_1 + \frac{1}{2}[dr_0, dr_1] + \frac{1}{12}([dr_0, [dr_0, dr_1]] + [dr_1, [dr_1, dr_0]])$$

To compute the approximation with the Taylor method, and so to compute the equation 2.15, 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 2.15 of the log composition, indicating $dt^\star = (dt^y, -dt^x)$ it can be proved easily by induction that

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

And so the series involved in the equation 2.15 become

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

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, using 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]:

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

where

$$\mathbf{J}(dr_0, dr_1) = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\theta_0/2-\tan(\theta_0/2)}{\theta_0\tan(\theta_0/2)}dt_0^x + \frac{1}{2}dt_0^y & \frac{\theta_0/2}{\tan(\theta_0/2)} & -\theta_0/2 \\ -\frac{1}{2}dt_0^x - \frac{\theta_0/2-\tan(\theta_0/2)}{\theta_0\tan(\theta_0/2)}dt_0^y & \theta_0/2 & \frac{\theta_0/2}{\tan(\theta_0/2)} \end{pmatrix}$$

therefore the corresponding numerical method indicated with the function Tl as

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

The approximation of the log-composition using parallel transport is a straightforward application of the equation 2.16:

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

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

As previously said in section 1.2.1, the passage from the finite to the infinite dimensional case is not free of deceptions. We will investigate in the next two subsections in particular the following facts that happen for matrices:

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

2. Lie logarithm and Lie exponential are local isomorphisms.

And do not happen in general for diffeomorphisms.

3.2.1 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 importance of the existence of a bigger algebra is not only a theoretical problem: the power series expansions of the exponential and the logarithm 2.3 2.2 as well as expressions as 2.4 and 2.5 would me meaningless without the possibility of expressing the sum of two elements of a multiplicative group. In addition if the bigger algebra that contains both Lie group and Lie algebra exists, a unique norm in this space can be defined and used to compare elements in both of the subspaces.

In the case of diffeomorphism of $\Omega \subset \mathbb{R}^d$, 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 definition at hand.

We define the set of *deformations*, the set of continuous functions from Ω to Ω (a strategy to avoid issues on the boundary of Ω is to consider the deformations over the whole $\mathbb{R}^d$ that are the identity outside Ω; for our purposes we will consider these definitions equivalent). If a deformation is invertible with continuous inverse then it is called *homeomorphism*; the set of homeomorphisms forms a group with the operation of function composition , indicated with $\mathrm{Hom}(\Omega)$. 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 function that at each point of Ω, associates a vector of $\mathbb{R}^d$; the set of velocity vector fields, indicated with $\mathcal{V}(\Omega)$ forms a vector space, and considering the multiplication with the Jacobian matrix at each point, it forms a non-commutative group:

$$\mathbf{u} \cdot \mathbf{v} := J_{\mathbf{u}}\mathbf{v} \qquad \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}(\Omega)$$

And thanks to the linearity of the Jacobian, $\mathcal{V}(\Omega)$ forms an algebra.

There are two ways to associate a velocity vector field to a diffeomorphisms φ. The first one is subtracting 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 *deformation 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 deformation 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. 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 3.2). The very same happens for matrices, as noticed in equations 3.1.

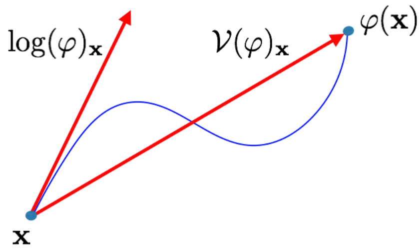

Figure 3.2: for small deformations, the deformation field $\log(\varphi)$ and the tangent field $\mathcal{V}(\varphi)$ are close to each others.

At this point it is important to notice that, while a deformation 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.

3.2.2 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 always 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 and therefore can not be related with any element in the tangent space by an ODE (see the counterexample in [LP13], pag. 6 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, investigate this feature is not only purely a theoretical issue. As happened in the previous section, it is necessarily to have some definitions at hand before moving in this direction.

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 *Diff*(Ω) (see for example [KW08]) 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, by the 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 $\frac{d\varphi_1(t)}{dt} = V_{\varphi_1(t)}$ and $\frac{d\varphi_2(t)}{dt} = U_{\varphi_2(t)}$, then their composition $\varphi_1 \circ \varphi_2$ does not satisfy any stationary ordinary differential equation, but it satisfy a 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 $\psi^{(t)}(\mathbf{x})$. It is important to notice 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 spatial dimension of the space. The vector that influence the satellite is an SVF for every point in the space-time.

It follows that tationary ODE and non-stationary ODE have solutions that belong to *Diff*$^1(\Omega)$ and *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 *Diff*(Ω), but for two different instant of time, the solution can belongs to two different one parameter subgroup.

To conclude this part, 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 but the one that satisfies 3.7, and therefore corresponds to TVVF does. 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 1.2.2) TVVFs are initially considered, while with the paper of Arsigny [ACPA06a] and in subsequent works the attention has been restricted to SVF, 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.

3.2.3 Norm for elements in the one-parameter subgroup

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

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

and induces a metric to compute the distance between stationary velocity fields. In the Lie group $\mathit{Diff}^1(\Omega)$ do not possess any norm, but the corresponding deformation fields defined by $\mathcal{V}$, as tangent vector fields does. Given two diffeomorphisms φ_0, φ_1 we have

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

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 $\mathit{Diff}^1(\Omega)$ and SVF:

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

The next

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 deformation fields are 5-dimensional matrices

$$M = M(x_i, y_j, z_k, t, d) \qquad (i, j, k) \in L, \quad t \in T \quad d = 1, 2, 3 \tag{3.8}$$

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 $\mathit{Diff}^1(\Omega)$ 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:

$$[\mathbf{u}, \mathbf{v}] := J_u\mathbf{v} - J_u\mathbf{v} \qquad \forall \mathbf{u}, \mathbf{v} \in \mathfrak{g} \tag{3.9}$$

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 2.14 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 2.16:

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

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 [ACPA06a] we use the scaling and squaring, keeping in mind that this choice impact on the results.

Chapter 4

Log-Algorithm using Log-composition

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]). Here we are interested in the numerical iterative algorithm for the computation of the Lie logarithm, called here *log-algorithm*, presented for the first time in [BO08]. In this chapter we present a strong relation between the log-algorithm and the log-composition: in consequence of this, each numerical methods presented in these pages can be applied to find a numerical method to solve the logarithm computation problem.
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 of the previous chapter for the particular case of $SE(2)$, if the matrix dr is small enough we can approximate $\exp(dr)$ with $1 + dr$. Aim of this section is to generalize the same approximation for the SVF.
We define two approximating functions:

$$\begin{aligned} \mathrm{app} : \mathfrak{g} &\longrightarrow \mathfrak{g}^{\sim} \\ \mathbf{u} &\longmapsto \exp(\mathbf{u}) - 1 \end{aligned}$$

$$\begin{aligned} \mathrm{App} : \mathbb{G} &\longrightarrow \mathbb{G}^{\sim} \\ \exp(\mathbf{u}) &\longmapsto 1 + \mathbf{u} \end{aligned}$$

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

$$\mathfrak{g}^{\sim} := \{\exp(\mathbf{u}) - 1 \mid \mathbf{u} \in \mathfrak{g}\} \cup \mathfrak{g}$$
$$\mathbb{G}^{\sim} := \{1 + \mathbf{u} \mid \mathbf{u} \in \mathbb{G}\} \cup \mathbb{G}$$

In general $\mathfrak{g}^{\sim} \neq \mathfrak{g}$ and $\mathbb{G}^{\sim} \neq \mathbb{G}$, but in the considered cases of $\mathfrak{se}(2)$ and SVF, when $\mathbf{u}$ is *small enough* it follows that $\exp(\mathbf{u}) - 1 \in \mathfrak{g}$ and $1 + \mathbf{u} \in \mathbb{G}$. Therefore the elements of $\mathfrak{g}^{\sim}$ are compatible with all of the operations of Lie algebra $\mathfrak{g}$ and the elements of $\mathbb{G}^{\sim}$ 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 $\mathbf{u}$ 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 of the same norm, then $e + \mathbf{u}$ is a diffeomorphism. With this condition holds that $\text{SVF}^{\sim} = \text{SVF}$.

Therefore, for each small enough $\mathbf{u}$ in $\mathfrak{se}(2)$ or SVF, and considering the definition of the log-composition (equation 2.6) the following properties holds:

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

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 Log-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:

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

Each strategy that we have examined to compute the Lie 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:

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathrm{BCH}^k(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) \end{cases} \tag{4.2}$$

For $k = 0$, the approximation $\mathbf{u}_{j+1}$ results simply the sum $\mathbf{u}_j + \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})$. When $k = 1$, it results

$$\begin{aligned} \mathrm{BCH}^1(\mathbf{u}_j, \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u})) &= \mathbf{u}_j + \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) \\ &= \mathbf{u}_j + \exp(-\mathbf{u}_j)\exp(\mathbf{u}) - 1 \end{aligned}$$

And $k = 2$ it follows

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

The following theorem presented in [BO08], provides an error bound when $k = \infty$ so when the BCH formula is used, instead one of its truncation.

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

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j \oplus \mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) \end{cases} \tag{4.3}$$

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 log-algorithm:

$$\begin{cases} \mathbf{u}_0 = \mathbf{0} \\ \mathbf{u}_t = \mathbf{u}_{t-1} + \exp(\frac{\mathbf{u}_{t-1}}{2}) \circ \exp(\delta\mathbf{u}_{t-1}) \circ \exp(-\frac{\mathbf{u}_{t-1}}{2}) - e \end{cases} \tag{4.4}$$

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

$$\begin{cases} \mathbf{u}_0 = 0 \\ \mathbf{u}_{j+1} = \mathbf{u}_j \oplus \frac{1}{2}\big(\mathrm{app}(-\mathbf{u}_j \oplus \mathbf{u}) + \mathrm{app}(\mathbf{u} \oplus -\mathbf{u}_j)\big) \end{cases} \tag{4.5}$$

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

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

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.*

In **chapter 1** the concept of log-composition is introduced, emphasizing its implications in medical imaging as a tool utilized in diffeomorphic registration and in the computation of the logarithm in the log-Euclidean framework. **Chapter 2** is devoted to the introduction of the underpinning mathematical theory: it defines formally the log-composition and presents three numerical methods for its computation:

1. Truncated BCH formula of degree $k = 1, \frac{3}{2}, 2, 3$ - *equation 2.14*.
2. Taylor expansion - *equation 2.15*.
3. Parallel transport - *equation 2.16*.

Before evaluating their results on the SVF, some tests in the are is evaluated for two groups of transformation, **chapter 3** introduces two groups of transformations:

1. The finite dimensional Lie group of euclidean transformation SE(2), where a closed form of the log-composition is known - *section 3.1*
2. The infinite dimensional Lie group diffeomorphisms, set of images of SVF through the Lie exponential map - *section 3.2*

For each of these groups it presents as well the numerical methods for the computation log-composition shown in the previous chapter for the general case. **Chapter 4** is about the algorithm for the computation of the Lie logarithm [BO08], called here log-algorithm. Thanks to the fact that this important piece in the jigsaw puzzle of the log-euclidean framework can be reformulated in term of the log-composition, it is possible to compute it using numerical methods introduced:

1. Truncated BCH formula of degree $k = 1, \frac{3}{2}, 2, 3$ - *equation 4.2*.
2. Parallel transport - *equation 4.4*.
3. Symmetric parallel transport - *equation 4.5*.

This last chapter is devoted to show some of the results of the numerical methods investigated. Computations are performed with a software written in Python (repository available on the cmic gitlab), based on the following libraries - numpy, matplotlib [Hun07], math, scipy, nibabel, timeit, random - as well as on the library NiftyBit, implemented by Pancaj Daga.

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.

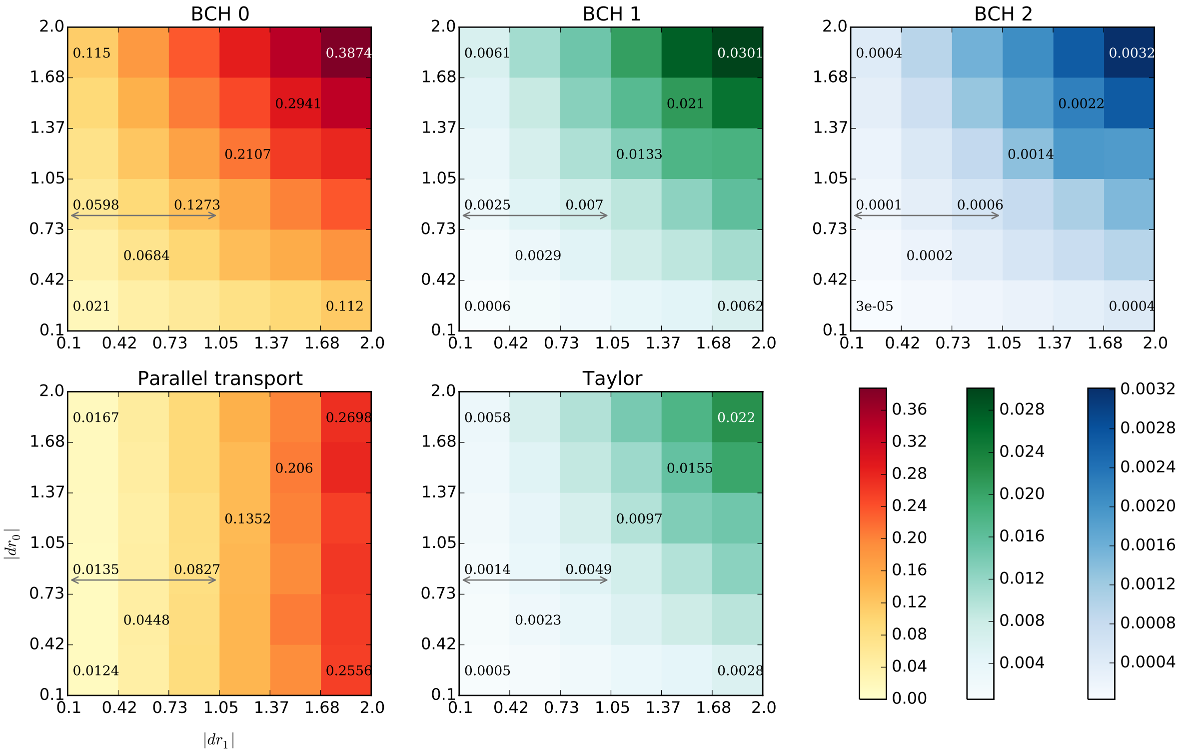

Figure 5.1: Comparison of the errors for each numerical method to compute the Log-composition $dr_0 \oplus dr_1$ in $\mathfrak{se}(2)$. Truncated BCH of degrees 0,1,2, parallel transport method and Taylor method are considered for different values of the norm of dr_1 (x-axes) and norm of dr_0 (x-axes). Values of each sub-square are 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, 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.

5.1.1 Methods and Results

To compare the errors the computation of the log-composition for the methods 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 $I =$ 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:

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

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

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

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:

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

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

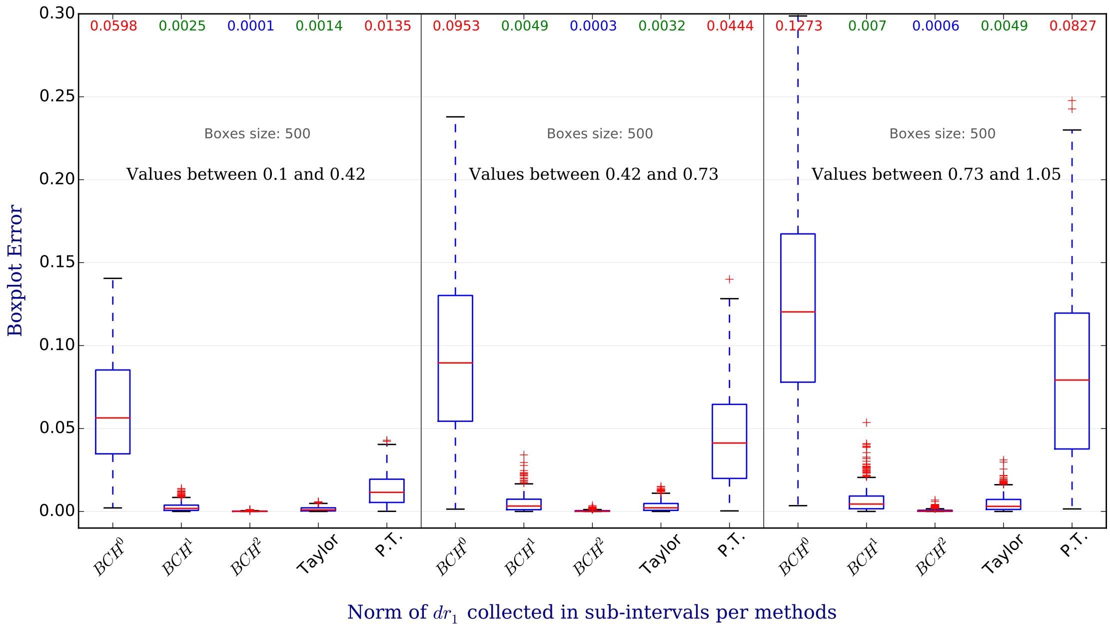

Figure 5.2: Errors of the numerical methods for the computation of the Log-composition of $dr_0 \oplus dr_1$ in $\mathfrak{se}(2)$. Norm of dr_0 is in the interval $[0.37, 1.05]$, norm of dr_1 in the interval $[0.1, 1.05]$ divided in 3 segments. Mean values of each box are shown in the first row in different 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.

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, ...).

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 $\mathfrak{se}(2)$.

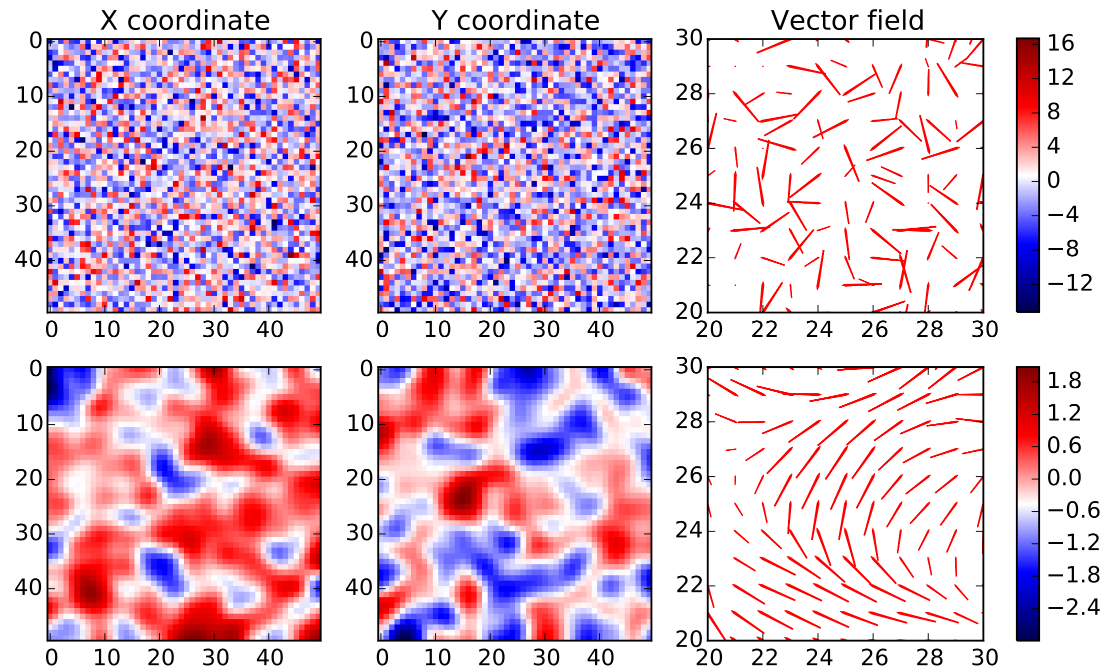

Figure 5.3: Random generated vector field before and after the Gaussian smoother: in the first row a random generated vector field of dimension $50 \times 50 \times 2$ where the value at each pixel are sampled from a random variable with normal distribution of mean 0 and sigma 4. The second row shows 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 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.

5.2.1 Methods: random generated SVF.

DRAFT:

- How to generate a random SVF - formula refer to figure 5.3
- Norm defined in both Lie algebra and Lie group, thanks to the fact that ... inheritance. formula
- How the norm changes with the space and with the filter: 5.4
- How the norm affect the Lie bracket:

We will exploit the parametrization of discretized SVF using matrices to have a ground truth to compare results.

Norm will be computed in the group as the L_2 norm of matrices that represents the SVF. Given $\mathbf{u}$ and $\mathbf{v}$ in $\mathit{Diff}^1(\Omega)$, $\mathbf{w}_{\text{ground}} = \mathbf{u} \oplus \mathbf{v}$ solution of the log-composition and $\mathbf{w}_{\text{app}}$ its approximation using a numerical method, then their difference is computed in the group as:

$$\text{error} = ||\exp(\mathbf{w}_{\text{ground}}) - \exp(\mathbf{w}_{\text{app}})||_{L^2}$$

where $\exp(\mathbf{w}_{\text{ground}})$ is computed as the composition of the exponentials of $\mathbf{u}$ and $\mathbf{v}$:

$$\exp(\mathbf{w}_{\text{ground}}) = \exp(\mathbf{u}) \circ \exp(\mathbf{v})$$

As previously said, the norm L^2 is considered improperly in a group structure. It can be done only thanks to the fact that the discrete SVF and the corresponding diffeomorphisms $\mathit{Diff}^1(\Omega)$ are implemented with 5-dimensional matrices (see equation 3.8).

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

- Other problems of the computation in consequence of the fact that the Jacobian is utilized in the Lie bracket from Marcos' paper. ...
- Table with errors!
- why we will take into account only the BCH 0 and the BCH 1 methods.

5.2.3 Log-composition for SVF: errors

Conclusions for syntetic data: where are the errors and why we are moving towards the real data.

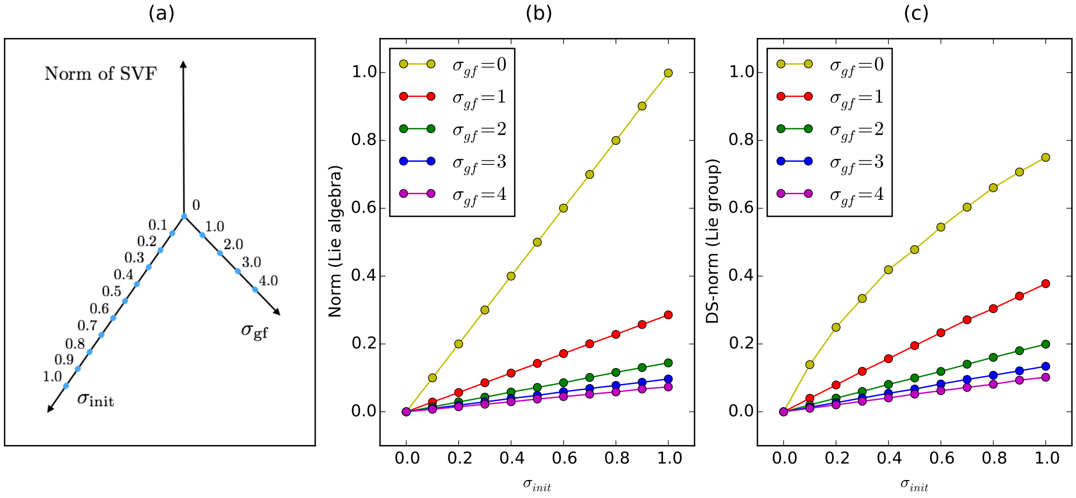

Figure 5.4: Relationship between the initial standard deviation, the standard deviation of the Gaussian filter and the norm. Figure (a) represents... Norm of random generated of SVF with initial standard deviation σ_{init} (on the x-axis) and Gaussian filter with standard deviation σ_{gf} (different colors). On the left is shown the the Frobenius norm computed on the SVF in the Lie algebra, while on the right the same norm is computed after the exponentiations. In this second case, the norm refers to the norm of the matrix data structure (DS-norm) utilized to parametrize the SVF. Each dot represents the mean of the norm of 10 an SVF randomly generated with the parameters indicated on the axes and in the legend. We observe that the exponential bend the shape of the random SVF when the Gaussian filter is 0 (thus we talk about an improper SVF). The decrease in the slope when $\sigma_{\text{gf}} = 0$ do not appears for any other value.

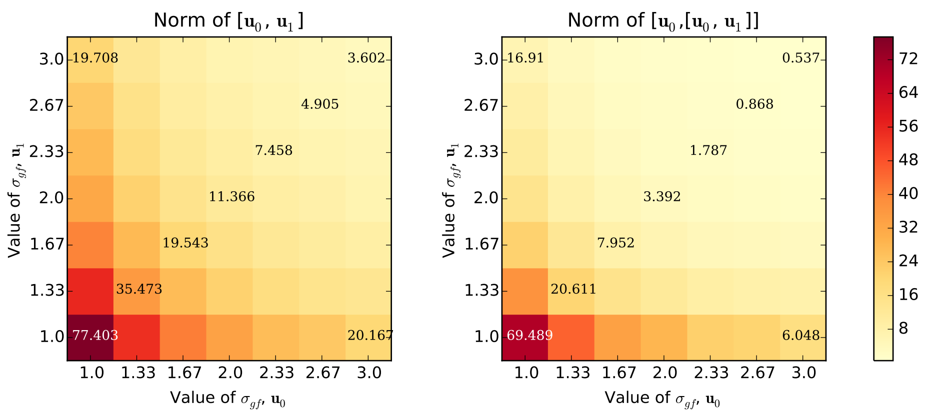

Figure 5.5: Relationship between the norm of the Lie bracket and the Gaussian smoother. For the provided SVF... Here put every other data!

5.3 Log-composition of real cases SVF

comment figures 5.6, 5.7, 5.8.

zzz see of you have time!

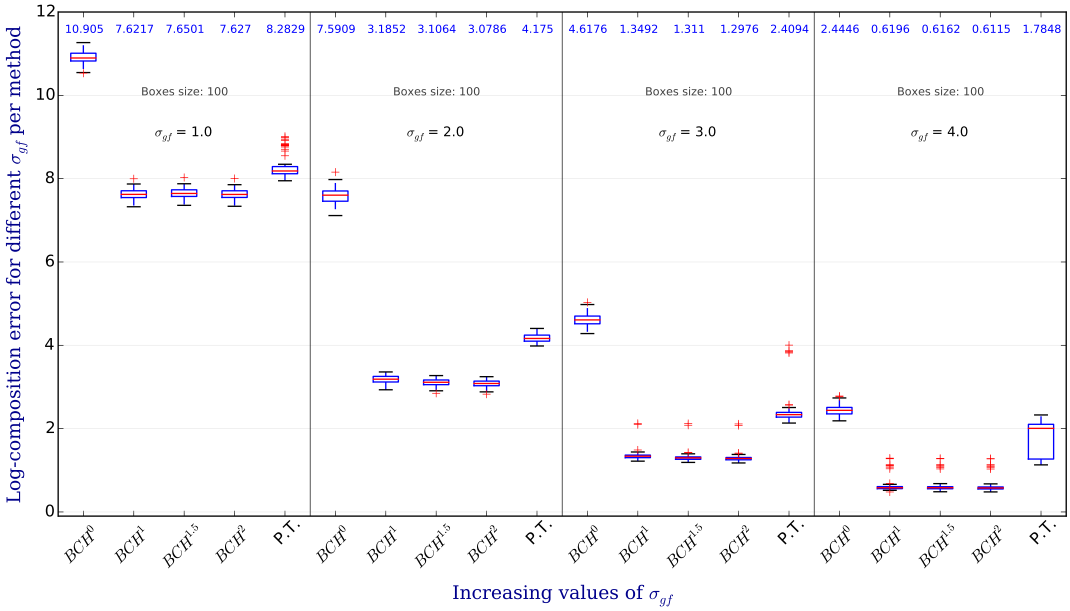

Figure 5.6: Boxplot to compare the error between various truncated BCH methods for SVF

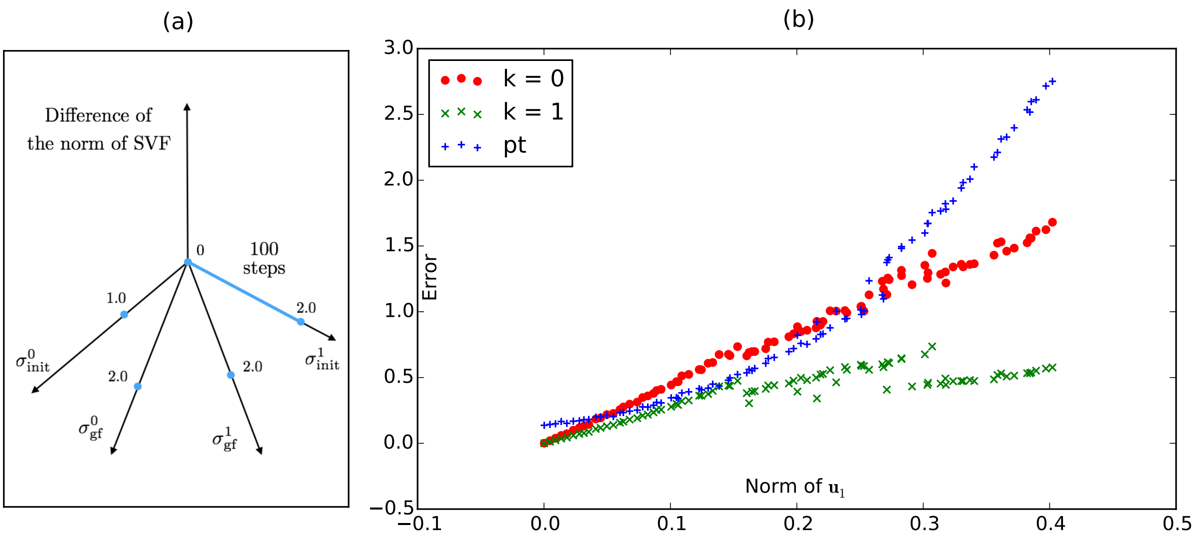

Figure 5.7: Log-composition for SVF computed using numerical methods of truncated BCH of degree 0,1 and parallel transport.

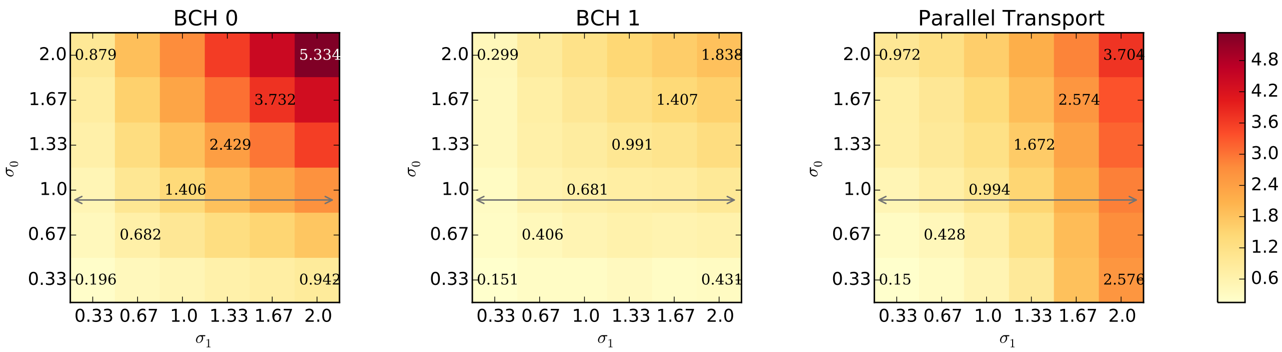

Figure 5.8: Log-composition for SVF; the operation $\mathbf{u}_0 \oplus \mathbf{u}_1$ is computed using numerical methods of truncated BCH of degree 0,1 and parallel transport. Respective standard deviation of the random generated SVF given by σ_0 and σ_1, ranges between 0.3 and 2.0 for σ_0 and between 0.2 and 2.0 for σ_1. Each value in the image scale is the mean of 10 results of the log-computation of random SVF generated with given standard deviation. For lower values of $\mathbf{u}_1$, that in the image registration algorithms are given by the update, parallel transport method and truncated BCH of degree 1 have comparable results. We can also notice that for truncated BCH methods the results are symmetric, while for parallel transport, as expected from the formula, results are not symmetric respect to the size of the input vectors.

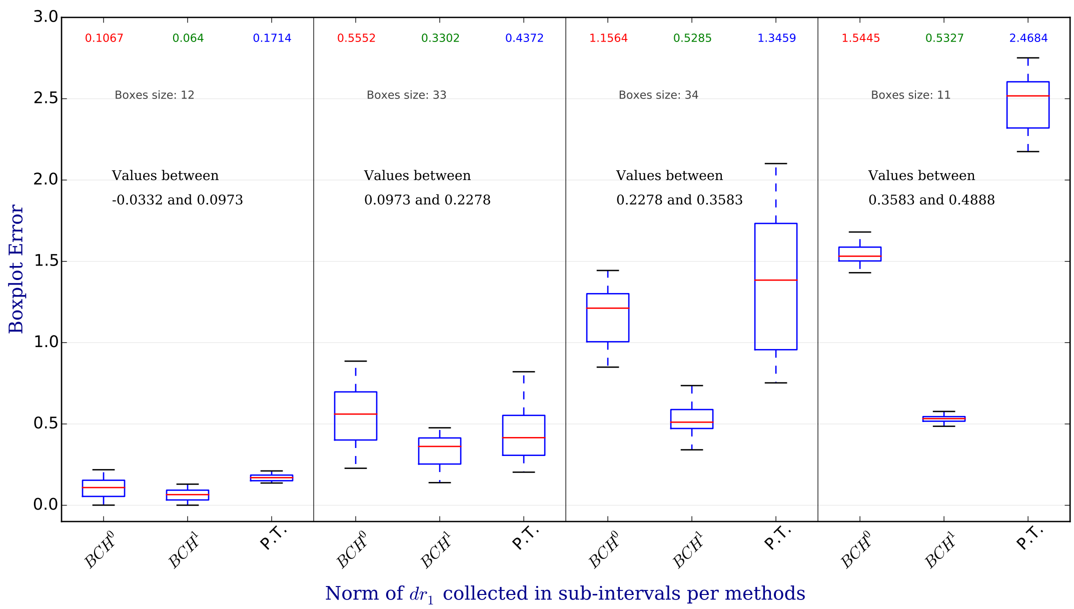

Figure 5.9: Log-composition for SVF computed using numerical methods of truncated BCH of degree 0,1 and parallel transport, represented in a boxplot.

5.4 Log-Algorithm for SVF

5.4.1 Methods

5.4.2 Results

5.5 Empirical Evaluations of Computational Time

5.6 Conclusions and Further Research

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

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.

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,

Bibliography

[ACPA06a] 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.

[ACPA06b] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. Statistics on diffeomorphisms in a log-euclidean framework. In *1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical, Statistical and Registration Methods for Modeling Biological Shape Variability*, pages 14–15, 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.

[Ash07] J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113, 2007.

[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.

[BMTY05] M Faisal Beg, Michael I Miller, Alain Trouvé, and Laurent Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision*, 61(2):139–157, 2005.

[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.

[Bro81] Chaim Broit. *Optimal Registration of Deformed Images.* PhD thesis, Philadelphia, PA, USA, 1981. AAI8207933.

[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.

[CBD+03] Pascal Cachier, Eric Bardinet, Didier Dormont, Xavier Pennec, and Nicholas Ayache. Iconic feature based nonrigid registration: the pasha algorithm. *Computer vision and image understanding*, 89(2):272–298, 2003.

[CRM96] Gary E Christensen, Richard D Rabbitt, and Michael I Miller. Deformable templates using large deformation kinematics. *Image Processing, IEEE Transactions on*, 5(10):1435–1447, 1996.

[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.

[DGM98] Paul Dupuis, Ulf Grenander, and Michael I. Miller. Variational problems on flows of diffeomorphisms for image matching, 1998.

[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.

[EMP06] DAVID G Ebin, GERARD Misiołek, and STEPHEN C Preston. Singularities of the exponential map on the volume-preserving diffeomorphism group. *Geometric and Functional Analysis*, 16(4):850–868, 2006.

[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.

[GM98] Ulf Grenander and Michael I Miller. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics*, 56(4):617–694, 1998.

[Gon] Rafael C Gonzalez. Digital image processing using matlab.

[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.

[HBO07] Monica Hernandez, Matias N Bossa, and Salvador Olmos. Registration of anatomical images using geodesic paths of diffeomorphisms parameterized with stationary vector fields. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

[HS81] Berthold K Horn and Brian G Schunck. Determining optical flow. In *1981 Technical Symposium East*, pages 319–331. International Society for Optics and Photonics, 1981.

[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.

[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.

[LAF$^+$13] Marco Lorenzi, Nicholas Ayache, Giovanni B Frisoni, Xavier Pennec, and Alzheimer's Disease Neuroimaging Initiative. LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. *NeuroImage*, 81:470–483, 2013.

[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.

[Lem97] László Lempert. The problem of complexifying a lie group. *Contemporary Mathematics*, 205:169–176, 1997.

[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.

[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.

[PCA99] Xavier Pennec, Pascal Cachier, and Nicholas Ayache. Understanding the "demon's algorithm": 3d non-rigid registration by gradient descent. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI'99*, pages 597–605. Springer, 1999.

[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.

[Thi98] J-P Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical image analysis*, 2(3):243–260, 1998.

[Tro98] Alain Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.

[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.

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

[YAL$^+$02] Terry S Yoo, Michael J Ackerman, William E Lorensen, Will Schroeder, Vikram Chalana, Stephen Aylward, Dimitris Metaxas, and Ross Whitaker. Engineering and algorithm design for an image processing api: a technical report on itk-the insight toolkit. *Studies in health technology and informatics*, pages 586–592, 2002.

[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.

[ZF03] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.