Toward a BCH-free Computation of the Composition of Stationary Velocity Fields

Sebastiano Ferraris

Universtiy College London
Medical Physiscs and Biomedical Engineering

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

May 12, 2015

Supervisor
Tom Vercauteren

Co-Supervisor
Marc Modat

Abstract

10 lines of abstract! This must stay all in 1 page !!!

Thesis' Organization

Four chapters, all alike in dignity... about the general framework in image registration and parametrization of diffeomorphisms for LDDMM and SVF is followed by three distinct part. The first one presents basic tools of differential geometry and parallel transport, with emphasis on the computational side. The second part is about the principal objects utilized to explore new numerical techniques and to compare them with the one currently utilized. The last part is devoted to the results for the computation on synthetic dataset and on patient images.

Chapter ??: This first part Introduction and Frameworks. After a first section about the main definitions and concepts utilized throughout the thesis, general registration framework's features are presented. Particular attention is given to the pros ad cons of in the use of diffeomorphism as set of transformation between anatomies and some considerations about methods currently used for their parametrization: the LDDMM and SVF.

Chapter 2: Main mathematical elements and tools from Lie group theory directly involved in the image registration techniques are formally defined with a particular attention to flows, left translation, push forward, Lie logarithm and Lie exponential. We define the concept of Log-composition around which the research gravitates: it originates form the need of generalized the BCH formula. In this context the BCH become one possible way to compute the Log-composition. The second way to compute it in finite dimensional case, is provided by the Taylor expansion, presented in the last section. zzzBCH and Taylor expansion are two possibility to compute the Log-composition. A third one presented in this chapter originates by a geometrical approach and it is given by the parallel transport. First and Second sections are devoted to present the theoretical tools to define formally the parallel transport. Last section is about two strategies to compute the parallel transport without involving the Christoffel symbols: the Schild's Ladder and the Pole Ladder.

Chapter ??: Validity of results in the Log-composition computations are tested with two groups of transformations commonly used in medical image registration: the group of rigid body transformation and the group of diffeomorphisms (expressed in the application as the set of Stationary Velocity Fields). These chapters are aimed to present them in details and they are oriented to the application.

zzzThis is the central part of the research. The Log-composition is analyzed as s valuable tool in image registration, within the framework presented in chapter 1.2. A summary of the methods for its computation is presented as possible numerical approximation to be utilized for image registration: BCH formula, Taylor expansion, parallel transport and Accelerating Convergences series. zzz The algorithm for the Lie-logarithm computation presented in *A new algorithm for the computation of the group logarithm of diffeomorphism* [BO08] is based on the computation of the BCH formula. If reformulated with the Log-composition, each of its numerical approximation is a valid tool to improve its performance. Of particular interests are the methods that avoid the computation of the BCH formula on which the algorithm was initially based.

Chapter 4: This chapter is devoted to experimental results. Performance of the Log-composition applied to rigid body transformation and diffeomorphisms are separately computed and compared. In addition a version of NiftyReg based on various we present the results of the numerical methods presented in the previous section, on synthetic data as well as on clinical data within a version of the LCC-Demons customized with parallel transport. zzz Conclusion of what has been done so far (with a shameless and challenging emphasis of what is missing and what is still to be done).

Contents

Chapter 1

Introduction and Framework

This part is devoted to...!

1.1 Motivations

Two instruments from Machine Vision and Image Processing are mostly utilized in Medical Imaging to reveal internal features and compare anatomies: segmentation and registration. Segmentation consists in enhance contours, detect edges and reveal hidden structure while registration is the process of determining correspondences between two or more images acquired from patients scans.
Dealing with image registration problem means search for a solution to an ill-posed problem. Transformations between anatomies are not unique, and the impossibility of recover the spatial or temporal evolution of a anatomical transformation from temporally isolated images, makes any validation a difficult, if not impossible task. Among all of the possible voxelwise mappings that transform one image into another, interest may vary according to the requirements of each specific clinical situation[1].
In brain imaging, for example, registration can be performed to examine differences between subjects and distinguish affected from unaffected patients. This often results in a better understanding of the disease's features. Another case may require to compare different acquisition of the same subject, before and after a surgery or after a fixed period of time: in each case parameters and features of the transformation are completely different. Customized registration techniques are required also when dealing with distortion motion correction caused by cardiac pulses or respiratory cycles, or for the mosaicing of several images acquired with a limited field of view, for example through a optical fiber probe. Also the correction of motion distortion in image's acquisition phase, the mosaicing of several images or the construction of the model of cardiac and respiratory motion require customized image registration techniques.
Any approach is therefore varied and flexible: this led to a wide range of variants that has been proposed by researchers in the last decades[2].
There are several features that distinguish image registration's algorithms, but the most relevant it is the choice of the family to whom the transformation belongs. Since anatomies are in a continuous process of modification over time, in general without any variation in

[1] A recent survey in medical image registration can be found in [SDP13].
[2] A quick glance to Google scholar reveals about 1200000 papers in *medical image registration* (55% of the whole *image registration* resources).

the topological features, the use of diffeomorphisms to model transformation appears one of the most natural.
Continuous nature of these functions appears to be in contrast with the discrete nature of voxel images as well as with any computer's parametrization ability.
The approach of modeling with richer structure for simpler elements is actually very common even in everyday math: for example when measuring the diagonal of a 1 meters side squared table. The decimal unlimited non periodical $\sqrt{2}$ doesn't help until we don't consider it as an answer belonging to a larger-than-reality mathematical structure: computations and theorem (as the Pythagorean one) are well defined and meaningful.
The simple structure of a raster image as 3-dimensional matrix is really limited if compared to the continuous object they represent. Modeling with continuous function provides a structure that reflects the object's reality. In addition enable us to apply mathematical features from differential geometry and dynamical system theory.
The first idea of using smooth and continuous function for image registration goes back to the idea of using the Navier-Cauchy partial differential equation to model the deformation of images as two balancing forces applied to an elastic body [Bro81]. The solution's domain restriction to diffeomorphisms for the solution of the Lagrange transport equation for medical imaging registration appears in [DGM98a] and [Tro], and with his many variants is an active subject of study for research in mathematic applied to medical imaging. An important framework for the computation of image registration of diffeomorphism is provided by the Large Deformation Diffeomorphic Metric Mapping (LDDMM). Here diffeomorphisms are parametrized as ending point of integral curves of vector field on the Lie group of diffeomoprhism equipped with a Riemannian metric [DGM98b], [BMTY05]. In this framework, solid mathematical foundations are payed in term of computational complexity. Aimed to solve this issue, different parametrization of diffeomorphisms, as Stationary Velocity Fields [ACPA06b] has been embedded in the LDDMM framework. This approach gave birth to the DARTEL [Ash07] and the Stationary LDDMM [HBO07]. Starting from the Tririon's DEMON algorithm [Thi98], a different framework for diffeomorphic image registration was presented as Diffeomorphic Daemon [VPPA07] and the LCC-daemon [LAF$^+$13] . A comparison between stationary LDDMM and Diffeomorphic Demons with emphasis in both theoretical and practical aspects can be found in [HOP08].
The theme of diffeomorphism do not recurs only in medical application but it is a continuously improved subject of research also theoretical studies as [Mil84a], [BBHM11], [BHM10] or studies applied to other domains of science [Arn] [OKC92]. Geometry remains an important underpinning structure for many achievement in medical imaging and advanced researches are actively utilized for applications.
from the necessity of having statistics to infer the variability of anatomical structure while performing image registration, bring the field of Patter Analysis from Machine Vision to Medical Imaging, with the name of Computational Anatomy (cite survey computational anatomy). Statistics on diffeomorphisms are introduced using a metric space defined over the Lie algebras of the Lie group of diffeomorphisms: this idea took the name of log-euclidean framework. Proposed for the first time in 2004 and refined in by the same authors in 2006 [AFPA06], as a faster improvement to the affine invariant Riemannian approach, has found successful applications in many domains of medical imaging (in vivo mosaicing [Ver08], brain Alzheimer detection [Lor12], cardiac image analysis [MPS$^+$11], mandible imaging using polyaffine registration [CSR11]) and has been continuously improved.
Thus the use of diffeomorphisms brings several aspect to be studied and to consider in their utilization: theoretical research, their practical utilization, their study in statistical analysis and in the anatomical deformations' evolutions.
In this thesis we investigate some numerical methods to compute their composition in the

log-euclidean framework aimed to the diffeomorphic image registration. Numerical computations are performed using parallel transport and accelerating convergence series, and are compared with the currently used BCH formula, keeping into account theoretical and practical aspects of the matter.

1.2 Image Registration Framework

1.2.1 Introductory Definitions

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 the investigation of features and parameters of 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 $\mathbf{x} \in \Omega_R$ the element $M(\varphi(\mathbf{x}))$ and $F(\mathbf{x})$ are closed as possible according to a chosen metric. The function $M \circ \varphi$ is called *warped image*.
The underpinning idea can be represented by the following diagram, where φ is the solution that makes f the identity function.

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

This definition leave two degrees of freedom in searching for a solution: the domain of the transformation (also called deformation model), and the metric to measure the similarity between images.
Once these are chosen, they can be used as constituent of an *image registration framework*: an iterative process that at each step provides a new function φ that approaches f to the identity. Additional degrees of freedom can be defined in a refined framework: the metric can be considered with an additive regularization term, that introduces a constraint based on prior knowledge about the searched solution:

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

where Sim is a function that measure the similarity, while Reg is the regularization term. Other two degrees of freedom are optimization algorithm on which the optimizer is based and the resampling - process of resize the image from one dimension to another - strategy. It can be chosen over several possibilities (nearest neighborhood, linear, cubic, ...).

1.2.2 Iterative Registration Algorithm

The definition of registration problem and the iterative algorithm described raise several issues. For example there are no reasons to believe that such a correspondence is unique and

that there is at least one of them whose behaviour corresponds to a reasonable biological transformation between anatomies. One way to deal with this problem is tp add some constraints on the transformation φ, such that it models realistic changes that can occur in biological tissues. The kind and quality of the constraints are one of the feature that distinguish registration algorithms one from the other.
Conscious of the risk involved in any rigid classification, in particular when about rapidly evolving fields, we can see the image registration framework as a device with some knobs, each with its range:

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

and with an optimization technique, borrowed from the various from literature, as gradient descent, aimed to optimize the sum of the similarity and the regularization. Under the hood of this ideal device we may see something that can be schematically represented as in figure 1.1.

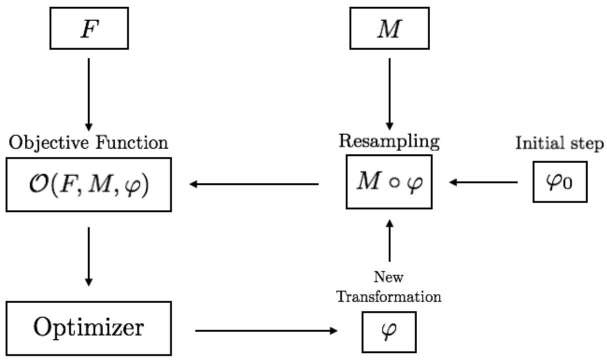

Figure 1.1: Image registration framework scheme.

Just from this simplified scheme we can see that at each step moving and fixed images are defined within their domain, not necessarily coincident. It is always possible to define a common domain $\Omega = \Omega_R$, called *background space*, on which both of the images are defined. Far from being a complete overview of all of the possible framework does not take into account the fact that each version or implementation inevitably involves different needs and consequent challenges. Solutions found for each case may fall outside this simplification scheme: for example the parametrization of the transformations (or the deformation field's update) at each iterative step do not appear in this picture, even though is a fundamental feature.
In the following section we will going from the generalized framework to some specific aspects: we are interested in performing numerical computation to compute diffeomorphisms (bijective differentiable maps with differentiable inverse[3]). They are particularly appealing

[3]Take the real valued $f(x) = x^3$: is bijective differentiable but the inverse is not everywhere differentiable.

in medical imaging since in many situation biological modification do not involve any change in tissues' topology. We will see some details of parametrization of diffeomorphisms in five different frameworks: the LDDMM, the Shooting LDDMM, the Stationary LDDMM (or SVF), the DARTEL and the Demon. In addition some details about the implementation of the image registration toolbox NiftyReg, used for some experiments in this research, can be found in the appendix.
Recent surveys in image registration can be found in [SDP13] and [HBHH01].

1.2.3 Using Diffeomorphisms: Utility and Liability

Idea of using diffeomorphisms can originate from two different perspective: if we imagine motion of captured images as motion of a fluid, we will start to approach the model using fluid dynamics; if we image motion as optical flow we will rely on differential geometry. Remarkably medical imaging application reveals common aspects of both underpinning theories, and requires some more since the discrete nature of actual input and output.
If we consider the rigid body transformation from classical mechanics, our transformations will be elements in $SE(3)$, the Lie group defined by combination of spatial rotations and translations. Five degrees of freedom, he possesses are enough to align 3d images; for continuous motions that maintains anatomical features between images they are not. We consider the Lie group of diffeomorphism $\text{Diff}(\Omega)$. As already said, these functions are particularly appealing in computational anatomy since their topology-preserving nature, but their mathematical formalization has encountered many issues.
Attempt to provide this object some handle to manipulate it easily was done for the first time in 1966 in [Arn66]: to solve differential equation in hydrodynamic $\text{Diff}(M)$ is considered as a Lie group with its Lie algebra. This assumption is not formally prosecuted in accordance to the problem-oriented nature of this paper[4]. The initial idea to consider $\text{Diff}(M)$ as a differentiable manifold involves to have it locally in correspondence with some generalized "infinite-dimensional euclidean "space. Attempt to set this correspondence showed that for some infinite-dimensional group the transition functions are smooth over Banach spaces. This led to the idea of Banach Manifolds. Unfortunately the group of diffeomorphisms do not belongs to the category of Banach manifold but requires a more generals space on which the transition map are smooth: the Frechet spaces. Thus the approach to the mathematical formalization of the general infinite dimensional Lie group involves Frechet differentiable manifolds. In this space we do not have anymore important theorems from analysis, as the inverse function theorem, or the main results from the Lie group theory in a finite dimensional settings, as Lie correspondence theorems.
These difficulties led some researcher in approaching the set of diffeomorphisms from other perspectives: for example, instead of treating $\text{Diff}(M)$ as group equipped with differential structures it is seen as a quotient of other well behaved group [Woj94].
Without denying the importance of fundamentals and underestimating the doors research in this domain may open, 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, where the fundamental definition of infinite dimensional Lie group is a generalization of the finite dimensional case left to the intuition. We work then mostly on finite dimensional settings, relying on important theorems and easy close formulas, and we will extend methods and results developed here in the infinite case -clearly - with proper precautions.

[4] Subsequent steps in the exploration of the set of diffeomorphisms as a Lie group are [MA70] and [Les83], [Omo70]. A survey on early development of infinite dimensional Lie group can be found in [Mil84b].

LDDMM

LDDMM [BMTY05] framework originates by considering motion between images as the motion of a fluid, and borrows its analytic tools from fluid dynamics. In this context the full set of homeomorphisms $\text{Hom}(\Omega)$ (continuous function from the background space Ω to itself with continuous inverse) is considered as a group with the composition. This group act[5] on the set of images defined on the background space $\mathcal{I}_\Omega$ as

$$\begin{aligned} \text{Hom}(\Omega) \times \mathcal{I}_\Omega &\longrightarrow \mathcal{I}_\Omega \\ (\varphi, F) &\longmapsto F \circ \varphi^{-1} \end{aligned}$$

The orbits of the subgroup of diffeomorphisms $\mathbb{G}$ on the image F, defined as

$$\mathcal{O}_{\mathbb{G}}(F) = \{F \circ \varphi^{-1} \mid \varphi \in \mathbb{G}\}$$

consists on all the images having the same topology[6] of F.
In consequence of this, the Similarity term will be the distance between the moving image and the fixed image in the same orbit:

$$\text{Sim}(F, M, \varphi) = \frac{1}{\sigma^2} ||F(\varphi^{-1}) - M||^2_{L^2}$$

To define the regularization term that provides the optimal φ at each step, in LDDMM (and subsequent frameworks) is considered the norm of the velocity vector field tangent to the transformation. Limiting is length means limiting the speed of the transformation at each step. We consider a generic time varying vector field (TVVF) as the continuously differentiable map

$$\begin{aligned} v : [0,1] &\longrightarrow \text{Vect}(\Omega) \\ t &\longmapsto v_t : \Omega \longrightarrow \mathbb{R}^d \\ &\qquad\quad \mathbf{x} \longmapsto v_t(\mathbf{x}) \end{aligned}$$

where $\text{Vect}(\Omega)$ is the set of all of the vector field over Ω. With this notation v is a vector field that changes continuously over a time parameter defined between 0 and 1. Once initial conditions are given, at each TVVF, corresponds a time varying homomorphisms defined by the following ODE

$$\frac{d\phi_t(\mathbf{x})}{dt} = v_t(\phi_t(\mathbf{x})) \tag{1.2}$$

where

$$\begin{aligned} \phi : [0,1] &\longrightarrow \text{Hom}(\Omega) \\ t &\longmapsto \phi_t : \Omega \longrightarrow \Omega \\ &\qquad\quad \mathbf{x} \longmapsto \phi_t(\mathbf{x}) \end{aligned}$$

[5] In this action it is preferable to consider the composition with the inverse, because this same action in differential geometry, called pull-back play the role of the contravariant operator of the push-forward, widely used to make a vector field act on a domain different from the one has been originally defined.

[6] The idea of having the same topology is an analytical consequence of the definition of diffeomoprhism. Separated domains remains separated, although if they gets close enough, let say that their distance is less than the size of a voxel for a significant region, then the discretization will not maintain analytical topology and could broke it anyway.

Now the transformation φ between fixed and moving images ($R \circ \varphi^{-1} = T$) in which we where originally interested, can be defined within the couple (v_t, ϕ_t). For $t = 0$, $\phi_t = Id$, identity of the group of homeomorphisms, and for $t = 1$, $\phi_t = \varphi$:

$$\varphi = \phi_1 = \phi_0 + \int_0^1 v_t(\phi) dt$$

The set $\{v_t(\phi) \mid t \in [0, 1]\}$ is a path of transformations that varies continuously over the parameter t, starting at the identity and ending in the one the registration framework is looking for.
To have an efficient algorithm and a meaningful constraint on the resulting transformation, it is reasonable to consider ϕ_t as the shortest path between the identity and φ, so to have v_t as the one that makes minimal the distance between transformations[7]:

$$l = \inf_{v_t \;:\; \dot{\phi}_t(\mathbf{x}) = v_t(\mathbf{x})} \int_0^1 ||v_t||^2_{L^2} dt$$

In the LDDMM, ending points of path on the set of diffeomorphisms, whose tangent vector field (that varies over time) and can be used as regularization term:

$$\mathrm{Reg}(F, M, \varphi) = \int_0^1 ||Lv_t||^2_{L^2} dt \qquad \dot{\phi}_t(\mathbf{x}) = v_t(\mathbf{x}) \quad \phi_0 = Id \quad \phi_1 = \varphi$$

Were φ in the registration framework is the one provided by the optimization algorithm at the previous step, and L is a linear operator that can be dependent on some parameters that makes the approach even more general[8]. The operator L is defined as $L = (\alpha\nabla + \gamma)$ for α and γ real parameters and ∇ the Laplace operator.
From the differential equation 1.2, and in consequence of the definition of φ the optimization function 1.1 is defined as:

$$\operatorname*{argmin}_{v_t \;:\; \dot{\phi}_t(\mathbf{x}) = v_t(\mathbf{x})} \mathcal{O}(F, M, \varphi) = \operatorname*{argmin}_{v_t \;:\; \dot{\phi}_t(\mathbf{x}) = v_t(\mathbf{x})} \int_0^1 ||Lv_t||^2_{L^2} dt + \frac{1}{\sigma^2} ||F(\varphi^{-1}) - M||^2_{L^2}$$

And so the optimizer, at each step of the registration will look for

$$\hat{v} = \operatorname*{argmin}_{v_t \;:\; \dot{\phi}_t(\mathbf{x}) = v_t(\mathbf{x})} \int_0^1 ||Lv_t||^2_{L^2} dt + \frac{1}{\sigma^2} ||F(\varphi^{-1}) - M||^2_{L^2}$$

With this definition we see that is the $TVVF$, instead of the transformation φ, to be the output of the registration algorithm. On the computational side this appears even more natural in software implementation, the action of a diffeomorphisms φ on an image is easily computed if the transformation is provided in term of discretized vector field v_t.
If G is a cubic grid defined as the discretization of the background space Ω, then a vector field dependent on time is parametrized as a 5-dimensional matrix

$$A = A(x, y, z, t, d) \qquad (x, y, z) \in G, \quad t \in [0, 1] \quad d = 1, 2, 3$$

[7] This next equation can provide a metric on the manifold of the transformations, making it a Riemannian manifold. On the other side starting with a metric previously defined on the manifold, the consequence existence of geodesics may avoid the computation of the inf. In both cases this "Riemannian approach" makes unavoidable the passage toward a metric, and makes the LDDMM a metric based algorithm.

[8] Recent approaches more image-oriented, proposed to use a kernel instead of an Operator.

where (x, y, z) are discrete position of the grid, t is the time parameter and d is index of the coordinate axis. So the tangent vector $\mathbf{v}_\tau(x_0, y_0, z_0)$ at the point of the grid (x_0, y_0, z_0), at time τ, has Cartesian coordinate

$$\mathbf{v}_\tau(x_0, y_0, z_0) = (A(x_0, y_0, z_0, \tau, 1), A(x_0, y_0, z_0, t, 2), A(x_0, y_0, z_0, \tau, 3))$$

In the LDDMM approach, each transformation, in particular input and output of the optimization algorithm are discretized time varying velocity fields; the update at each step is given by

$$\mathbf{v}^{k+1} = \mathbf{v}^k - \epsilon \nabla d\mathcal{O}$$

where $\mathbf{v}^k$ is the k-th step of the approximation, $d\mathcal{O}$ is the discretized version of the optimization function and ϵ is the gradient descent step size. Details of the algorithms are provided in [BMTY05]; for purpose of this thesis it is important to notice that diffeomorphisms are used for the underpinning theory, as the solution of the differential equation 1.2, and they are considered in the implementation while computing the similarity function; they are not used to compute the update at each step.

Shooting LDDMM

A direct upgrade of the

Euler poincare equation to compute the momentum equation.

Only geodesics with initial condition.

[VRRC12]

1.2.4 Stationary Velocity Fields LDDMM

The approach of using time varying velocity field (TVVF) is a consequence of the differential equation involved in the registration problem. Solutions to differential equation depends on time. Another approach, is to start from the manifold structure on which the solution lives, the integral domain. In this context higher terms of differential equations are element of the tangent bundle. These objects do not originates from any differential equation, and retain their dignity of vector field even if independent form time parameters. The first time the passage form TVVF as solution to differential equation, to SVF as elements living in the tangent bundle of the integral domain of differential equation was done in [ACPA06a]. Using SVF in image registration algorithms have many advantages: a greedy strategy is adopted, so that at each iteration an update only of the stationary velocity field is computed and added to the one of the previous step. This makes the algorithm faster and less memory consuming. On the other side we have two main drawback: the possibility to perform statistics on SVF become less straightforward and the set of SFV is a subset of the diffeomorphisms with no group structure. More details on this will be seen in section 3.2.1.

xxx parametrization of SVF

xxx the discretization

xxx the composition of SVF

xxx how statistics can be performed relying on the tangent spaces

In daemonology

Dartel, LDDMM shooting, do not requires any log composition. Log-composition is used in the 1) Log-demon - Tom 2) Bossa algorithm 3) Log euclidean by arsigny 4) Schild's Ladder - marco

In fact, given two TVVF u_t and v_t there is no straightforward operation that provides the TVVF that corresponds to the composition of the diffeomorphisms that generates these TVVF. To compute this, the vector field must be integrated to obtain the corresponding transformation ψ and ϕ. Corresponding transformation must be then composed and derived, to obtain again a TVVF that corresponds to the tangent vector field of the composition of u_t and v_t.

1.3 Stationary Velocity Fields and the Composition of Diffeomorphisms in the Tangent Space

xxx computational anatomy and the need for statistics.
xxx introduction of log and exp concepts (formally later, in section 2.1.3).
xxx starting from vectors, ending with one vector: the log composition, idea, not formally defined.
defined as the vector $\mathbf{w}$ in the Lie algebra $\mathfrak{g}$ that reflects the composition in the Lie group of two vectors $\mathbf{u}, \mathbf{v}$ in the same tangent space:

$$\mathbf{w} = \log(\exp(\mathbf{u}) \circ \exp(\mathbf{v})) \qquad \forall \mathbf{u}, \mathbf{v} \in \mathfrak{g}$$

xxx the BCH formula, as the first method to compute the log-composition.

Chapter 2

Tools from Differential Geometry

2.1 A Lie Group Structure for the Set of Transformation

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

We consider every group $\mathbb{G}$ as a group of transformation acting on $\mathbb{R}^d$, having in mind the particular case $d = 2, 3$ for 2-dimensional or 3-dimensional images. We will focus out 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 transformation 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 in general 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 gives us new handles to hold 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. The abstract idea of vector field over a manifold will be concretized for image registration introducing the concepts of *displacement field*, *deformation field* and *velocity field (stationary or time varying)* that will be there presented. Avoid pedantry is as important as to avoid confusions on notations and definitions, therefore it is necessary to call back a few concepts from differential geometry tailored for rigid-body and diffeomorphic image registration, before getting into the heat of the applications.

2.1.1 Velocity Vector Fields and Flows

Let $\gamma(t)$ be a (continuous) path over a Lie group $\mathbb{G}$, such that $t \in (-\eta, \eta) \subseteq \mathbb{R}$ and $\gamma(0) = p$. If (C, ψ) is a local chart, neighborhood of p, the tangent vector of γ at the point p can be expressed as

$$\mathbf{u} = \frac{d}{dt}(\psi \circ \gamma)(t)\Big|_{t=0}$$

For different choice of γ passing through p, we obtain different tangent vectors.

$$\begin{tikzcd}
 & C \subseteq \mathbb{G} \arrow[dr, "\psi"] & \\
\mathbb{R} \arrow[ur, "\gamma"] \arrow[rr, "\psi \circ \gamma"] & & \mathbb{R}^n
\end{tikzcd}$$

It can be proved that the set of all of the tangent vector at the point p defines a vector space: the tangent space at p, indicated with $T_p\mathbb{G}$. It can be proved that this construction do not depend on the local chart's choice.
Taking into account the disjoint union of all of the tangent spaces of $\mathbb{G}$ we obtain the tangent bundle $T\mathbb{G}$; it can be proven that it is, in its turn, a differentiable manifold.
Be $\mathbb{G}$ n-dimensional Lie group. A *vector field* over $\mathbb{G}$ is a function that assigns at each point p of $\mathbb{G}$, a tangent vector V_p in the tangent space $T_p\mathbb{G}$, such that V_p is differentiable respect to p.
If $(C; x_1, \dots, x_n) = (C, \psi)$ is a local chart of $\mathbb{G}$, neighborhood of p, then V_p can be expressed locally as:

$$V_p = \sum_{i=1}^{n} v_i(p) \frac{\partial}{\partial x_i}\Big|_p \qquad\qquad v_i \in \mathcal{C}^{\infty}(C)$$

Using the Einstein summation convention V_p is sometime expressed as $V_p = v^i(p)\partial_i\big|_p$. The smooth functions v_i define the vector fields in the base $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$. The idea of expressing the elements of the base in terms of differential operator reveals the possibility to consider each vector field as a directional derivative over the algebra of smooth functions defined on the manifold.
The set of all vector field over M, indicated with $\mathcal{V}(M)$, is a real vector space and a module over $\mathcal{C}^{\infty}(M)$:

$$\begin{aligned}
&(V+W)_p = V_p + W_p && \forall V, W \in \mathcal{V}(M)\\
&(aV)_p = aV_p && \forall a \in \mathbb{R}\\
&(fV)_p = f(p)V_p && \forall V \in \mathcal{V}(M) \quad \forall f \in \mathcal{C}^{\infty}(M)
\end{aligned}$$

Moreover $\mathcal{V}(M)$ acts over $\mathcal{C}^{\infty}(M)$ as follows

$$\begin{aligned}
\mathcal{V}(M) \times \mathcal{C}^{\infty}(M) &\longrightarrow \mathcal{C}^{\infty}(M)\\
(V, f) &\longmapsto Vf : \mathcal{C}^{\infty}(M) \longrightarrow \mathbb{R}\\
&\qquad\qquad p \longmapsto (Vf)(p) = V_p f
\end{aligned}$$

In the local chart the real number $V_p f$ is given by

$$(Vf)(p) = V_p f = \sum_{j=1}^{n} v_j(p) \frac{\partial f}{\partial x_j}\Big|_p$$

and represents the directional derivative of f along the vector $V_p \in T_pM$.
If $(C_2; y_1, \ldots, y_n)$ is another local chart, $p \in C_2$, then the change of coordinates can be expressed as follows:

$$V_p = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} v_i(p) \frac{\partial y_j}{\partial x_i} \Big|_p \Big) \frac{\partial}{\partial y_j} \Big|_p$$

A vector field can be time-dependent if each of its vectors varies smoothly within a parameter t, otherwise it is time-independent. In this case a continuous function over the set of times T is defined:

$$\begin{aligned} \mathbb{R} \supseteq T &\longrightarrow \mathcal{V}(\mathbb{G}) \\ t &\longmapsto V^{(t)} : \mathbb{G} \longrightarrow T\mathbb{G} \\ &\qquad\qquad p \longmapsto V_p^{(t)} \end{aligned}$$

where $V_p^{(t)}$ has local coordinates

$$V_p^{(t)} = \sum_{i=1}^{n} v_i(p,t) \frac{\partial}{\partial x_i} \Big|_p \qquad v_i \in \mathcal{C}^\infty(C \times T)$$

Let V a vector field over a differentiable manifold $\mathbb{G}$, an *integral curve* of V is given by

$$c : (a,b) \longrightarrow \mathbb{G} \quad \text{such that} \quad \dot{c}(t) = V_{c(t)} \in T_{c(t)}\mathbb{G} \ \forall t \in (a,b)$$

To get the equations of the integral curves, we consider the local expression

$$V = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i} \qquad\qquad v_i \in \mathcal{C}^\infty(C)$$

and the unknown curve in the same local chart

$$c(t) = (c_1, c_2, \ldots, c_n) \qquad \dot{c}(t) = \sum_{i=1}^{n} \frac{dc_i(t)}{dt} \frac{\partial}{\partial x_i} \Big|_{c(t)}$$

Imposing the condition $\dot{c}(t) = V_{c(t)}$ we get:

$$\sum_{i=1}^{n} \frac{dc_i(t)}{dt} \frac{\partial}{\partial x_i} \Big|_{c(t)} = \sum_{i=1}^{n} v_i(c(t)) \frac{\partial}{\partial x_i} \Big|_{c(t)}$$

For a given point of the manifold, and considering the integral curves passing for this point we obtain the initial condition $c(0) = p$ for a Cauchy problem :

$$\begin{cases} \frac{dc_i(t)}{dt} = v_i(c_1, t_2, \ldots, c_n) \\ c_i(0) = p_i \end{cases} \tag{2.1}$$

Thanks to the Cauchy theorem it has a unique solution $\gamma(t)$. The unique integral curve passing through p when $t = 0$ is noted by $c^{(p)}(t)$.
Integral curves can be divided in 2 classes: the one whose domain can be extended to the whole real line $\mathbb{R}$ (in this case V is called *completely integrable vector field*) and the one

whose domain is a strict subset of $\mathbb{R}$. We reminds that the *flow* of the vector field V is the defined as:

$$\begin{aligned}\Phi_V : S \times \mathbb{G} &\longrightarrow \mathbb{G}\\ (t,p) &\longmapsto \Phi_V(t,p) = c^{(p)}(t)\end{aligned}$$

where $S = \mathbb{R}$ or $S \subset \mathbb{R}$ if V is or is not respectively completely integrable. Fixing the point p, the flow become simply the integral curve passing through p; keep t fixed and letting p varying over the manifold, we get the position of each point on the manifold subject to the vector field V at the time t. This last idea gives raise to the *one-parameter subgroup*:

$$\forall p \in M \qquad \Phi_V(t,p) = \varphi_t \qquad G = \{\varphi_t \ : \ t \in S\}$$

$$\begin{aligned}G \times G &\longrightarrow G\\ (\varphi_{t_1}, \varphi_{t_2}) &\longmapsto \varphi_{t_1+t_2}\end{aligned}$$

Despite the name, the fact that G forms a group is less important[1] than considering the compatibility between a sum on the real line and a product between functions. This property will be largely used when dealing with Lie logarithms. In general a continuous function

$$f : \mathbb{R} \supseteq (-\eta, \eta) \longrightarrow \mathbb{G} \qquad f(0) = p$$

satisfies the *one parameter subgroup property* if $f(t+s) = f(t)f(s)$ where the last multiplication is the composition on the group.

2.1.2 Push-forward, Left, Right and Adjoint Translation

Given two Lie group $\mathbb{G}$ and $\mathbb{H}$ linked by the differentiable map $F : \mathbb{G} \to \mathbb{H}$, then the *push forward* at the point p is defined as the covariant operator

$$\begin{aligned}(F_\star)_p : T_p\mathbb{G} &\longrightarrow T_{F(p)}\mathbb{H}\\ V_p &\longmapsto (F_\star V_p) : \mathcal{C}^\infty(\mathbb{H}) \longrightarrow \mathbb{R}\\ &\qquad f \longmapsto (F_\star V_p)(f) = V_p(f \circ F) = v(p)^i \partial_i (f \circ F)\big|_p\end{aligned}$$

When the point p is implicit by the context it will be omitted: namely $(F_\star)_p = F_\star$.

In general the push forward gives the right to the vector field V defined over $\mathbb{G}$ to act as a derivative on another manifold $\mathbb{H}$. Push forward is well defined since a vector field is completely determined by its action over $\mathcal{C}^\infty(\mathbb{H})$. It can be proved that it is linear, satisfies the Leibnitz rules, and $(G \circ F)_\star = G_\star \circ F_\star$; moreover, the push forward of the identity is the identity map between vector spaces, and if F is a diffeomorphism, $F_\star$ is an isomorphism of vector spaces.
The *pull-back*, is defined on the dual space of $\mathbb{G}$ and $\mathbb{H}$ as the contravariant operator of the push forward[2]:

$$\begin{aligned}F^\star : \mathcal{C}^\infty(\mathbb{H}) &\longrightarrow \mathcal{C}^\infty(\mathbb{G})\\ f &\longmapsto F^\star f := f \circ F\end{aligned}$$

[1] In this context: from a group theory point of view the action of the group $(\mathbb{R}, +)$ over the manifold has as its orbits, the set of disjoints integral curves.

[2] Push-forward is defined between vector spaces, pull-back between space of functions and $V_p(F^\star f) = (v(p)^i \partial_i\big|_p)(f \circ F) = v(p)^i \partial_i (f \circ F)\big|_p = V_p(f \circ F)$.

The following diagram relates pull-back and push-forward:

$$
\begin{tikzcd}
T_p\mathbb{G} \arrow[r, "F_\star"] \arrow[d, dashed] & T_p\mathbb{H} \arrow[d, dashed] & \\
\mathbb{G} \arrow[r, "F"] \arrow[rr, bend right, "F^\star f"'] & \mathbb{H} \arrow[r, "f"] & \mathbb{R}
\end{tikzcd}
$$

Here we restrict our attention at the case $\mathbb{G} = \mathbb{H}$, thus F is a map that moves the points of $\mathbb{G}$ smoothly. Each element p of a Lie group $\mathbb{G}$ defines three maps on $\mathbb{G}$:

1. *left-translation*:
$$\begin{aligned} L_p : \mathbb{G} &\longrightarrow \mathbb{G} \\ q &\longmapsto pq \end{aligned}$$

2. *right-translation*:
$$\begin{aligned} R_p : \mathbb{G} &\longrightarrow \mathbb{G} \\ q &\longmapsto qp \end{aligned}$$

3. *adjoint map*
$$\begin{aligned} \mathrm{Ad}_p : \mathbb{G} &\longrightarrow \mathbb{G} \\ q &\longmapsto pqp^{-1} \end{aligned}$$

The push forward for the vector field V at the point q are given by:

1. *left-translation*:
$$(L_p)_\star V_q f = V_q(f \circ L_p) = \sum_{i=1}^{n} v_i(q) \frac{\partial f \circ L_p}{\partial x_i} \Big|_q = \sum_{i=1}^{n} v_i(q) \frac{\partial f}{\partial x_i} \Big|_{pq}$$

2. *right-translation*:
$$(R_p)_\star V_q f = V_q(f \circ R_p) = \sum_{i=1}^{n} v_i(q) \frac{\partial f \circ R_p}{\partial x_i} \Big|_q = \sum_{i=1}^{n} v_i(q) \frac{\partial f}{\partial x_i} \Big|_{qp}$$

3. *adjoint map*
$$(\mathrm{Ad}_p)_\star V_q f = V_q(f \circ \mathrm{Ad}_p) = \sum_{i=1}^{n} v_i(q) \frac{\partial f \circ \mathrm{Ad}_p}{\partial x_i} \Big|_q = \sum_{i=1}^{n} v_i(q) \frac{\partial f}{\partial x_i} \Big|_{pqp^{-1}}$$

We note that in each expression the coefficient $v_i(q)$ remains the same even if the partial derivative is not applied at the point q. Therefore the linear combination of the constant coefficients $v_i(q)$ can be considered as a scalar product with the elements of the base applied at the function f. Left and right translation of the vector $\mathbf{u}$ can be expressed as scalar product with the *differential*, equivalent concept as the push forward, that emphasizes the scalar product implied in the definition:

$$\begin{aligned}(DL_p)_q : T_q\mathbb{G} &\longrightarrow T_{pq}\mathbb{G}\\ \mathbf{u} &\longmapsto (DL_p)_q \cdot \mathbf{u}\end{aligned}$$

$$\begin{aligned}(DR_p)_q : T_q\mathbb{G} &\longrightarrow T_{qp}\mathbb{G}\\ \mathbf{u} &\longmapsto (DR_p)_q \cdot \mathbf{u}\end{aligned}$$

where $(DL_p)_q$, $(DR_p)_q$ are properly defined vectors that can be expressed local coordinates as follow

$$(DL_p)_q = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Big|_{pq} \qquad (DR_p)_q = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Big|_{qp}$$

Or equivalently linear operators defined as:

$$\begin{aligned}(DL_p)_q : \mathcal{C}^\infty(M) &\longrightarrow \mathbb{R}\\ f &\longmapsto (DL_p)_q(f) = \frac{\partial f}{\partial x_i}\Big|_{pq}\end{aligned}$$

$$\begin{aligned}(DR_p)_q : \mathcal{C}^\infty(M) &\longrightarrow \mathbb{R}\\ f &\longmapsto (DR_p)_q(f) = \frac{\partial f}{\partial x_i}\Big|_{qp}\end{aligned}$$

A change of notation $V_q = \mathbf{u}$ makes push-forward and differential strikingly equivalent. This holds also for the generic map F:

$$(DF)_q(f) = \sum_{i=1}^{n} \frac{\partial f \circ F}{\partial x_i}\Big|_q \qquad (DF)_q(f)\cdot\mathbf{u} = \sum_{i=1}^{n} u_i \frac{\partial f \circ F}{\partial x_i}\Big|_q$$

The subscript q in $(DL_p)_q$ can be omitted when the tangent space of $\mathbf{u}$ is clear by the context.

A vector field V defined over a manifold is *left-invariant* if it is invariant for each left translation. It means that $(L_q)_\star V_p = V_p$ for any choice of p and q. If we consider all of the possible push forward of the left translation applied to a single tangent vector at the origin $\mathbf{v}$ of $T_e\mathbb{M}$ we have a unique left-invariant vector field defined as $\mathbf{v}^L$ such that

$$\mathbf{v}^L_q := (L_q)_\star\mathbf{v} \qquad \forall q \in M$$

Vice versa every left-invariant vector field V is uniquely represented by V_e. The set of all of the left-invariant vector fields form a linear subspace of the space of the vector field, indicated with left$\mathcal{V}(M)$. This can be easily proved by:

$$(L_g)_\star(aV + bW) = a(L_g)_\star V + b(L_g)_\star W \qquad \forall V, W \in \mathcal{V}(\mathbb{G}) \quad \forall a, b \in \mathbb{R}$$

In fact for each $h \in \mathbb{G}$ and for each $f \in \mathcal{C}^\infty(\mathbb{G})$ the linearity property holds:

$$\begin{aligned}(L_g)_\star(aV_h + bW_h)f &= (aV_h + bW_h)(f \circ L_g)\\ &= aV_h(f \circ L_g) + bW_h(f \circ L_g)\\ &= a(L_g)_\star V_h f + b(L_g)_\star W_h f\end{aligned}$$

The linearity property leads to the definition of the group of homomorphism over $\mathbb{G}$. It is the set of all the Lie group homomorphism from $\mathbb{R}$ to $\mathbb{G}$:

$$Hom(\mathbb{R}, \mathbb{G}) = \{\varphi : \mathbb{R} \to \mathbb{G} \mid \varphi(a+b) = \varphi(a) \circ \varphi(b) \quad \forall a, b \in \mathbb{R}\}$$

Tangent spaces, flows, one-parameter subgroup and Lie group homomorphisms are bounded together by the following remarkable result, which is a most important precondition for the definition of the Lie group exponential, and so deserve to be written in form of a lemma and formally proved.

Lemma 2.1.1. Let $\mathbb{G}$ be a Lie group. For each $\mathbf{v}$ in the tangent space $T_e\mathbb{G}$, exists a unique homomorphism $\gamma_{\mathbf{v}}$ in $Hom(\mathbb{R}, \mathbb{G})$ (or equivalently a function satisfying the one-parameter subgroup property) such that

$$\dot{\gamma}_{\mathbf{v}}(0) = \mathbf{v}$$

Proof. The homomorphism $\gamma_{\mathbf{v}}$ coincides with the integral curve Φ of the left invariant vector field generated by $\mathbf{v}$ passing through the identity. Its uniqueness is then a consequence of the Cauchy theorem. The same theorem also specifies the existence for a small enough neighbour $(-\eta, \eta) \subset \mathbb{R}$. To extend the solution to the whole $\mathbb{R}$ it is enough to consider that $\gamma_{\mathbf{v}}(t+s) = \gamma_{\mathbf{v}}(t)\gamma_{\mathbf{v}}(s)$ for each $s, t \in (-\eta, \eta)$:

$$\gamma_{\mathbf{v}}(t+s) = \Phi(t+s, e) = \Phi(t, \gamma_{\mathbf{v}}(s)) = \gamma_{\mathbf{v}}(t)\gamma_{\mathbf{v}}(s)$$

□

We observe that $\gamma_{\mathbf{v}}$ is exactly the one parameter subgroup of $\mathbf{v}^L$ defined above, and then we can write $\gamma_{\mathbf{v}}(t) = \Phi(t, e) = \varphi_e(t)$.

We conclude this paragraph remembering the definition of the Lie algebra of a Lie group that take into account every feature so far introduced:

Definition 2.1.1. Given a Lie group $\mathbb{G}$, its Lie algebra $\mathfrak{g}$ is defined as:

1. The vector space $T_e\mathbb{G}$ of all of the tangent vector at the identity (or at any other point of the manifold): $\mathfrak{g} := T_e\mathbb{G}$.
2. The set of the left invariant vector Field over $\mathbb{G}$: $\mathfrak{g} := \text{left}\mathcal{V}(\mathbb{G})$.
3. The set of all of the flows passing through e: $\mathfrak{g} := \{\Phi(e, t) : t \in S \subseteq \mathbb{R}\}$.
4. The set of homomorphism $Hom(\mathbb{R}, \mathbb{G})$.

The Lie algebra can be also defined independently from a Lie group as a vector space endowed with Lie bracket (bilinear form, antisymmetric, that satisfies the Jacobi identity). In the finite dimensional case given a Lie algebra $\mathfrak{g}$ it can be proved that exists always a Lie group $\mathbb{G}$ such that $\mathfrak{g}$ is the Lie algebra defined over $\mathbb{G}$. This property (third Lie theorem) do not holds anymore infinite dimensional Lie algebra of diffeomorphisms.

2.1.3 Lie Exponential, logarithm and Log-composition

Let $\mathbf{v}$ be an element in the tangent space $\mathfrak{g}$ and $V \in \text{left}\mathcal{V}(\mathbb{G})$ the unique vector field defined by $\mathbf{v}$ over a local coordinate system around the origin. Let Φ_V be the flow associated with the vector field and $\gamma(t)$ the unique integral curve of V passing through the identity of the group. The *Lie exponential* is defined as

$$\begin{aligned} \exp : \mathfrak{g} &\longrightarrow \mathbb{G} \\ \mathbf{v} &\longmapsto \exp(\mathbf{v}) = \gamma(1) \quad \dot{\gamma}(t) = V_{\gamma(t)}, \gamma(0) = e \end{aligned}$$

It satisfies the following properties:

1. $\exp(\mathbf{v}) = \Phi_V(e, 1)$.
2. $\exp(t\mathbf{v}) = \gamma(t) = \Phi_V(e, t)$.
3. $\exp(\mathbf{v}) = e$ if $\mathbf{v} = \mathbf{0}$.
4. $\exp(\mathbf{v}) \circ \exp(-\mathbf{v}) = e$
5. The exponential function satisfies the one parameter subgroup property:
$$\exp((t+s)\mathbf{v}) = \gamma(t+s) = \gamma(t) \circ \gamma(s) = \exp(t\mathbf{v}) \exp(s\mathbf{v})$$
6. $\exp(\mathbf{v})$ is invertible and $(\exp(\mathbf{v}))^{-1} = \exp(-\mathbf{v})$.
7. exp is a diffeomorphism between a neighborhood of $\mathbf{0}$ in $\mathfrak{g}$ to a neighborhood of Id in $\mathbb{G}$.

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[3].
When we deal with a matrix Lie group of dimension n, we have the following remarkable property:

1. for all $\mathbf{v}$ in a matrix Lie algebra $\mathfrak{g}$:
$$\exp(\mathbf{v}) = \sum_{k=0}^{\infty} \frac{\mathbf{v}^k}{k!}$$
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, $||\exp(\mathbf{v})|| \leq \exp(||\mathbf{v}||)$.
6. $\exp(\mathbf{u}+\mathbf{v}) = \lim_{m\to\infty} (\exp(\frac{\mathbf{v}}{m}) \exp(\frac{\mathbf{v}}{m}))^m$
7. If $\exp(\mathbf{w}) = \exp(\mathbf{u}) \circ \exp(\mathbf{v})$ then $\exp(-\mathbf{w}) = \exp(-\mathbf{v}) \circ \exp(-\mathbf{u})$.
8. For Ad adjoint map[4] we have $\exp(\mathrm{Ad}_{\mathbf{u}}\mathbf{v}) = \mathrm{Ad}_{\mathbf{u}} \exp(\mathbf{v})$

[3] Here we define cut locus starting from the exp and log function, and in both domains. Traditionally it is defined only on Riemannian manifolds and using the geodesics (see [dCV92], p. 267). For Levi Civita connection we have that the definition are coincident.

[4] Lie group action defined as as

$$\begin{aligned} \mathrm{Ad} : \mathbb{G} &\longrightarrow Aut(\mathbb{G}) \\ \mathbf{u} &\longmapsto \mathrm{Ad}_{\mathbf{u}} : \mathbb{G} \longrightarrow \mathbb{G} \\ &\qquad\qquad \mathbf{v} \longmapsto \mathrm{Ad}_{\mathbf{u}}\mathbf{v} = \mathbf{u}\mathbf{v}\mathbf{u}^{-1} \end{aligned}$$

that preserves the lie bracket in the Lie algebra: $\mathrm{Ad}_{\mathbf{u}}[\mathbf{v}, \mathbf{w}] = [\mathrm{Ad}_{\mathbf{u}}\mathbf{v}, \mathrm{Ad}_{\mathbf{u}}\mathbf{w}]$.

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

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

where $\mathbf{v}$ is the tangent vector having p as it exp.
The idea of the names seems to be justified by the following example:

Example 2.1.1. If we take the unitary circle in the complex plane $\mathbb{S}^1$, and the vertical line $x = 1$ as its tangent at the point $(0,0)$. Each element θ of the group of rotation of the plane corresponds a point on the circle $\cos(\alpha) + i\sin(\alpha)$, and so the group of rotation can be identified with the circle. Thanks to the Euler's formula we can write

$$\mathbb{S}^1 = \{e^{i\alpha} \mid \alpha \in (-\pi, \pi]\}$$

The Lie algebra of this group of rotations is the tangent line to the circle at the neutral

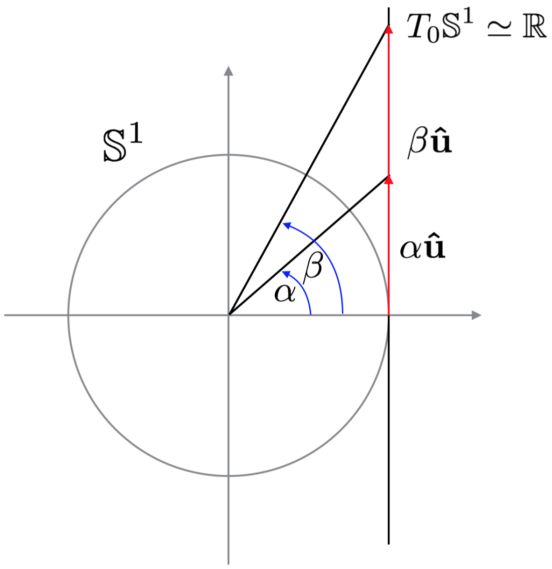

Figure 2.1: Lie algebra of the Lie group of plane's rotation.

element $\alpha = 0$, and it is isomorphic to $\mathbb{R}$. Lie logarithm and Lie exponential for this particular case corresponds exactly with the usual logarithm and exponential:

$$\begin{aligned} \log : \mathbb{S}^1 &\longrightarrow T_0\mathbb{S}^1 \\ \exp(i\alpha) &\longmapsto \log(\exp(i\alpha)) = i\alpha \end{aligned} \qquad \begin{aligned} \exp : T_0\mathbb{S}^1 &\longrightarrow \mathbb{S}^1 \\ i\alpha &\longmapsto \exp(i\alpha) \end{aligned}$$

The internal cut locus of the lie group is $(-\pi, \pi)$.

If $\mathbb{G}$ is a matrix Lie group of dimension n, the following properties hold:

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

$$\log(\mathbf{v}) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\mathbf{v} - I)^k}{k!}$$

   where I is the identity matrix.

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

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

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:

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

Here we may see the beginning of the problem we have to deal with for the rest of the research, when passing from the finite dimensional case to the infinite dimensional case.
The domain of the logarithm is the matrix Lie group in which only the composition is defined. Nevertheless it is possible to compute $I + \mathbf{c}$, and this still make sense (and satisfy remarkable properties) when applied to the log. On the other side the domain of the exponential is the matrix Lie algebra, but the exponential can be nevertheless applied to a generic matrix.
This can be done thanks to the fact that for matrices, $\mathfrak{g}$ and $\mathbb{G}$ are subset of a bigger algebra, the algebra of invertible matrix: in this context operation of sum is still defined over the group that admit only compositions. The sum between element of a group can be performed on a Lie group, every time he and its Lie algebra are subset of a bigger algebra (Kirillov).
In these cases infinite series are doors to passes from the structure of group and the algebra.
When presenting the rigid body transformation in chapter 3.1 we will see a second couple of access doors based on numerical approximations.

2.1.4 Definition of Lie Log-Composition

We define the Lie Log-composition (Lie to distinguish it from the Affine Log-composition of the next chapter) as inner binary operation on the Lie algebra that reflects the composition on the lie group:

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

Following properties holds for the Lie log-composition:

1. $\mathfrak{g}$ with the Lie log-composition $\star$ 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) $(\mathbf{u}_1 \star \mathbf{u}_2) \star \mathbf{u}_3 = \mathbf{u}_1 \star (\mathbf{u}_2 \star \mathbf{u}_3)$ for all $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ in $C_{\mathfrak{g}}$.
   (b) $\mathbf{u} \star \mathbf{0} = \mathbf{0} \star \mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.
   (c) $\mathbf{u} \star (-\mathbf{u}) = \mathbf{0}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.
2. For all t, s real, such that $(t + s)\mathbf{v}$ is in $C_{\mathfrak{g}}$,

$$(t\mathbf{v}) \star (s\mathbf{v}) = (t + s)\mathbf{v}$$

   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 $\mathfrak{g}$.
3. For all $\mathbf{u}$ and $\mathbf{v}$:

   xxx property involving metric and log-composition... may results in something interesting for computing statistics.

$$||\mathbf{u} \star \mathbf{v}|| = ?(||\mathbf{u}||, ||\mathbf{v}||)$$

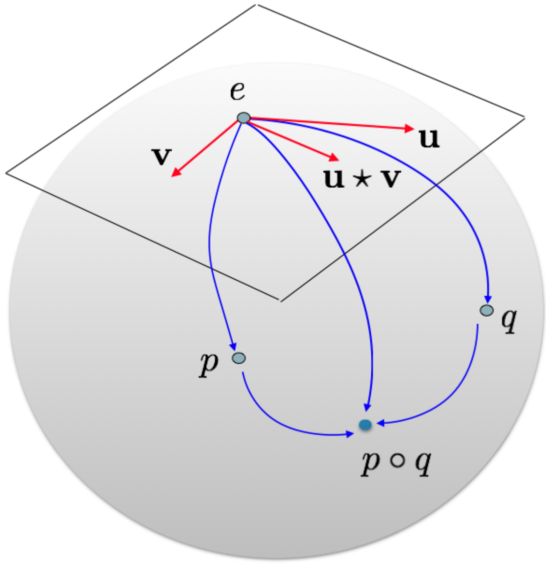

Figure 2.2: graphical visualization of the Lie log-composition.

2.1.5 BCH formula for the Computation of Log-composition

To compute the Lie Log composition, literature provides the BCH formula, defined as the solution of the equation $exp(\mathbf{w}) = \exp(\mathbf{u}) \circ \exp(\mathbf{v})$, for $\mathbf{u}$ and $\mathbf{v}$ in the Lie algebra $\mathfrak{g}$:

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

xxx derivation of the bch formula, constraints on the Lie algebra elements involved in its computation.

2.1.6 Taylor Expansion for the Computation of Log-composition

Once *adjoint action* of $\mathbf{u}$ on the Lia algebra is defined, nested Lie bracket can be reformulated as multiple composition of this operator:

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

So

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

In the appendix of xxx Klarsfeld xxx adjoint action are used to provide an expansion of the BCH formula. This can be rewritten as

$$\mathbf{u} \star \mathbf{v} = \mathbf{u} + \frac{ad_{\mathbf{u}} \exp(ad_{\mathbf{u}})}{\exp(ad_{\mathbf{u}}) - 1}\mathbf{v} + O(\mathbf{v}^2)$$

xxx intermediate passages to be written from zachos, blane!

The functional applied to $\mathbf{v}$ can be rewritten as

$$\frac{ad_{\mathbf{u}} \exp(ad_{\mathbf{u}})}{\exp(ad_{\mathbf{u}}) - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} ad_{\mathbf{u}}^n$$

Where $\{B_n\}$ is the sequence of the second-kind Bernoulli number[5].

2.2 Parallel Transport Tool

In this section we present parallel transport for the finite dimensional Lie group and we make the assumption that obtained results hold in the infinite dimensional case.

2.2.1 Connections and Geodesics

Given a Lie Group $\mathbb{G}$, a connection ∇ is an operator which assign to each vector field U in $\mathcal{V}(\mathbb{G})$ the map

$$\begin{aligned} \nabla_U : \mathcal{V}(\mathbb{G}) &\longrightarrow \mathcal{V}(\mathbb{G}) \\ V &\longmapsto \nabla_U V \end{aligned}$$

such that for all f, g in $\mathcal{C}^\infty(\mathbb{G})$ and for all V, W in $\mathcal{V}(\mathbb{G})$ the following conditions are satisfied:

1. $\nabla_{fU+gV} = f\nabla_U + g\nabla_V$
2. $\nabla_U(fV) = f\nabla_U(V) + (Uf)V$

where in the second condition we have used the structure of $\mathcal{C}^\infty$-module of $\mathbb{G}$ and the fact that

$$\begin{aligned} Uf : \mathbb{G} &\longrightarrow \mathbb{R} \\ p &\longmapsto U_p f \end{aligned}$$

Geometrically the vector field $\nabla_U(V)$ associates at each point of the manifold the projection on the tangent plane of the covariant derivative of U in the direction of V. The definition seems cryptic but the connection appears to be that general tool that provides geodesics and curvature over manifold on which no Riemannian metric has been defined [dCV92]. In fact on a Lie group $\mathbb{G}$ a *geodesic* between two of its points p and q can be defined as the curve γ such that:

$$\gamma : [0,1] \longrightarrow \mathbb{G} \qquad \gamma(0) = p,\ \gamma(1) = p,\ \nabla_{\dot\gamma}\dot\gamma = 0$$

Note that in this case the concept of geodesic did not involves any metric defined on the surface of the manifold. If also a Riemannian metric is defined on $\mathbb{G}$, then geodesics defined by the metric coincides with the geodesics defined by the connection only for the particular Levi-Civita connection. A connection is said to be left invariant if it is closed for left invariant vector fields, i.e. if for any $V, W \in Left\mathcal{V}(\mathbb{G})$ their connection $\nabla_U V$ is still left invariant.

[5]If first-kind Bernoulli number is used then each term of the summation must be multiplied for $(-1)^n$, as did for example in ....Klarsfeld.

2.2.2 Affine Exponential, Logarithm and Log-Composition

If $\mathbb{G}$ is endowed with a connection ∇, then a new kind of exponential from the Lie algebra to the Lie group can be defined, using geodesics. This time the tangent plane that defines the Lie algebra is considered at the generic point p of the Lie group and $\mathbf{v} \in T_p\mathbb{G} \simeq \mathfrak{g}$ is a tangent vector at the point p:

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

The curve $\gamma(t; p, \mathbf{v}) = \gamma(t)$ on on $\mathbb{G}$ is the unique one with the following features:

$$\gamma(0) = p \qquad \dot{\gamma}(0) = \mathbf{v} \qquad \nabla_{\dot{\gamma}}\dot{\gamma} = 0$$

To distinguish the affine exp and log from the Lie exp and log presented in the previous chapter, the affine will always have the subscript of the point of application even when it is the identity.
The following properties hold:

1. If ∇ is a Cartan connection then $\exp_e$ and exp coincides.
2. For all p in $\mathbb{G}$, $\mathbf{v} \in T_p\mathbb{G}$ and t real
$$\exp_p(t\mathbf{v}) = \gamma(t; p, \mathbf{v})$$
3. Given $\mathbf{u} \in T_e\mathbb{G}$, $\mathbf{v} \in T_{\exp_e(\mathbf{u})}\mathbb{G}$, exists a $\mathbf{w} \in T_e\mathbb{G}$ such that
$$\exp_e(\mathbf{w}) = \exp_{\exp_e(\mathbf{u})}(\mathbf{v}) \circ \exp_e(\mathbf{u})$$
4. If V is a unitary left-invariant vector field, then for $V_e \in T_e\mathbb{G}$
$$\exp_e(2V_e) = \exp_{\exp_e(V_e)}(V_{\exp_e(V_e)}) \circ \exp_e(V_e)$$

Last two properties provides 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 possess 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 Lie algebra structure) may not be oriented in the same way.
xxx parallel transport example on the sphere.
To approach the inverse of the affine exponential we consider the affine logarithm:

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

Where $\mathbf{v}$ is the vector at the tangent plane defined at p such that the curve on $\mathbb{G}$ with the following features

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

has as its tangent in p the vector $\mathbf{v}$.
xxx properties of affine log and exp xxx Any Lie group $\mathbb{G}$ considered with a left-invariant connection ∇ can be equipped with a metric, based on the elements of its tangent space and on the log, and not necessarily coincident with the Riemannian one:

$$\text{dist}(x, y) := ||\log_e(x^{-1} \circ y)|| \qquad \forall x, y \in \mathbb{G}$$

Definition of Affine Log-Composition

We need to extend the definition of Lie log-composition to the Affine Log-composition. The first step is to extend the definition of internal cut locus of the Lie algebra, even when not centered at the zero. If the Lie algebra, considered as tangent space, is not considered at e of $\mathbb{G}$ but at the point p instead, we still have a diffeomorphism between a neighborhood of $\mathbf{0}$ in $\mathfrak{g}$ to a neighborhood of p in $\mathbb{G}$. The internal cut locus of $\mathfrak{g}$ this time is based on p and it is denoted with $C_{\mathfrak{g}}(p)$.

Given a point p_1 and a vector $\mathbf{v}_1$ on its tangent plane $T_{p_1}\mathbb{G}$ the *affine Log-composition* is defined as the operator operation $\tilde{\star}$ over the $\mathbb{G}$ fiber bundle such that

$$\begin{aligned} \cdot\ \tilde{\star}\ \mathbf{v}_1 : T_{\exp_{p_1}(\mathbf{v}_1)}\mathbb{G} &\longrightarrow T_{p_1}\mathbb{G} \\ \mathbf{v}_2 &\longmapsto \mathbf{v}_2\ \tilde{\star}\ \mathbf{v}_1 = \log_{p_1}(\exp_{\exp_{p_1}(\mathbf{v}_1)}(\mathbf{v}_2) \circ \exp_{p_1}(\mathbf{v}_1)) \end{aligned}$$

Note that not necessarily $\mathbf{v}_1\ \tilde{\star}\ \mathbf{v}_2$ is a vector belonging to the internal cut locus based on the starting point p_1.

2.2.3 Parallel Transport through Examples

Parallel transport will play a role in the computation of both Lie and Affine log-composition.

Definition 2.2.1. Let $\mathbb{G}$ be a finite dimensional connected Lie group defined with a connection ∇. Given $p, q \in \mathbb{G}$ and $\gamma : [0,1] \to \mathbb{G}$ such that $\gamma(0) = p$ and $\gamma(1) = q$, then the vector $V_p \in T_p\mathbb{G}$, belonging to some vector field V is parallel transported along γ up to $T_q\mathbb{G}$ if for all $t \in [0,1]$ $\nabla_{\dot{\gamma}} V_{\gamma(t)} = 0$.
The parallel transport is the function that maps V_p from $T_p\mathbb{G}$ to $T_q\mathbb{G}$ along γ:

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

xxx examples of parallel transport: manifold as surfaces and matrices! Calculemus!

Property 2.2.1 (Inversion)**.** $\mathbb{G}$ Lie group, ∇ connection, $p, q \in \mathbb{G}$. Given γ such that $\gamma(0) = p$, $\gamma(1) = q$ and $\dot{\gamma}(0) = \mathbf{u} \in T_p\mathbb{G}$, we have:

$$\Pi(\gamma)_p^q(-\mathbf{u}) = -\Pi(\gamma)_p^q(\mathbf{u})) \tag{2.2}$$

$$p = \exp_q(\mathbf{u}) \iff q = \exp_p(-\Pi(\gamma)_q^p(\mathbf{u})) \tag{2.3}$$

xxx proof, see notebook!

In the next property we explore how does behave the affine exponential expressed as a composition when changed of sign. It shows how the usefulness of the parallel transport in extending the property -if $\exp(\mathbf{w}) = \exp(\mathbf{u}) \circ \exp(\mathbf{v})$ then $\exp(-\mathbf{w}) = \exp(-\mathbf{v}) \circ \exp(-\mathbf{u})$- at the case of the affine exponentials.

Property 2.2.2 (change of signs of the composition for affine exponential)**.** $\mathbb{G}$ Lie group, ∇ connection, $a, b \in \mathbb{G}$, $\mathbf{u} \in T_a\mathbb{G}$, $\mathbf{v} \in T_b\mathbb{G}$. Let β be the tangent curve to $\mathbf{u}$ at a and $c = \exp_b(\mathbf{v})$. Given $\mathbf{w} \in T_c\mathbb{G}$ such that

$$\exp_a(\mathbf{w}) = \exp_b(\mathbf{v}) \circ \exp_a(\mathbf{u})$$

Then

$$\exp_a(-\mathbf{w}) = \exp_{\tilde{b}}(-\Pi(\beta)_b^{\tilde{b}}(\mathbf{v})) \circ \exp_a(-\mathbf{u})$$

where $\tilde{b}$ is the affine exponential of $-\mathbf{u}$ or the element $\beta(-1)$.

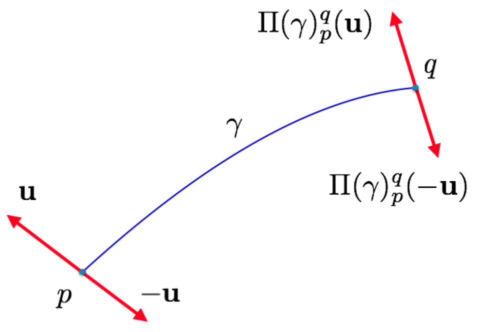

Figure 2.3: First inversion property.

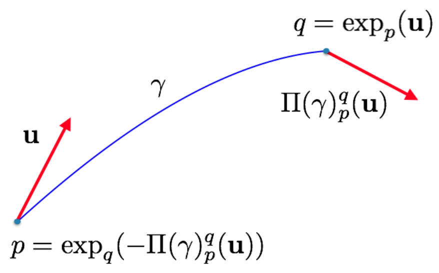

Figure 2.4: Second inversion property.

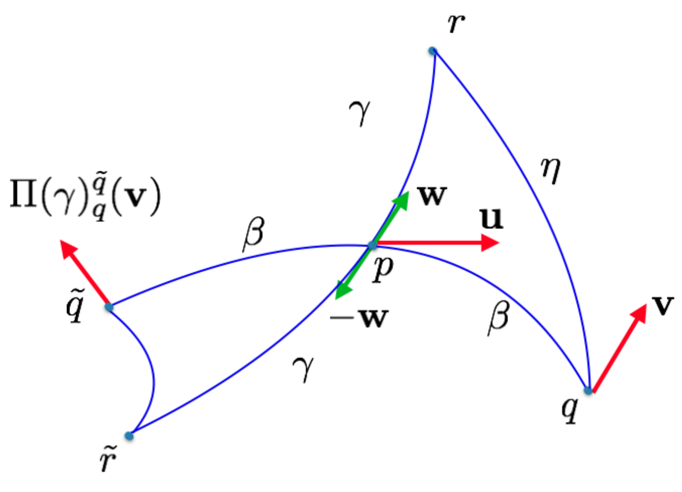

Figure 2.5: Change of sign property.

xxx proof, see notebook!
xxx Examples of these properties applied to real case!

2.2.4 Change of Base Formulas with and without Parallel Transport

Using the derivative of the left-translation L_p it is possible to bring back the exp and the log function based at the point p of the manifold to the one evaluated at the identity using the following formulas:

$$\log_p(q) = DL_p(e)\log_e(q)$$
$$\exp_p(\mathbf{u}) = p \circ \exp_e(DL_p(e)^{-1}\mathbf{u})$$

xxx Is this the same as the parallel transport? See examples and try to find a proof!!!

2.2.5 Parallel Transport in Practice: Schild's Ladder and Pole Ladder

Lemma 2.2.1. $\mathbb{G}$ Lie group, ∇ connection, $a \in \mathbb{G}$, $\mathbf{u} \in T_e\mathbb{G}$. Let γ be a curve defined on $\mathbb{G}$ such that $\gamma(0) = e$, $\gamma(1) = a$, $\dot{\gamma}(0) = \mathbf{u}$. Let β be the curve over $\mathbf{G}$ defined as $\beta(t) = a \circ \gamma(t)$, then the two following conditions hold:

1. If ∇ is a Cartan connection then β is a geodesic.
2. For $\mathbf{u}_a := D(L_a)_e(\mathbf{u}) \in T_a\mathbb{G}$:

$$\exp_a(t\mathbf{u}_b) = b \circ \exp_e(tD(L_{a^{-1}})_a(\mathbf{u}_a)) = b \circ \exp_e(t\mathbf{u}) \tag{2.4}$$

xxx proof, see notebook!

Theorem 2.2.1. Let $\mathbb{G}$ be a finite dimensional connected Lie group defined with a Cartan connection ∇. If, for each couple of linearly independent vectors $\mathbf{u}, \mathbf{v} \in T_e\mathbb{G}$, we consider the following elements:

$$a = \exp_e(\mathbf{u}) \qquad b = \exp_e(\mathbf{v})$$
$$\mathbf{u}^{\|} = \Pi(\alpha)_e^b(\mathbf{u})$$
$$\gamma : [0,1] \to \mathbb{G} \quad \gamma(0) = e \quad \dot{\gamma}(0) = \mathbf{v}$$

Then, for $\mathbf{u}_e^{\|} := D(L_{b^{-1}})_e(-\Pi(\alpha)_a^b(\mathbf{u}))$, the approximation

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

holds.

Proof. As a consequence of the construction we have the following considerations:

$$\gamma(t) = \exp(t\mathbf{w}) = a \circ \exp_e(D(L_{ba^{-1}})_e(t\mathbf{w})) = \exp_e(\mathbf{u}) \circ \exp_e(D(L_{a^{-1}})_e(t\mathbf{w}))$$
$$m = \alpha(\frac{1}{2}) = \exp_e\big(\frac{\mathbf{v}}{2}\big) = \gamma(1) = \exp_a(\mathbf{w})$$
$$\exp_e(D(L_{a^{-1}})_e(\mathbf{w})) = \exp_e(-\mathbf{u}) \circ \exp_e\big(\frac{\mathbf{v}}{2}\big)$$

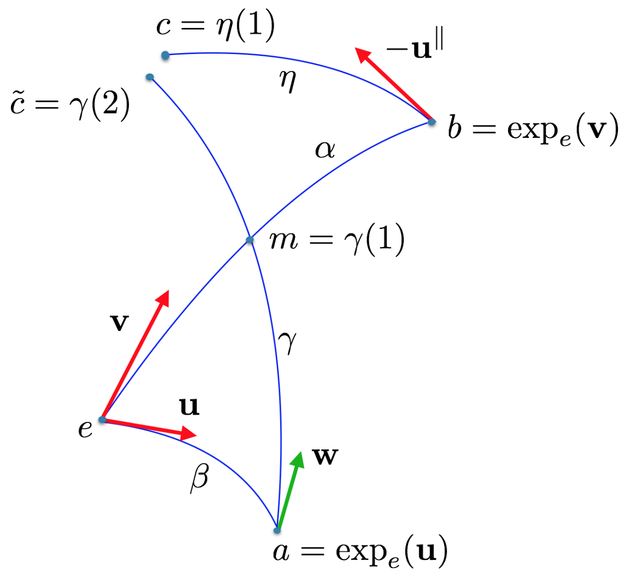

Figure 2.6: Pole ladder applied to parallel transport.

Let η be the integral curve of $-\Pi(\alpha)_a^b(\mathbf{u})$ starting at b. If $c := \eta(1)$ and $\tilde{c} := \gamma(1)$, then on one side we have:

$$\begin{aligned}
\tilde{c} = \gamma(1) &= \exp_a(2\mathbf{w}) = a \circ \exp_e(D(L_{a^{-1}})_e(2\mathbf{w})) \\
&= \exp_e(\mathbf{u}) \circ \exp_e(D(L_{a^{-1}})_e(\mathbf{2w})) \\
&= \exp_e(\mathbf{u}) \circ \exp_e(2D(L_{a^{-1}})_e(\mathbf{w})) \\
&= \exp_e(\mathbf{u}) \circ \big(\exp_e(D(L_{a^{-1}})_e(\mathbf{w}))\big)^2 \\
&= \exp_e(\mathbf{u}) \circ \big(\exp_e(-\mathbf{u}) \circ \exp_e(\frac{\mathbf{v}}{2})\big)^2 \\
&= \exp_e\big(\frac{\mathbf{v}}{2}\big) \circ \exp_e(-\mathbf{u}) \circ \exp_e\big(\frac{\mathbf{v}}{2}\big)
\end{aligned}$$

On the other side:

$$\begin{aligned}
c = \eta(1) &= \exp_b(-\mathbf{u}^{\|}) = b \circ \exp_e(D(L_{b^{-1}})_e(-\mathbf{u}^{\|})) \\
&= \exp_e(\mathbf{v}) \circ \exp_e(D(L_{b^{-1}})_e(-\mathbf{u}^{\|})) \\
&= \exp_e(\mathbf{v}) \circ \exp_e(-\mathbf{u}_e^{\|})
\end{aligned}$$

where $D(L_{b^{-1}})_e(\mathbf{u}^{\|})$ has been written $\mathbf{u}_e^{\|}$ for brevity. If we consider $c \simeq \tilde{c}$ it follows that:

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

which implies

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

As a consequence of property of the signs inversion it follows that

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

$\square$

Corollary 2.2.2. xxx attempt to measure the error in the formula... to be done in a more effective way! If, with previous notations, the condition (1) is an approximation

$$\exp_C(\frac{\mathbf{k}}{2}) = \exp(\xi) \circ \exp_M(\frac{\mathbf{k}}{2})$$

for some ξ in $\mathfrak{g}$ such that $\| \xi \| < \delta$ then the approximation has error

$$O(\| \delta\mathbf{u}^{\|} \|^2) + O(\| \mathbf{u} + \delta\mathbf{u} \|^3) + \text{xxx something that must be investigated depending on } \delta$$

2.3 Accelerating Convergences Series

xxx Think about it after 14 of may!

Space of the series of elements of $\mathfrak{g}$

$$S(\mathfrak{g}) = \{\sum_{j=0}^{\infty} \mathbf{u}_j \mid \mathbf{u}_j \in \mathfrak{g}\}$$

Series generate by

$$S(\mathbf{u}) = \sum_{j=0}^{\infty} \mathbf{u}^j$$

$$\exp(\mathbf{u}) = S_1 \cdot S(\mathbf{u})$$

where $S_1 = \sum_{k=0}^{\infty}(\frac{1}{k!})$ in the space of coefficient series $\cdot$ is the (infinite) scalar product in the space of series.

k-th series truncation:

$$S^k(\mathfrak{g}) = \{\sum_{j=0}^{k} \mathbf{u}_j \mid \mathbf{u}_j \in \mathfrak{g}\}$$

This notation may make sense as a starting point to define $\exp(\mathbf{u})$. The restriction to the first order truncation of the exp is the starting point the numerical approximation

$$\exp(\mathbf{u}) = 1 + \mathbf{u} \in S^1(\mathfrak{g})$$

Chapter 3

A Lie Group Perspective on Spatial Transformations

3.1 The Group of Rigid Body Transformations

People know or dimly perceive, that if thinking is not kept pure and keen, if spirit's contemplation do not holds, even mechanics of automobiles and ships will soon cease to run. Even engineer's slide rule, computations of banks and stock exchanges will wonder aimlessly for the lost of authority, and chaos will ensue.
-Hermann Hesse, *Magister Ludi*

A rigid body transformation in a normed vector space is a transformation that preserves distances. The set of rigid body transformations is constructed as any combination of rotations, translations and reflection. We are interested in two things about them: their expression in matrix form, and the Lie group structure they form. In particular we will develop the close formulas for the computation of the log composition, and for each approximation technique presented. In part **??**, they will be compared with the same technique applied to SVF where no ground truth is known.

3.1.1 Lie Logarithm and Exponential for SE(2)

The set of all of the 3×3 matrices with real entries is denoted with $M_3(\mathbb{R})$. Its subset, defined by all the matrices with non-zero determinant, and thus by all the invertible matrices, is denoted with $GL_3(\mathbb{R})$. A matrix group is any subgroup of $GL_3(\mathbb{R})$. We are interested in writing the group of rigid body transformation

$$\mathbb{G} = \{(\theta, tx, ty) \mid \theta \in [0, 2\pi), tx, ty \in \mathbf{R}^2\}$$

using matrices, so as a subgroup of $GL_3(\mathbb{R})$. Rotation in the plane can be expressed using matrix of the orthogonal group $SO(2)$, linear subgroup of $GL_2(\mathbb{R})$, so that rotations' actions on planes' points are simply defined as a product:

$$SO(2) = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mid \theta \in [0, 2\pi) \right\}$$

To include the translation we can add its $(tx, ty)^T$ parameter to the action of the rotation over the initial point $(x_i, y_i)^T$ to obtain the transformed $(x_t, y_t)^T$. So each element of the group $\mathbb{G}$ act over $\mathbf{R}^2$ as

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} tx \\ ty \end{pmatrix}$$

Another way to express rigid body transformation group's elements is to include the translation in a bigger matrix, subgroup (not linear, since the translation is not linear) of $GL_3(\mathbb{R})$. This is defined as the group $SE(2)$:

$$SE(2) = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) & t_x \\ \sin(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1 \end{pmatrix} \mid \theta \in [0, 2\pi), (tx, ty) \in \mathbf{R}^2 \right\}$$

Expressed in this way the matrices act on the point of the plane represented as the elements of the vector space $\{1\} \times \mathbf{R}^2$.
The passage between the restricted form $\mathbb{G}$ and $SE(2)$ is defined by the injection:

$$\begin{aligned} \rho_{\mathbb{G}} : \mathbb{G} &\longrightarrow SE(2) \\ (\theta, tx, ty) &\longmapsto \begin{pmatrix} \cos(\theta) & -\sin(\theta) & tx \\ \sin(\theta) & \cos(\theta) & ty \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

We are now interested the Lie algebra of the Lie group $SE(2)$. It is defined as:

$$\mathfrak{se}(2) = \left\{ \begin{pmatrix} 0 & -\theta & dt_x \\ \theta & 0 & dt_y \\ 0 & 0 & 1 \end{pmatrix} \mid \theta \in [0, 2\pi), (tx, ty) \in \mathbf{R}^2 \right\}$$

Expressing $r \in SE(2)$ as:

$$\mathbf{r} = \begin{pmatrix} R(\theta) & t \\ 0 & 1 \end{pmatrix} \qquad R(\theta) \in SO(2) \quad t \in \mathbb{R}^2$$

for t plane translation and $R(\theta)$ in $SO(2)$, then the element of the Lie algebra can be expressed as:

$$d\mathbf{r} = \begin{pmatrix} dR(\theta) & dt \\ 0 & 1 \end{pmatrix} \qquad R(\theta) \in SO(2) \quad t \in \mathbb{R}^2$$

Both $SE(2)$ and $\mathfrak{se}(2)$ are in bijective correspondence with $\mathbb{G}$, and both are subset of the bigger algebra of, The algebra $\mathfrak{se}(2)$ do not form a group with the operation of composition, but it is provided with the lie bracket defined by the commutator:

$$[d\mathbf{r}, d\mathbf{s}] = d\mathbf{r}d\mathbf{s} - d\mathbf{s}d\mathbf{r}$$

The Lie logarithm between Lie group and Lie algebra is given by:

$$\begin{aligned} \log : \mathfrak{se}(2) &\longrightarrow SE(2) \\ \begin{pmatrix} R(\theta) & t \\ 0 & 1 \end{pmatrix} &\longmapsto \begin{pmatrix} dR(\theta) & dt \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Where

$$dR(\theta) = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$$

and $dt = L(\theta)t$ for

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

The inverse function, Lie exponential is given by:

$$\begin{aligned} \exp : SE(2) &\longrightarrow \mathfrak{se}(2) \\ \begin{pmatrix} dR(\theta) & dt \\ 0 & 1 \end{pmatrix} &\longmapsto \begin{pmatrix} R(\theta) & t \\ 0 & 1 \end{pmatrix} \end{aligned}$$

where $t = L(\theta)^{-1}dt$ for

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

The proposed exponential function is not well defined over all $\mathfrak{se}(2)$.
In fact the elements of $\mathbb{G}$ can be identified with no risk with their matrices, while the same thing do not happen for the element of the Lie algebra $\mathfrak{g}$ of $\mathbb{G}$. If we formalize the passage between $\mathfrak{g}$ and $\mathfrak{se}(2)$ with the function:

$$\begin{aligned} \rho_{\mathfrak{g}} : \mathfrak{g} &\longrightarrow \mathfrak{se}(2) \\ (\theta, dtx, dty) &\longmapsto \begin{pmatrix} 0 & -\theta & dtx \\ \theta & 0 & dty \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

it is not an injection if we do not restrict its domain. In addition, given two elements (θ_0, dtx_0, dty_0) and (θ_1, dtx_1, dt_1) in $\mathfrak{g}$, with $\theta_1 \neq 0$, we have that for each $k \in \mathbb{Z}$, if

$$\theta_0 = \theta_1 + 2k\pi$$

and

$$(dtx_0, dty_0) = \frac{\theta_0}{\theta_1}(dtx_1, dty_1)$$

then

$$\exp(\theta_0, dtx_0, dty_0) = \exp(\theta_1, dtx_1, dty_1)$$

The exponential is then well defined only on the quotient of $\mathfrak{g}$ over the relation $\sim$, defined by

$$(\theta_0, dtx_0, dty_0) \sim (\theta_1, dtx_1, dt_1) \iff \exp(\theta_0, dtx_0, dty_0) = \exp(\theta_1, dtx_1, dt_1)$$

The quotient set $\mathfrak{g}/\sim$ coincides the neighborhood U of the identity on which the function $\rho_{\mathfrak{g}}$ becomes an injection

$$\rho_{\mathfrak{g}/\sim} : \mathfrak{g} \longrightarrow \mathfrak{se}(2)$$

and exp is a bijection having log as its inverse. What said so far can be summarize in the following commutative diagram:

$$
\begin{tikzcd}
\mathfrak{g} \arrow[rrr, "\rho_{\mathfrak{g}}"] \arrow[dr, "\pi"] & & & \mathfrak{se}(2) \\
& \mathfrak{g}/\sim \arrow[r, "\rho_{\mathfrak{g}/\sim}"] \arrow[dl, "\exp"] & U \subset \mathfrak{se}(2) \arrow[ur, hook] \arrow[dr, "\exp"] & \\
\mathbb{G} \arrow[uu, "\log"] \arrow[rrr, "\rho_{\mathbb{G}}"] & & & SE(2) \arrow[uu, "\log"']
\end{tikzcd}
$$

We can see that the function $\rho_{\mathfrak{g}}$ is the inverse of a restriction of the general vectorization function that aligns column vector in a single vector. This will be particularly useful for our purposes.

Matrix Vectorization

xxx this part must be set after subsection 2.5 is done, to avoid repetitions and circular properties!
We can see that the function $\rho_{\mathfrak{g}}$ is the inverse of a restriction of the general vectorization function that aligns column vector in a single vector:

$$
\begin{aligned}
\mathrm{Vect} : M_3(\mathbb{R}) &\longrightarrow \mathbb{R}^{3\times 3} \\
[A_1|A_2|A_3] &\longmapsto (A_1^t, A_2^t, A_3^t)
\end{aligned}
$$

Thanks to this adjoint action can be defined as an action over The vectorization, in combination with Lie bracket, Kronecker product, adjoint action and adjoint map, satisfies the following properties:

- $\mathrm{Vect}([M, X]) = (I \otimes M - M^t \otimes I)\mathrm{Vect}(X)$
- $\mathrm{Vect}([X, M]) = (M^t \otimes I - I \otimes M)\mathrm{Vect}(X)$

These are still valid for its restriction

$$
\begin{aligned}
\mathrm{Vect}^{\sim} : M_3(\mathbb{R}) &\longrightarrow \mathbb{R}^{3} \\
[A_1|A_2|A_3] &\longmapsto (a_{2,1}, a_{3,1}, a_{3,2})
\end{aligned}
$$

that respects the Lie group operations between the restricted representation $\mathfrak{g}$ and the matrix representation $SE(2)$:
and will be used in the next subsection to compute the log composition.

3.1.2 Close Formula for the Log-composition

In the finite-dimensional case, investigate here the log-composition can be computed with a close formula:

$$d\mathbf{r}_1 \star d\mathbf{r}_2 = \log(\exp(d\mathbf{r}_1) \circ \exp(d\mathbf{r}_2))$$

which results

$$d\mathbf{r}_1 \star d\mathbf{r}_2 = xxx\text{On some lost paper... to be computed again!}$$

3.1.3 Taylor Approximation to compute the Log-composition

We can apply the Taylor expansion formula for the computation of the affine log-composition to matrices in $SE(2)$. From previous subsection we have:

$$\mathbf{u} \star \mathbf{v} = \mathbf{u} + \frac{ad_{\mathbf{u}} \exp(ad_{\mathbf{u}})}{\exp(ad_{\mathbf{u}}) - 1}\mathbf{v} + O(\mathbf{v}^2) \qquad \frac{ad_{\mathbf{u}} \exp(ad_{\mathbf{u}})}{\exp(ad_{\mathbf{u}}) - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} ad_{\mathbf{u}}^n$$

Where $\{B_n\}$ is the sequence of the second-kind Bernoulli number[1].

3.1.4 Parallel Transport to compute the Log-composition

3.1.5 Log and Exp Approximations for little rotations

Computations of logarithm and exponential obtained so far are a consequence of these formula:

$$\exp(\mathbf{r}) = \sum_{k=0}^{\infty} \frac{\mathbf{v}^k}{k!} \qquad \log(\mathbf{r}) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\mathbf{v} - I)^k}{k!}$$

Remarkably, infinite series of elements of a group (whose sum is not even defined within the group structure) is an element into an associated algebra, while another infinite series of matrices of the algebra appears to be the natural way to going backward. A second door to passing from one structure to the other, when $\mathbf{r}$ is little appears to be the following approximation:

$$\exp(\mathbf{r}) \simeq I + \mathbf{r} \qquad \log(d\mathbf{r}) \simeq d\mathbf{r} - I$$

In fact for little θ, $\sin(\theta) \simeq \theta$, $\cos(\theta) \simeq 0$ and $L(\theta)^{-1} \simeq I$.
xxx this may deserve an investigation about the errors in the approximations error!

3.2 The Set of Stationary Velocity Fields

Accurate reckoning: the entrance into knowledge of all existing things and all obscure secrets.
- Ahmes, 1800 B.C.

The set of diffeomorphisms can be seen as an infinite dimensional Lie group. For these reasons xxx we reduce the set of transformation to the SVF. This has the following positive consequences xxx and it has been applied in xxx . Nevertheless reducing the set of transformation to the SVF has bring new issues and challenges xxx - limitations -

3.2.1 Set and Set Only

Let Ω be an open connected subset of $\mathbb{R}^d$ containing the origin. We define Diff(Ω) the infinite dimensional Lie group of diffeomorphism over Ω with neutral element e:

$$\text{Diff} := \{ f : \mathbb{R}^d \longrightarrow \mathbb{R}^d \mid \text{ diffeomorphism } \}$$

xxx Short non-formal part about recognizing Diff(Ω) as a Lie group. Banach manifold and Frechet manifold. What does imply the naive theory of infinite dimensional manifold.

[1] If first-kind Bernoulli number is used then each term of the summation must be multiplied for $(-1)^n$, as did for example in ....Klarsfeld.

3.2.2 Some Tools for the Infinite Dimensional Case

It can be proved that the Lie algebra of $\text{Diff}(\mathbb{R}^d)$ is isomorphic to the Lie algebra of the vector field over $\mathbb{R}^d$.

$$Lie(\text{Diff}(\mathbb{R}^d)) = \mathcal{V}(\mathbb{R}^d) \tag{3.1}$$

To Visualize the meaning of this isomorphism we can consider the following diagram:

$$\begin{tikzcd}
 & T_{\mathbf{0}}\mathbb{R}^D \arrow[r, "\rho_\star"] \arrow[d] & T_{\rho(\mathbf{0})}\mathbb{R}^D \arrow[d] \\
\mathbb{R} \arrow[ur, dashed, "\gamma'(0)"] \arrow[urr, dashed, "(\rho\circ\gamma)'(0)"'] \arrow[r, "\gamma"] & \mathbb{R}^D \arrow[r, "\rho"] \arrow[dr, dashed, "f\circ\rho"] & \mathbb{R}^D \arrow[d, "f"] \\
 & & \mathbb{R}
\end{tikzcd}$$

where $(\rho_\star)_{\mathbf{0}}$ is the push forward of ρ, defined as follows:

$$\begin{aligned}
\rho_\star : T_{\mathbf{0}}\mathbb{R}^D &\longrightarrow T_{F(\mathbf{0})}\mathbb{R}^D \\
\mathbf{v} &\longmapsto \rho_\star\mathbf{v} : \mathcal{C}^\infty(\mathbb{R}^D) \longrightarrow \mathbb{R} \\
&\qquad\qquad f \longmapsto \rho_\star\mathbf{v}f := \mathbf{v}(f\circ\rho)
\end{aligned}$$

We can consider the first floor of the diagram as the group of diffeomorphism of $\mathbb{R}^d$ and the second floor of the diagram as the algebra of the continuous function from $\mathbb{R}^d$ to $\mathbb{R}^d$. For $\rho \in \text{Diff}(\mathbb{R}^D)$ and $\gamma : (-\eta, \eta) \to \mathbb{R}^D$ such that $\gamma(0) = \mathbf{0}$, then $(\rho\circ\gamma)'(0)$ belongs to $T_{\rho(\mathbf{0})}\mathbb{R}^D$ and $(\rho_\star)$ is a continuous function from $\mathbb{R}^D$ to $\mathbb{R}^D$ and belongs to $\text{Lie}(\text{Diff}(\mathbb{R}^D))$.

Lemma 3.2.1 (existence)**.** Be $p \in \text{Diff}(\mathbb{R}^d)$, then exists a $\mathbf{v}$ in the Lie algebra $\mathcal{V}(\mathbb{R}^d)$, such that

$$||p - \exp(V)|| < \delta$$

for some δ and for some metric in $\text{Diff}(\mathbb{R}^d)$.

Proof. Investigate a proof to define δ. □

Lemma 3.2.2 (identity lemma)**.** Be $p \in \text{Diff}(\mathbb{R}^d)$, such that $p = \exp(\mathbf{v})$ for some $\mathbf{v} \in \mathfrak{g}$. Then

$$\exp(-\mathbf{v}) \circ p = p \circ \exp(-\mathbf{v}) = e$$

Proof.

$$\exp(-\mathbf{v}) \circ p = \exp(-\mathbf{v}) \circ \exp(\mathbf{v}) = \varphi_1\varphi_{-1} = \varphi_0 = e$$
$$p \circ \exp(-\mathbf{v}) = \exp(\mathbf{v}) \circ \exp(-\mathbf{v}) = \varphi_{-1}\varphi_1 = \varphi_0 = e$$

□

Property 3.2.1. If $\mathbf{v}$ is close to the origin $\exp(\mathbf{v})$ can be numerically approximated with:

$$\exp(\mathbf{v}) = e + \mathbf{v}$$

Proof. xxx !! □

xxx Exp and Log function in the infinite dimension case

xxx we can not express the exp function using the Taylor expansion in the infinite dimensional Lie group Diff(Ω). We define it as an unknown function with some features related to the 1-parameter subgroup structure over $\mathbb{G}$: We define exp function as

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

The following properties are satisfied:

1. exp is well defined and surjective (at least near 0).
2. If $\exp(\mathbf{v}) = \gamma(1)$ then $\exp(t\mathbf{v}) = \gamma(t)$.
3. It satisfies the one parameter subgroup property.
4. It satisfies the differential equation

$$\frac{d}{dt}\exp(t\mathbf{v})\Big|_{t=0} = \mathbf{v}$$

We observe that exp respects the one parameter subgroup structure of the Lie group $\mathbb{G}$: stretching the tangent vector $\mathbf{v}$ by a parameter t, the same stretch is reflected in $\exp(V)$ along the same integral curve.

In addition exp respect the 1-parameter subgroup structure:

$$\exp((t+s)\mathbf{v}) = \varphi_{t+s} = \gamma(t+s) \qquad \forall t, s \in \mathbb{R}$$

moreover, if two elements p_1, p_2 of $\mathbb{G}$ belongs to integral curve passing in e of the same integral curve defined by $\mathbf{v}$, their log function are a vectors having the same direction:

$$p_1 = \gamma(t_1),\ p_2 = \gamma(t_2) \Rightarrow \exists \mathbf{v} \in \mathfrak{g},\ \lambda \in \mathbb{R} \mid \log(p_1) = \mathbf{v},\ \log(p_2) = \lambda\mathbf{v}$$

It follows that for a fixed $t \in \mathbb{R}$ and $\gamma(t) = \exp(\mathbf{v})$ for some $\mathbf{v}$ in $left\mathfrak{X}(\mathbb{G})$, then $\gamma(1) = \exp(\frac{1}{t}\mathbf{v})$.

xxx Issue related to the image of exp for stationary velocity fields in the finite dimensional case. define $\text{Diff}_s(\Omega)$ as the subset of Diff(Ω) defined by the images of exp from the tangent space to the Lie group.

xxx We define *log* function as

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

such that for p in $\mathbb{G}$ we have $\exp(log(p)) = p$ when $log(p)$ is defined.

3.2.3 SVF in Practice

Three main definitions around which the whole theory of diffeomorphic image registration gravitate are introduced in this subsection.
xxx Define here displacement and deformation.
xxx the set of time dependent spatial transformation. We can express it as the set of continuous functions from Ω to $\mathbb{R}^d$ depending on a real parameter in $T \subseteq \mathbb{R}$:

$$\mathcal{V}_T(\mathbb{R}^d) = \mathcal{V}_T := \{V : \Omega \times T \longrightarrow \mathbb{R}^d \mid \text{ continuous } \}$$

its elements are called *time varying velocity field* (TVVF) and can be expressed as

$$V(\mathbf{x}, t) = \sum_{i=1}^{d} v_i(\mathbf{x}, t) \frac{\partial}{\partial x_i} \Big|_{\mathbf{x}} \qquad v_i \in \mathcal{C}^{\infty}(\Omega \times T)$$

In case $V(\mathbf{x}, t) = V(\mathbf{x}, s)$ for all s, t real, then V is a *stationary velocity field* (SVF), and the set of the stationary velocity field, second item presented in this subsection, is defined as

$$\mathcal{V}(\mathbb{R}^d) = \mathcal{V} := \{V : \Omega \longrightarrow \mathbb{R}^d \mid \text{ continuous } \}$$

Their elements can be expressed as

$$V(\mathbf{x}) = V_{\mathbf{x}} = \sum_{i=1}^{d} v_i(\mathbf{x}) \frac{\partial}{\partial x_i} \Big|_{\mathbf{x}} \qquad v_i \in \mathcal{C}^{\infty}(\Omega)$$

While $\mathcal{V}$ and $\mathcal{V}_T$ are Lie algebra, Diff is a Lie group with the operation of composition.

If we imagine a particle starting at the point $\mathbf{x}$ of $\Omega \subseteq \mathbb{R}^d$ at time 0, with velocity vector for each instant of time given by $V(\mathbf{x}, t)$, then its trajectory $\gamma = \gamma(t)$ is determined by the ODE:

$$\frac{d\gamma}{dt} = V(\mathbf{x}, t)$$

In case $V(\mathbf{x}, t)$ is a stationary velocity field the equation is *stationary* or *autonomous*.
xxx if $V^{(t)} = V^{(s)}$ for all s, t real, then we call this vector field *stationary velocity field* (SVF), otherwise are called *time varying velocity field* (TVVF).
xxx The set of the SVF can be expressed as

$$\text{SVF} := \{\varphi_t(e) \mid t \in \mathbb{R}, \dot{\varphi}_t(e) = V_{\varphi_t(e)}, V \in \mathfrak{V}(\Omega)\}$$

(note that in this way V is not an element of the Lie algebra!! We should have said $V \in left\mathfrak{V}(\text{Diff})$).
xxxWe know that SVF are geodesics-complete if a norm over diff is defined, while SVF are not complete. $\varphi_t(e)$ do not spans Diff i.e. for each point of Diff may not always pass an integral curve of a left-invariant vector field over Diff. We will consider only the element of Diff of the form $\varphi_t(e)$. We assume also that each vector field is complete. (THIS MUST BE INVESTIGATED LATER!)
xxxThanks to the Dini theorem we have that a SVF can be considered locally as an element of a local expression of Diff. Moreover to each spatial transformation vector field corresponds an element of the one parameter subgroup of local transformation over $\mathbb{R}^2$ (not sure...).

xxx practical aspects, discretization, structure as they are considered into the practical side!

Chapter 4

Applications

We believe that we know something about the things themselves when we speak of trees, colors, snow, and flowers; and yet we possess nothing but metaphors for things — metaphors which correspond in no way to the original entities.
-Nietzsche, *On Truth and Lies in extra-moral sense.*

zz write chapter intro.

4.1 Numerical Approximation of the Log-Composition

xxx Statistics on anatomies:
xxx It is of fundamental importance to have the possibility to going from an element of a group of spatial transformations to a tangent space, in which each vector corresponds to the tangent vector field that this transformation causes on the space. It makes possible to lean on the group structure a structure of vector space, which implies the possibility to compute statistics on the group of transformation as well as compose velocity fields in the tangent space passing through the corresponding transformation (both of them are made possible thanks to the local bijection between the Lie group and the Lie algebra).

4.1.1 Group composition at the Service of Image Registration

xxx Log-composition in diffeomorphic image registration.

where $p_1 \in \mathbb{G}$, $p_2 = \exp_{p_1}(\mathbf{v}_1)$ and Affine exponential and Affine logarithm are considered. Posing $p_3 = \exp_{p_2}(\mathbf{v}_2)$ the second offset group composition is the tangent vector in p_1 that corresponds to the composition between p_1 and p_2, i.e. $\log_{p_1}(p_3 \circ p_2)$.

4.1.2 BCH formula to compute the Log-composition

The BCH formula is the exact solution to the Log-composition. 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. It can be practically computed using its *approximation of degree k* defined as the sum of the BCH terms having no more than k nested Lie bracket.

For example:

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

These numerical approximations of the group composition leave the difficulty of managing the problem of the error carried by each term. In some cases the increase of the degree of the BCH approximation do not necessarily implies a decrease in error:
.... Add an example in which this happens.

We present other ways to compute the Lie group composition in the following subsubsections.

4.1.3 Accelerating convergences applied to the Log-composition

....

4.1.4 Taylor expansion to compute the Log-composition

....

4.1.5 Parallel transport to compute the Log-composition

....Here will be made the strong assumption according to which the parallel transport defined for the finite dimensional case, works also in the infinite dimensional case....

4.2 Numerical Approximation to Compute the Lie logarithm

The problem of the computation of the logarithm computation can be stated as follows: given $p \in \mathbb{G}$ the goal is to find $\mathbf{u}$ such that $\exp(\mathbf{u})$ is the best possible approximation of p. This chapter is devote to the numerical computation of the logarithm, using an iterative algorithm based on the Log-composition. In this context each of the presented techniques are suitable to perform this computation.

Tools for the Computation from Truncated Series

xxx These must be posed in the truncated series context after chapter 4 is investigated.

We emphasize the fact that if $\mathfrak{g}$ and $\mathbb{G}$ are subset of a bigger algebra, then exp and log can be considered as infinite series. Remarkable consequence is the approximation of $\exp(\mathbf{v})$ with $1+\mathbf{v}$ if the transformation $\mathbf{v}$ is small. This approximation is the base of what follows in this chapter.
In parallel with the log computation given by:

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

we define two approximating functions:

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

and

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

Where $\mathfrak{g}^{\sim}$ is a space of approximations of elements of $\mathfrak{g}$,and $\mathbb{G}^{\sim}$ is a space of approximations of elements in $\mathbb{G}$ (xxx that requires some more investigations and formal definition in conjunction with truncated series).
Consequence of this definition is the fact that

$$\mathbf{u} \simeq \mathrm{app}(\mathbf{u}) \qquad\qquad \exp(\mathbf{u}) \simeq \mathrm{App}(\exp(\mathbf{u}))$$

xxx errors can be investigated and maybe can become known elements in the computations!
The two following straightforward properties, that holds for all $\mathbf{u}, \mathbf{v}$ in the Lie algebra

1. $\mathbf{u} = \mathbf{v} \star (-\mathbf{v} \star \mathbf{u})$
2. $\mathrm{app}(\mathbf{v} \star \mathbf{u}) = \exp(\mathbf{v}) \exp(\mathbf{u}) - 1 \in \mathfrak{g}^{\sim}$

lead us to consider the algorithm presented in [BO08] under a new perspective.

4.2.1 A reformulation of the Bossa 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

$$\mathbf{u} = \mathbf{u}_j \star (-\mathbf{u}_j \star \mathbf{u}) \Longrightarrow \mathbf{u} \simeq \mathbf{u}_j \star \mathrm{app}(-\mathbf{u}_j \star \mathbf{u})$$

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

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

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

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

Each strategy that we have examined to compute the Lie composition, become a numerical method for the computation of the logarithm.

BCH Strategy

At each step, we compute the approximation $\mathbf{v}_{j+1}$ with the k-th truncation of the BCH formula:

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

thus, for the first degree we have

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

For the second degree we have:

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

Theorem 4.2.1 (Bossa)**.** The iterative algorithm (4.2) converges to $\mathbf{v}$ with error $\delta_n \in \mathbb{G}$, where

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

Parallel Transport Strategy

If we use the parallel transport for the computation of the log-composition, we obtain:

$$\begin{cases}\mathbf{u}_0 = \mathbf{0} \\ \mathbf{u}_t = \mathbf{u}_{t-1} + \exp(-\frac{\mathbf{u}_{t-1}}{2}) \circ \exp(\delta\mathbf{u}_{t-1}) \circ \exp(\frac{\mathbf{u}_{t-1}}{2}) - e\end{cases} \tag{4.3}$$

Symmetrization Strategy

The algorithm for the computation of the group logarithm can be improved considering a symmetric version of the underpinning strategy. In this version we use the first order approximation of the BCH formula (see equation (4.6) in the following proof), compensating with the fact that the symmetrization should decrease the error involved. It gives birth to the following algorithm:

$$\begin{cases}\mathbf{v}_0 = \mathbf{0} \\ \mathbf{v}_{t+1} = \mathbf{v}_t + \frac{1}{2}(\tilde{\delta}\mathbf{v}_t^L + \tilde{\delta}\mathbf{v}_t^R)\end{cases} \tag{4.4}$$

Where $\tilde{\delta}\mathbf{v}_t^R = \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) - e$ and $\tilde{\delta}\mathbf{v}_t^L = \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) - e$.

Proof. To show why it works we remind that the starting point was

$$p = \exp(\mathbf{v}) = \exp(\mathbf{v}_0) \circ \exp(\delta\mathbf{v}_0)$$

where $\exp(\delta\mathbf{v}_0) = \exp(-\mathbf{v}_0) \circ p$.
An equivalent starting point would have been $\exp(\mathbf{v}) = \exp(\delta\mathbf{v}) \circ \exp(\mathbf{v}_0)$ for $\exp(\delta\mathbf{v}) = p \circ \exp(-\mathbf{v}_0)$.
This idea leads to the definition of

$$\begin{aligned}\exp(\delta\mathbf{v}_t^R) &:= p \circ \exp(-\mathbf{v}_t) = \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) \\ \exp(\delta\mathbf{v}_t^L) &:= \exp(-\mathbf{v}_t) \circ p = \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v})\end{aligned}$$

It follows that

$$\exp(\mathbf{v}) = \exp(\mathbf{v}_0) \circ \exp(\delta \mathbf{v}_0^R)$$
$$\exp(\mathbf{v}) = \exp(\delta \mathbf{v}_0^L) \circ \exp(\mathbf{v}_0)$$

Using $\exp(\delta \mathbf{v}_t^R) \approx e + \delta \mathbf{v}_t^R$ and $\exp(\delta \mathbf{v}_t^L) \approx e + \delta \mathbf{v}_t^L$ we can use the following approximation to define the symmetric algorithm:

$$\begin{aligned}
\exp(\delta \mathbf{v}_t^R) &= \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) \\
e + \tilde{\delta} \mathbf{v}_t^R &= \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) \\
\tilde{\delta} \mathbf{v}_t^R &= \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) - e
\end{aligned}$$

$$\begin{aligned}
\exp(\delta \mathbf{v}_t^L) &= \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) \\
e + \tilde{\delta} \mathbf{v}_t^L &= \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) \\
\tilde{\delta} \mathbf{v}_t^L &= \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) - e
\end{aligned}$$

Which gives birth to iterative algorithm, for a given initial value V_0:

$$\begin{cases} \mathbf{v}_0 \\ \mathbf{v}_{t+1} = \mathrm{BCH}(\mathbf{v}_t, \tilde{\delta} \mathbf{v}_t^R) \end{cases} \quad \begin{cases} \mathbf{v}_0 \\ \mathbf{v}_{t+1} = \mathrm{BCH}(\tilde{\delta} \mathbf{v}_t^L, \mathbf{v}_t) \end{cases} \tag{4.5}$$

If follows that

$$\mathbf{v}_{t+1} = \frac{1}{2}(\mathrm{BCH}(\tilde{\delta} \mathbf{v}_t^L, \mathbf{v}_t) + \mathrm{BCH}(\mathbf{v}_t, \tilde{\delta} \mathbf{v}_t^R))$$

Taking the first order approximation of the BCH formula:

$$BCH(\tilde{\delta} \mathbf{v}_t^L, \mathbf{v}_t) \approx \tilde{\delta} \mathbf{v}_t^L + \mathbf{v}_t \tag{4.6}$$
$$BCH(\mathbf{v}_t, \tilde{\delta} \mathbf{v}_t^R) \approx \mathbf{v}_t + \tilde{\delta} \mathbf{v}_t^R \tag{4.7}$$

we get

$$\mathbf{v}_{t+1} = \mathbf{v}_t + \frac{1}{2}(\tilde{\delta} \mathbf{v}_t^L + \tilde{\delta} \mathbf{v}_t^R)$$

□

We observe that the symmetric approach do not requires to use the BCH formula at each passage, having considered the approximation at the first order of the BCH.
We conclude with a formula that relates $\tilde{\delta} \mathbf{v}_t^L$ with $\tilde{\delta} \mathbf{v}_t^R$:

Theorem 4.2.2. Be $\tilde{\delta} \mathbf{v}_t^R = \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) - e$ and $\tilde{\delta} \mathbf{v}_t^L = \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) - e$ as before, then

$$\delta \mathbf{v}_t^L \approx \exp(-\mathbf{v}_t) \circ \delta \mathbf{v}_t^R \circ \exp(\mathbf{v}_t)$$

Proof. Since $\exp(\mathbf{v}_t) \circ \exp(\delta \mathbf{v}_t^R) \approx exp(\delta \mathbf{v}_t^L) \circ \exp(\mathbf{v}_t)$ it follows

$$\exp(\delta \mathbf{v}_t^R) = \exp(-\mathbf{v}_t) \circ \delta \mathbf{v}_t^L \circ \exp(\mathbf{v}_t)$$

Using $\exp(\delta \mathbf{v}_t^R) = e + \delta \mathbf{v}_t^R$ and $\exp(\delta \mathbf{v}_t^L) = e + \delta \mathbf{v}_t^L$ we get

$$\begin{aligned}
e + \delta \mathbf{v}_t^R &= \exp(-\mathbf{v}_t) \circ (e + \delta \mathbf{v}_t^L) \circ \exp(\mathbf{v}_t) \\
\delta \mathbf{v}_t^R &= \exp(-\mathbf{v}_t) \circ \delta \mathbf{v}_t^L \circ \exp(\mathbf{v}_t)
\end{aligned}$$

□

Symmetric-Parallel Transport Strategy

If we are not satisfied to having take only the firs order approximation of the BCH in the equation (4.6) we use at this stage the parallel transport in the method presented in this subsection. Going back to the algorithm 4.4 we can apply to

$$\mathbf{v}_{t+1} = \frac{1}{2}(\mathrm{BCH}(\tilde{\delta}\mathbf{v}_t^L, \mathbf{v}_t) + \mathrm{BCH}(\mathbf{v}_t, \tilde{\delta}\mathbf{v}_t^R))$$

the parallel transport to get

$$\begin{aligned}\mathbf{v}_{t+1} &= \frac{1}{2}((\tilde{\delta}\mathbf{v}_t^L)^{\parallel} + \mathbf{v}_t + \mathbf{v}_t + (\tilde{\delta}\mathbf{v}_t^R)^{\parallel}) \\ &= 2\mathbf{v}_t + \frac{1}{2}((\tilde{\delta}\mathbf{v}_t^L)^{\parallel} + (\tilde{\delta}\mathbf{v}_t^R)^{\parallel})\end{aligned}$$

Applying the definition of parallel transport we get

$$(\tilde{\delta}\mathbf{v}_t^L)^{\parallel} + (\tilde{\delta}\mathbf{v}_t^R)^{\parallel} = \exp(-\frac{\mathbf{v}_t}{2}) \circ (\tilde{\delta}\mathbf{v}_t^L + \tilde{\delta}\mathbf{v}_t^R) \circ \exp(\frac{\mathbf{v}_t}{2})$$

where

$$\begin{aligned}\tilde{\delta}\mathbf{v}_t^L &= \exp(\mathbf{v}) \circ \exp(-\mathbf{v}_t) - e \\ \tilde{\delta}\mathbf{v}_t^R &= \exp(-\mathbf{v}_t) \circ \exp(\mathbf{v}) - e\end{aligned}$$

Then a new improvement of the algorithm **??** is

$$\begin{cases}\mathbf{v}_0 = 0 \\ \mathbf{v}_t = 2\mathbf{v}_{t-1} + \frac{1}{2}(\exp(-\frac{\mathbf{v}_{t-1}}{2}) \circ (\tilde{\delta}\mathbf{v}_{t-1}^L + \tilde{\delta}\mathbf{v}_{t-1}^R) \circ \exp(\frac{\mathbf{v}_{t-1}}{2}))\end{cases} \tag{4.8}$$

(This must be investigated!)

4.3 Experimental Results

4.3.1 Toy Examples to Compare the Log-Compositions

4.3.2 NiftyReg Applications

4.3.3 BCH-free Computation of the Lie Logarithm

4.4 Further Research and Conclusion

Appendix: NiftyReg and NiftyBit

xxx This will be a short a description of the tools you used to make experiments for this thesis.

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.

[Arn]

[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.

[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.

[CSR11] X. Pennec C. Seiler and M. Reyes. Geometry-aware multiscale image registration via obbtree-based polyaffine log-demons, 2011.

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

[DGM98a] Paul Dupuis, Ulf Grenander, and Michael I. Miller. Variational problems on flows of diffeomorphisms for image matching, 1998.

[DGM98b] Paul Dupuis, Ulf Grenander, and Michael I Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of applied mathematics*, 56(3):587, 1998.

[HBHH01] Derek L G Hill, Philipp G Batchelor, Mark Holden, and David J Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1, 2001.

[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.

[HOP08] Monica Hernandez, Salvador Olmos, and Xavier Pennec. Comparing algorithms for diffeomorphic registration: Stationary lddmm and diffeomorphic demons. In *2nd MICCAI workshop on mathematical foundations of computational anatomy*, pages 24–35, 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.

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

[Lor12] Marco Lorenzi. *Deformation-based morphometry of the brain for the development of surrogate markers in Alzheimer's disease*. Ph.d. thesis, University of Nice Sophia Antipolis, December 2012.

[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.

[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.

[MPS+11] Tommaso Mansi, Xavier Pennec, Maxime Sermesant, Hervé Delingette, and Nicholas Ayache. iLogDemons: A demons-based registration algorithm for tracking incompressible elastic biological tissues. *International Journal of Computer Vision*, 92(1):92–111, 2011.

[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.

[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.

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

[Tro]

[Ver08] Tom Vercauteren. Image registration and mosaicing for dynamic in vivo fibered confocal microscopy, 2008.

[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.

[VRRC12] François-Xavier Vialard, Laurent Risser, Daniel Rueckert, and Colin J Cotter. Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision*, 97(2):229–241, 2012.

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