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

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

June 3, 2015

Supervisor
Tom Vercauteren

Co-Supervisor
Marc Modat

Abstract

Medical imaging employs techniques and tools belonging to several branches of mathematics and physics that, when applied to morphometric and statistical study of biological shape variability, are collected under the name of *computational anatomy*. One of these techniques, image registration, is widely used in both academical studies and applications, and continuously challenges researchers to enhance accuracy, improve reliability and reduce the time of the computations. The use of diffeomorphisms in image registration and the concomitant introduction of the log-euclidean framework to compute statistics, provide an interesting options to model the organs' deformations and to quantify the variations of anatomies. Main challenges of this setting are the numerical computation of the Lie logarithm and Lie exponential, and their combinations as they appear in the BCH formula: concept at the core of the underpinning theory of the *log-demons* algorithm. The necessity of finding fast numerical computation techniques of the BCH formula gave birth to the concept of log-composition presented in this thesis, with some strategies for its computation.

Thesis' Organization

Chapter 1 The first chapter is devoted to the introduction of diffeomorphic image registration and the main recent advancements. Here are introduced advantages and disadvantages of using diffeomorphisms as well as some of the milestone frameworks that led to, or are directly affected by the log-composition.

Chapter 2 The second one is the chapter of mathematical elements and tools; called from Lie group theory and hired to service image registration, techniques and tools are formally defined with a particular attention to flows, left translation, push forward, Lie logarithm, Lie exponential, connections and parallel transport. The concept of log-composition, around which the research gravitates, is defined as consequence of the need of generalize the BCH formula. Since its computation is impractical, other paths as Taylor expansion, parallel transport and accelerating convergence series are proposed as feasible alternatives.

Chapter 3: Each of the numerical methods for the computation of the log-composition are not computed and tested directly on the group of diffeomorphisms. Initially it they are applied to the finite dimensional group of rigid body transformation, where logarithms and exponentials possess a closed form. This chapter is aimed to present the customization of the tools defined in the previous chapter for the group of rigid body transformation and for the group of diffeomorphisms.

Chapter 4: Since the algorithm to compute the logarithm (here called *log-computation algorithm*) presented in [BHO07] relies on the BCH formula, the log-composition provides a range of methodologies for its computation. Each of the tools presented here for the log-composition become naturally available to challenge its performance.

Chapter 5: Here we are at the experimental results. The log-composition applied to rigid and diffeomorphic registration is applied to synthetic data and clinical images.

Contents

Chapter 1

Introduction to Diffeomorphic Image Registration

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

1.1 Toward an ill-posed Problem

The process of determining correspondences between two or more images acquired from patients scans is a challenging task that has seen the application of a growing number of mathematical theories contributing to its solution.
The challenge and the concomitant difficulties in approaching the problem is a consequence of the fact that dealing with image registration problem means dealing with an ill-posed problem. Transformations between anatomies are not unique, and the impossibility to recover spatial or temporal evolution of an anatomical transformation from temporally isolated images, makes any validation a difficult, if not an impossible task. In addition each situation inevitably leads to consider some prior knowledge within the initial data, that may affect the problems' parameters and chose some constraints, that, of course, impact dramatically the range of possible results.
It is often the practical situation that provides the hint in choosing the optimal constraints, but it almost never provides enough information to reduce the large amount of options involved. A wide range of variants in methodologies and approaches to solve the registration problem has been thus proposed in the last decades: a quick glance to Google scholar reveals about 1200000 papers in *medical image registration* (55% of the whole *image registration* resources)[1].

1.1.1 Examples of applications

One of the main application of image registration is in the domain of brain imaging: it can be used to examine differences between subjects and distinguish biological features between subjects (cross-sectional studies) or to compare different acquisition of the same subject, before and after a surgery or after a fixed period of time (longitudinal studies). In both

[1] Surveys in medical image registration can be found in [SDP13], [ZF03].

cases an accurate comparison between images and the parameters of the transformation involved, result in a quantification of anatomical variability and in a better understanding of the patients' features. For example, brain atrophy is considered a biomarker to diagnose the Alzheimer disease and to analyze its evolution; most of the algorithms and techniques involved in the atrophy measurement requires longitudinal or cross-sectional scans to be aligned, and so are directly affected by the solution of the registration algorithm [FF97], [GWRNJ12], [PCL+15].
Also when dealing with motion correction, if the sequence of images is affected by the motion of cardiac pulses or respiratory cycles, registration algorithms are often used for the realignment. For example, in lungs radiotherapy, lungs' motion is directly computed using a registration algorithm, and it is related with the respiratory surrogate signal. The correspondence model is then used to direct the X-ray or electrons beam on the cancerous tissue, minimizing the damage on the unaffected cells [MHSK], [MHM+11].
Another application is to compose together several images to obtain a bigger picture (mosaicing): in this case image registration is used to align images in the overlapping regions [VPM+06], [Sze94].

1.1.2 State of the Art

In the attempt to classify image registration algorithms, the most relevant feature that distinguish them is the choice of the family to whom the transformation belongs and its parametrization. Since anatomies are in a continuous process of modification over time, in general without any variation in the topological features, the use of diffeomorphisms to model transformations of organs appears as one of the most natural. Not accidentally diffeomorphisms are as well an important class of solutions of partial and ordinary differential equation aimed to model the dynamics of fluids.
In the development of diffeomorphic image registration, we can broadly identify some steps that led to the concept of log-composition presented in this research:

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

1998-2004 ▷ Based on the concept of attraction, the Demons algorithm [Thi98], [PCA99] propose the computation of the transformation between images in an iterative framework, where the update of the transformation at each step is parametrized with a vector field that minimize at each step an energy function. This vector field is defined as the set of vectors (demons) that moves one image into the other.
Here diffeomorphisms are not directly involved and the vectors at each voxel are considered as independent elements. In the same year of [Thi98], the set of diffeomorphism was taken into account in image matching and computational anatomy, not only as the set of solution of some family of differential equations, but with its Lie group structure [DGM98, Tro98, GM98].

2005-2006 ▷ The almost concomitant appearance of the Large Deformation Diffeomorphic Metric Mapping (LDDMM) [BMTY05] and the further investigation on the tangent space to the Lie group of diffeomorphisms as the space where to perform statistics (the so called log-euclidean framework) [ACPA06, AFPA06] keep the valuable approach in using diffeomorphisms as Lie group and to consider them with their Lie algebra, to model the little deformation in the tangent space as well as to rely on this normed space in computing the distance between transformations.

2007-2013 ▷ The LDDMM revealed all the opportunities provided by differential geometry in considering Lie group and Lie algebra embedded in a framework for the computation of image registration. In this setting, the tangent vector field comes from the solution of the ODE that models the transformations and it consists of the set of the non-stationary vector field (also time varying vector field or TVVF). After the log-euclidean framework [ACPA06] aimed at the computation of statistics of diffeomorphisms, the set of tangent vector field is restricted to the time-independent vector field (also stationary velocity field or SFV); the same restriction was subsequently considered in some further improvements of LDDMM (DARTEL, Stationary LDDMM [Ash07], [HBO07]). Log-euclidean framework brought new life also to the Demons algorithm, that become, in 2007, the diffeomorphic demons (or log-demons) [VPPA07][2]. Subsequent approaches that involves the symmetrization of the energy function and a different norm (local correlation coefficient instead of L^2) are proposed in symmetric diffeomorphic demons [VPPA08] and LCC-demons [LAF$^+$13] respectively.

1.1.3 Using Diffeomorphisms: Utility and Liability

If the set of transformations is bonded to the rigid body transformations group $SE(3)$ (mostly utilized in robotics and classical mechanics), the registration algorithm will be adequate to align 3d images. In clinical practice, brain images are aligned using rigid registration to compensate the motion, or to compare differences in longitudinal and cross sectional scans. If we assume that the motion of internal anatomy is continuous, then the group $SE(3)$ is not versatile enough to transform images. The choice goes on the Lie group of diffeomorphism Diff, that appears particularly appealing in computational anatomy since their topology-preserving nature. Their mathematical formalization as Lie group, is on the other hand not of immediate understanding, and it is still an open field of research.
Attempt to provide this object some handles for easy manipulation was done for the first time in 1966 by Vladimir Arnold [Arn66] [3]: to solve differential equation in hydrodynamic Diff 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. Subsequent steps in the exploration of the set of diffeomorphisms as a Lie group are [MA70, EM70, Omo70, Les83]. A survey on early development of infinite dimensional Lie group can be found in [Mil84], while more recent results and applications on diffeomorphisms has been published in [OKC92, BHM10, Sch10, BBHM11].

Consider Diff as a differentiable manifold involves the idea of having 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 [KW08]. 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. Here, 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

[2] An accurate comparison between stationary LDDMM and Diffeomorphic Demons with emphasis in both theoretical and practical aspects can be found in [HOP08].

[3] With a more readable sequel [Arn98] for non-French speakers.

do not holds anymore.
These difficulties led some researchers in approaching the set of diffeomorphisms from other perspectives: for example, instead of treating Diff as a group equipped with differential structures it is seen as a quotient of other well behaved group [Woj94].
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 more to the intuition than to a robust formalization. We work then mostly on finite dimensional settings, relying on important theorems and available close forms, and we will extend methods and results developed here in the infinite case -clearly - with proper precautions.

Another limitation that the reader should be aware of do not comes down to the theoretical difficulties of handling diffeomorphisms, but from the necessity of deal with discrete images and softwares. Two subset of some space have the same topology if exists an homeomorphism between them, but this analytical definition do not holds if the objects involved are considered in a discretized space. Separated subset remains separated until their distance is less than the size of a voxel for a significant region; if this happen, even with a homeomorphic underpinning model, the discretization process do not preserve the topology.

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 $\mathbf{x} \in \Omega_F$ the element $M(\varphi(\mathbf{x}))$ and $F(\mathbf{x})$ are as closed as possible according to a chosen measure of similarity. The function defined as the composition of the moving image after the transformation, $M \circ \varphi$, 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:

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

If $\Omega_F \neq \Omega_M$, it is always possible to apply an homeomorphism to transform them into a common domain Ω, called *background space*, on which both of the images are defined. In practice, this transformation between domains can be performed with a resampling technique, eventually with the well known artefacts related problem.

The definition of image registration here proposed, leaves two degrees of freedom in searching for a solution: the transformation's domain (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*: we define it as an iterative process that, at each step provides a new function φ that approaches f to the identity. Each iteration involves the optimization of a function that measure the similarity between the fixed image and the warped image computed at the previous step. To refine the energy function, the metric can be considered with an additive regularization term, that introduces a constraint based on prior knowledge about the searched solution:

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

where Sim is a function that measure the similarity, while Reg is a regularization term. The optimization algorithm on which the framework is based and the resampling strategy - process of resize the image from one dimension to another - provide alternative options to define the registration algorithm based on this framework.

1.2.2 Iterative Registration Algorithm

The definition of registration problem and the iterative framework described above, 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 to 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 features that distinguish one registration algorithm from the other.
The image registration framework here presented can be see as a electronic device with 5 knobs, each with its range:

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

Under the hood of this ideal device we may see an engine that can be schematically represented as in figure 1.1.
Modulating on the value of each knob, changing for example the set of transformation or the resampling technique, we change between the possible registration algorithms that generally falls in this framework.

Far from being a complete overview of all of the possible frameworks, it 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

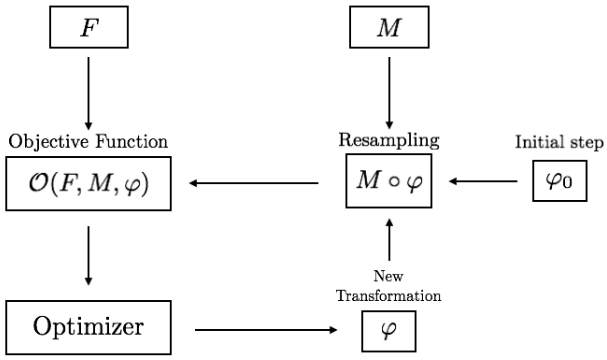


Figure 1.1: Image registration framework scheme.

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 subsections we will going from the generalized framework to some specific algorithms. We are interestedin particular in the parametrization of the *diffeomorphisms*[4] of the latest important algorithms: the LDDMM and the diffeomorphic Demons.

1.2.3 LDDMM: Classic, Shooting and Stationary

As previously done in the elastic registration [Bro81], the LDDMM framework [BMTY05] originates by considering motion between images as the motion of a fluid, and utilizes ODE from fluid dynamics to compute the deformation between reference and floating. Although, the discretized vector fields that can be stored in the computers' memory, are 5-dimensional matrix on a grid G over the background space Ω. The standard structure, used by most of the available software for image manipulation (Nibabel, py-NIfTI, niftilib, ...) is

$$M = M(x_i, y_j, z_k, t, d) \qquad (i, j, k) \in G, \quad t \in T \quad d = 1, 2, 3$$

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

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

Consider the set of homeomorphisms $\mathrm{Hom}(\Omega)$ (continuous function from the background space Ω to itself with continuous inverse) that act[5] on the set of images from the background

[4]bijective differentiable maps with differentiable inverse; take the real valued $f(x) = x^3$: it is bijective differentiable but the inverse is not everywhere differentiable.

[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. For definitions on group's action and orbits, we refer to [Art11]. For modern definitions of push-forward, pull-back and an introduction on Differential Geometry (in a non-Riemannian settings) [Lee12].

space $\mathcal{I}_\Omega$:

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

its orbit, given an image F and a subgroup of the homeomorphisms $\mathbb{G} \subset \mathrm{Hom}(\Omega)$, consists in the set of the images having the same topology of F:

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

The similarity term in the LDDMM is the L^2 norm[6] between the moving image and the fixed image in the same orbit:

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

while the regularization term that provides the optimal φ at each step is defined on the norm of the velocity vector field tangent to the transformation. Limiting its length means impose a constrain on the speed of the transformation at each step.
Let $\mathrm{Vect}(\Omega)$ be the set of all of the vector field over Ω. A generic time varying vector field (TVVF) is the continuously differentiable map defined as

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

Once initial conditions are given, at each TVVF, corresponds a unique time varying (or non-stationary) homomorphisms defined by the following ODE

$$\frac{d\phi_t(\mathbf{x})}{dt} = V^{(t)}(\phi_t(\mathbf{x})) \tag{1.2}$$

where

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

The transformation φ between fixed and moving images ($F \circ \varphi^{-1} = M$), can be defined by the couple $(V^{(t)}, \phi_t)$, such that, for $t = 0$, $\phi_t = Id$, identity of the group of homeomorphisms, and for $t = 1$, $\phi_t = \varphi$ is the sought homomorphism.

$$\varphi := \phi_1 = \phi_0 + \int_0^1 V^{(t)}(\phi)dt$$

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 minimize the distance between transformations[7]:

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

[6] For an introduction on the L^2 norm in the continuous we refer to chapter 4 of [SS09].

[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 (by a Levi-Civita connection so to be compatible with the shortest length of the integral curve), the consequent 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.

Ending points of path on the set of diffeomorphisms, whose tangent vector field (that varies over time), are used as regularization term:

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

Were L is a linear operator that can be dependent on some parameters that makes the approach even more general; it is defined as $L = (\alpha \vec{\nabla}^2 + \gamma)$ for α and γ real parameters and $\vec{\nabla}^2$ the Laplace operator. From the differential equation 1.2, and in consequence of the definition of φ the energy function 1.1 become

$$\mathcal{E}(F, M, \varphi) = \int_0^1 ||LV^{(t)}||^2_{L^2} dt + \frac{1}{\sigma^2} ||F(\varphi^{-1}) - M||^2_{L^2}$$

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

$$\hat{V} = \underset{V^{(t)} \,:\, \dot{\phi}_t(\mathbf{x}) = V^{(t)}(\mathbf{x})}{\text{argmin}} \int_0^1 ||LV^{(t)}||^2_{L^2} dt + \frac{1}{\sigma^2} ||F(\varphi^{-1}) - M||^2_{L^2}$$

Each transformation involved in the optimization algorithm are discretized time varying velocity fields; the update at each step is given by

$$\mathbf{v}_{k+1} = \mathbf{v}_k - \epsilon \nabla(\Delta \mathcal{E})$$

where $\mathbf{v}^k$ is the k-th step of the approximation of the velocity vector field V, $\Delta \mathcal{E}$ is the discretized version of the energy function and ϵ is the gradient descent step size.

A direct upgrade of the classical LDDMM performs the optimization on the geodesic flows, defined by a set of Hamiltonian equation (Shooting LDDMM [VRRC12]).
In this algorithm the iterative evolution of the deformation field, solution of the optimization algorithm, is regularized with the constraint imposed by an additional scalar field called *initial momentum*. As proved by the authors, the evaluation of this constraint at each step provides geodesics flows of homeomorphisms, but it is computationally expensive. Using the log-euclidean framework presented in [AFPA06], the algorithm proposed in [Ash07] uses a constraint based on the stationarity of the involved velocity field. Instead of considering time varying velocity fields constrained by a set of Hamiltonian equations, the domain of vector field is reduced to the stationary, which is an -almost- equivalent constraint, that considerably reduces the computational complexity. Resulting algorithm, the DARTEL (Diffeomorphic Anatomical Registration using Exponentiated Lie Algebra), was published in contemporary with [HBO07] based on the same concept of the parametrization of geodesics path of diffeomorphisms with stationary velocity fields, and with [VPPA07] that uses SVF on the Demons framework instead on the LDDMM.

1.2.4 Demonology: Classic, Pasha, Diffeomorphic, Symmetric and LCC

The first demons-based algorithm in image registration was proposed by [Thi98] in analogy with the Maxwell's demon in thermodynamics. This early version, called classic (or additive) demons, do not involves diffeomorphisms; the floating image is deformed with a vector field resulting from the computation of the optical flow[8] regularized by a gaussian filter at each

[8]Previously introduced for small deformation of optical sequence of images [HS81].

step. The optical flow is based on the idea that a voxel in the moving image is attracted by some force to all the points in the fixed with similar intensity, therefore it works under the hypothesis that the intensity of a moving object is constant over time.
Let $\mathbf{x}$ be a point in the background space Ω, the unknown vector field $V : \Omega \to \mathbb{R}^d$ is the function that at each point satisfies:

$$V(\mathbf{x}) \cdot \vec{\nabla} F(\mathbf{x}) = M(\mathbf{x}) - F(\mathbf{x}) \tag{1.3}$$

Thus it can be defined pointwise as

$$V(\mathbf{x}) = \frac{(M(\mathbf{x}) - F(\mathbf{x}))\vec{\nabla} F(\mathbf{x})}{||\vec{\nabla} F(\mathbf{x})||^2} \qquad \forall \mathbf{x} \in \Omega \tag{1.4}$$

To reduce the instability for small values of $||\vec{\nabla} F||$ and to satisfy the condition $V \to \mathbf{0}$ for little $\vec{\nabla} F$, the equation 1.4 is multiplied by $||\vec{\nabla} F(\mathbf{x})||^2/(||\vec{\nabla} F(\mathbf{x})||^2 + (M(\mathbf{x}) - F(\mathbf{x}))^2)$:

$$V(\mathbf{x}) = \frac{(M(\mathbf{x}) - F(\mathbf{x}))\vec{\nabla} F(\mathbf{x})}{||\vec{\nabla} F(\mathbf{x})||^2 + (M(\mathbf{x}) - F(\mathbf{x}))^2} \qquad \forall \mathbf{x} \in \Omega \tag{1.5}$$

Therefore, the classical demons algorithm defines the update of the transformation field at each step with 1.5 or with $V(\mathbf{x}) = \mathbf{0}$ when

$$||\vec{\nabla} F(\mathbf{x})||^2 + (M(\mathbf{x}) - F(\mathbf{x}))^2 < \epsilon$$

for ϵ spacing of floating point number.

In the paper that presents the PAHSA demons [CBD$^+$03], the authors underline the fact that the scalar product 1.3 underpinning the classic demon consists in the minimization of a local energy function, one per volxel. They reformulate the algorithm using a global energy function, aimed to makes the method easier to be analyzed and compared with others. With some modification, the Classic demons algorithm is accompanied back to the framework presented in the previous section, having as energy function

$$\mathcal{E}(F, M, C, V) = \frac{1}{\sigma_s}\mathrm{Sim}(F, M, C) + \frac{1}{\sigma_d}\mathrm{dist}(C, V) + \frac{1}{\sigma_r}\mathrm{Reg}(V) \tag{1.6}$$

where C and V are both vector fields defined on the background space Ω and each σ accounts for the uncertainty of each of the involved measure.
It differs from the energy function 1.1 because of the additional *hidden variable* C, whose purpose is to use alternating minimization algorithm[9] to optimize $\mathcal{E}$. Functions of involved in this new version of the energy function can be defined as

$$\begin{aligned} \mathrm{Sim}(F, M, C) &= ||F - M \circ C||^2 \\ \mathrm{dist}(C, V) &= ||C - V|| \\ \mathrm{Reg}(V) &= ||\vec{\nabla} V||^2 \end{aligned}$$

[9] The optimization algorithm is divided in two phase: at the step t, in the first phase the first two addend of the energy function are optimized to obtain the approximation of C, $\mathbf{c}_t$ holding the variable V to $\mathbf{v}_{t-1}$, approximation obtained from the previous step. In the second phase, the last two addend of the energy function are optimized to obtain the approximation of V for C keept constant at $\mathbf{c}_t$ obtained in the previous phase.

or with more advanced criteria proposed in [CBD+03]. In this context an equivalent version of the Classical demon, involving the global optimization function,can be reformulated from the preovious equations, keeping $C = V$:

$$\begin{aligned}
\mathrm{Sim}(F, M, V) &= ||F - M \circ V||^2 \\
\mathrm{dist}(C, V) &= 0 \\
\mathrm{Reg}(V) &= ||\vec{\nabla} V||^2
\end{aligned}$$

and applying a Gaussian smoother at each step to regularize the solution.
The update of the vector field obtained with the optimization of the energy function and the Gaussian smoother is summed at each step with the previously computed vector field[10]. Let $\{\mathbf{v}_j\}$ be the vector fields' sequence approximating V and $\delta\mathbf{v}$ the one computed at the iteration $j + 1$ by the optimization algorithm, then the *update* is computed with

$$\mathbf{v}_{j+1} := \mathbf{v}_j + \delta\mathbf{v} \tag{1.7}$$

Again the PASHA algorithm do not involves any diffeomorphism. This development was made in [VPM+06], with the diffeomorphic demons (or log-demon). Within the log-euclidean framework the involved vector fields are element in the tangent space of the group of diffeomorphisms. As in LDDMM, to each vector field $V \in \mathrm{Vect}(\Omega)$ is associated a diffeomorphisms p by the ODE $dp/dt = V^{(t)}(p)$.
In this settings, the update can not be computed simply with a sum, since it must reflect the composition of diffeomorphisms that defines the corresponding vector fields. Intuitively if $dp/dt = V^{(t)}(p)$ and $dq/dt = W^{(t)}(q)$ do not implies $dp \circ q/dt = V^{(t)}(p) + W^{(t)}(q)$. To define the composition in the tangent space, we need to use two functions to transform vector fields into diffeomorphisms and vice-versa; they are already provided by Lie group theory. The *Lie algebra*, usually denoted with $\mathfrak{g}$, contains the set of the tangent vector fields, while the Lie group, denoted with $\mathbb{G}$, contains the set of transformations; the function *Lie exponential* maps vector fields on the corresponding Lie group elements while the *Lie logarithm*, its inverse under some condition defined in chapter 2, maps each diffeomorphisms in the correspondent tangent vector field:

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

The addition of tangent vector fields 1.7 that defines the update of the additive demons, becomes, in the diffeomorphic demon, the logarithm of the composition of the two diffeomorphisms that corresponds to the vector fields in the Lie group:

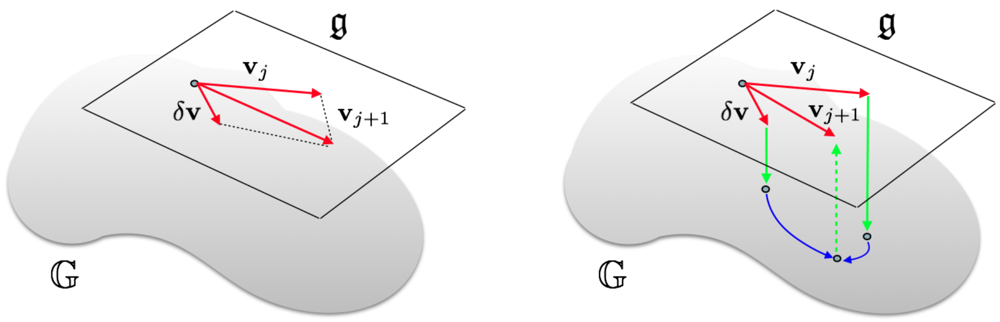
$$\mathbf{v}_{j+1} = \log(\exp(\mathbf{v}_j) \circ \exp(\delta\mathbf{v})) \tag{1.8}$$

And again the theory of Lie group provides a formula to compute this composition: the BCH formula. This involves an infinite series of nested Lie bracket, that do not makes its computation a straightforward task:

$$BCH(\mathbf{v}_j, \delta\mathbf{v}) := \mathbf{v}_j + \delta\mathbf{v} + \frac{1}{2}[\mathbf{v}_j, \delta\mathbf{v}] + \frac{1}{12}([\mathbf{v}_j, [\mathbf{v}_j, \delta\mathbf{v}]] + [\delta\mathbf{v}, [\delta\mathbf{v}, \mathbf{v}_j]]) + ...$$

Remarkably, if we set to 0 all the lie brackets in the BCH formula (so we neglect the curvature of the space), the resulting composition of vector field in the log domain will coincide with 1.7:

$$BCH(\mathbf{v}_j, \delta\mathbf{v}) \simeq \mathbf{v}_j + \delta\mathbf{v}$$



$$\mathbf{v}_{j+1} = \mathbf{v}_j + \delta\mathbf{v} \qquad\qquad \mathbf{v}_{j+1} = \log(\exp(\mathbf{v}_j) \circ \exp(\delta\mathbf{v}))$$

Figure 1.2: Each update computed in the additive demons algorithm is the sum of vector fields (left). In diffeomorphic demons algorithm vector fields are considered as elements of the lie algebra $\mathfrak{g}$, tangent space of the Lie group of diffeomorphisms. The update is compute as the composition of the correspondent transformations in the Lie group $\mathbb{G}$ using the canonical transformations exp and log (right).

An extension of the diffeomorphic demons, that involves the optimization of the energy function $\mathcal{E}(F, M, V)$, exploits the existence of the inverse transformation and optimizes instead a symmetrised version of the energy function [VPPA07]. If $V = log(p)$, then the *symmetric diffeomoprhic demons algorithm* finds at each step

$$\tilde{V}_{t+1} = \underset{p \in \mathbb{G}}{\operatorname{argmin}} \, \mathcal{E}(F, M, log(p)) + \mathcal{E}(F, M, \log(p^{-1}))$$

A further development of the diffeomoprhic demons algorithm, presented in [LAF$^+$13], preserve the structure of the symmetric version, but uses the LCC measure of similarity instead of L^2 to compute the energy function.

Last section of this chapter is devoted to introduce the intuition of this concept, while chapter 2 is devoted to the introduction the formal definitions and to its development.

1.3 The Composition of SVF in the Diffeomorphisms Group

Every non-rigid registration algorithm requires to be implemented and to work with discretized images. The nature of the computers' memory prevent from the possibility of storing the continuous fluid transformations that solves the differential equations of the LDDMM and the diffeomorphic demons approach: the only thing that we can do, unfortunately, is to store the discretized vector fields and resampling them with the images using one of the available techniques.
When relying on diffeomorphisms, we still have to consider discretized vector fields as input and output, but we need the Lie group of diffeomorphisms as support to compute the

[10] This is the reason why in demonology, the classical demons is sometime classified as additive demon.

composition 1.8 from the tangent space. This operation is baptized under the name of *log-composition* and it is defined as

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

The main aim of this research is to find and compare numerical ways to compute it.

A fast log-composition is not useful only for the diffeomorphic demons. It can be used to solve every problem that uses Lie groups and need to be implemented in a computer. In medical imaging it can be used for

1. Diffeomorphic demon, and symmetric diffeomorphic demon [VPM$^+$06, VPPA08].
2. Fast computation of the logarithm [BO08].
3. Calculus on diffusion tensor [AFPA06].
4. Compute the discrete ladder for Parallel Transport in Transformation Groups [LP14].

The next chapter is devoted to the formal definition of the log-composition, underpinned with the tools from differential geometry theory and to present the numerical technique developed so far to compute it.

Chapter 2

Tools from Differential Geometry

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

2.1 A Lie Group Structure for the Set of Transformation

We consider every group $\mathbb{G}$ as a group of transformations acting on $\mathbb{R}^d$, having in mind the particular case $d = 2, 3$ for 2-dimensional or 3-dimensional images. We will focus 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 heart 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 turns out to be

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

Of course for different choice of γ passing through p, we obtain different tangent vectors:

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

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.
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 $(C; x_1, \ldots, x_n) = (C, \psi)$ 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 $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$. 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 $\mathrm{Vect}(M)$, is a real vector space and a module over $\mathcal{C}^\infty(M)$:

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

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

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

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:

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

Let V be 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

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

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 :

$$\begin{cases} \frac{dc_i(t)}{dt} = v_i(c_1, c_2, \dots, c_n) \\ c_i(0) = p_i \end{cases} \tag{2.1}$$

Thanks to the Cauchy theorem, it has an 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 led to the concept of *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: we say in general that 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. Finally, 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:

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

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

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

The last definition and the one parameter subgroup property, in conjunction with 2.1 will be largely used when dealing with Lie exponentials and Lie logarithms, and are at the core of the LDDMM and subsequent framework (see equation 1.2).

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

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

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 $\mathcal{C}^\infty(\mathbb{H})$, but it is not easy to compute in practical applications defined in this way. 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. We are particularly interested when $\mathbb{G} = \mathbb{H}$, since the push-forward become a way to make the vector field V defined at some point p of the manifold, to act on the point $F(p)$ of the same manifold toward the function F.
The *pull-back*, is defined on the dual space of $\mathbb{G}$ and $\mathbb{H}$ as the contravariant operator of the

[1] Less important in this context: the name comes down to group theory where the action of the group $(\mathbb{R}, +)$ over the manifold has as its orbits the set of disjoints integral curves.

push forward[2]:

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

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

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

Restricting our attention to the case $\mathbb{G} = \mathbb{H}$, the map F becomes a smooth movement of the points of $\mathbb{G}$. Each element p of a Lie group $\mathbb{G}$ can be considered as the seed of three particularly interesting maps over $\mathbb{G}$:

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*:

$$(R_p)_\star V_q f = V_q(f \circ R_p) = \sum_{i=1}^{n} v_i(q) \frac{\partial f \circ R_p}{\partial x_i} \Big|_q = \sum_{i=1}^{n} v_i(q) \frac{\partial f}{\partial x_i} \Big|_{qp}$$

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

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 an 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 $\mathrm{leftVect}(M)$. This can be easily proved by:

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

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 an 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 $\gamma_\mathbf{v}(t) = \Phi(t, e) = \varphi_e(t)$.

Using the features so far introduced it is possible to recover the Lie algebra of a Lie group. We remember these equivalent definitions:

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 $\mathbb{G}$: $\mathfrak{g} := \mathrm{leftVect}(\mathbb{G})$.
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.

We conclude this section considering the twin of the adjoint map for the Lie algebra of a Lie group:

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

(notations in this research are, as usual, case sensitive). One of its interesting property is that it preserves the Lie bracket: $\mathrm{ad}_{\mathbf{u}}[\mathbf{v}, \mathbf{w}] = [\mathrm{ad}_{\mathbf{u}}\mathbf{v}, \mathrm{ad}_{\mathbf{u}}\mathbf{w}]$.

2.1.3 Lie Exponential, Lie logarithm and Lie Log-composition

Let $\mathbf{v}$ be an element in the tangent space $\mathfrak{g}$ and $V \in \mathrm{leftVect}(\mathbb{G})$ 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:

[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.

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 in the Lie algebra we have $\exp(\mathrm{Ad}_{\mathbf{u}}\mathbf{v}) = \mathrm{Ad}_{\mathbf{u}}\exp(\mathbf{v})$

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 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:
$$\begin{aligned} \log : \mathbb{S}^1 &\longrightarrow T_0\mathbb{S}^1 & \exp : T_0\mathbb{S}^1 &\longrightarrow \mathbb{S}^1 \\ \exp(i\alpha) &\longmapsto \log(\exp(i\alpha)) = i\alpha & i\alpha &\longmapsto \exp(i\alpha) \end{aligned}$$
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
$$||\log(I+\mathbf{c}) - \mathbf{c}|| \leq \alpha||\mathbf{c}||^2$$

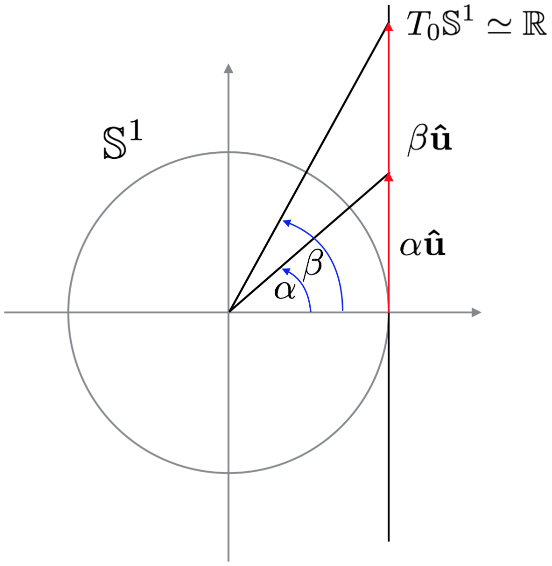


Figure 2.1: Lie algebra of the Lie group of plane's rotation.

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.

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.

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:

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

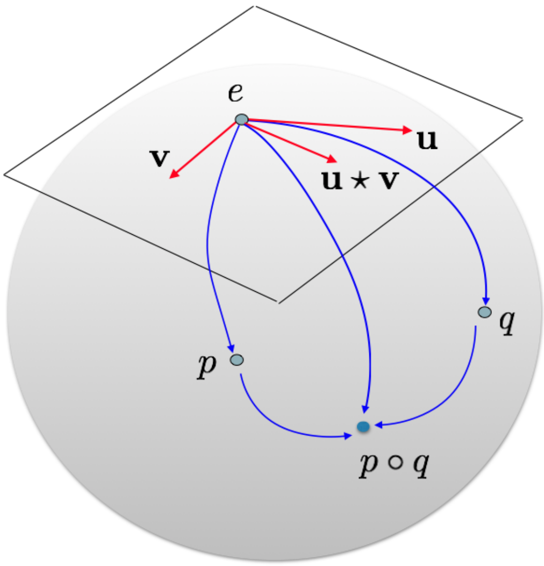


Figure 2.2: graphical visualization of the Lie log-composition.

Following properties holds for the Lie log-composition:

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

   (a) $(\mathbf{u}_1 \oplus \mathbf{u}_2) \oplus \mathbf{u}_3 = \mathbf{u}_1 \oplus (\mathbf{u}_2 \oplus \mathbf{u}_3)$ for all $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ in $C_{\mathfrak{g}}$.
   (b) $\mathbf{u} \oplus \mathbf{0} = \mathbf{0} \oplus \mathbf{u} = \mathbf{u}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.
   (c) $\mathbf{u} \oplus (-\mathbf{u}) = \mathbf{0}$ for all $\mathbf{u}$ in $C_{\mathfrak{g}}$.

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

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

   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.

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

2.1.4 Connections and Geodesics

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

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

such that for all f, g in $\mathcal{C}^\infty(\mathbb{G})$ and for all V, W in $\mathrm{Vect}(\mathbb{G})$ 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 $V, W \in Left\mathrm{Vect}(\mathbb{G})$ their connection $\nabla_U V$ is still left invariant.

2.1.5 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:

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

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 $\tilde{\oplus}$ over the $\mathbb{G}$ fiber bundle such that

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

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

2.2 Parallel Transport

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 Initial Definition and Some 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!

2.2.2 Introductory Properties

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:

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

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

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.

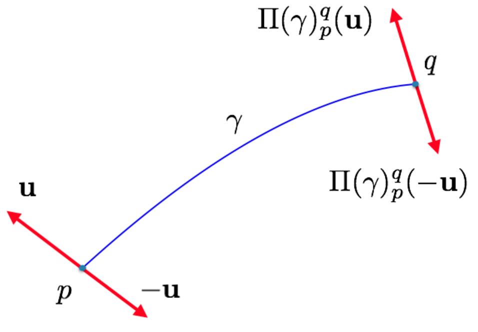


Figure 2.3: First inversion property.

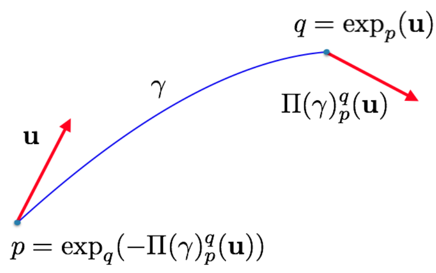


Figure 2.4: Second inversion property.

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 $\beta(-1)$.

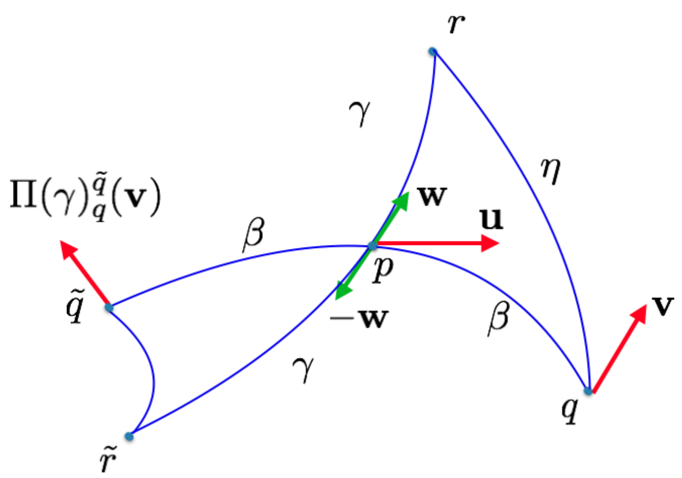


Figure 2.5: Change of sign property.

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

2.2.3 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!!!

2.2.4 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!

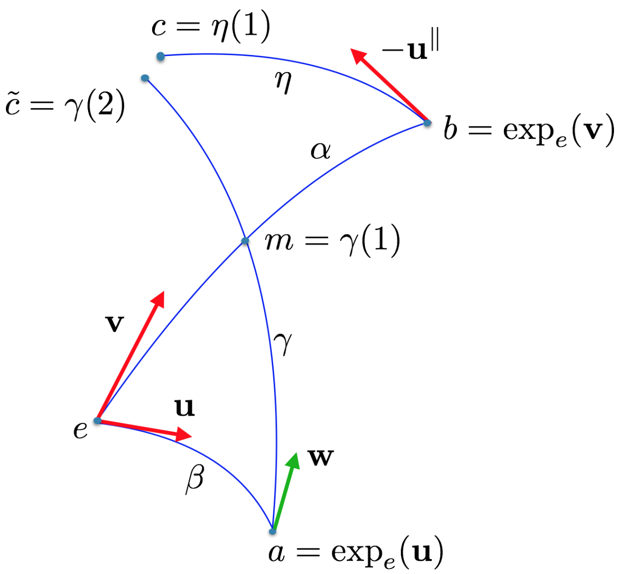


Figure 2.6: Pole ladder applied to parallel transport.

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:

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

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:

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

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

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

2.4 Four Strategies for the Computation of Log-composition

In this section we provide explicit formulas for the computation of the log composition, using the tools introduced in the previous sections.

2.4.1 Truncated BCH formula

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.

2.4.2 Taylor Expansion

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

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

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

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

xxx intermediate passages to be written from zachos, blane!

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

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

Where $\{B_n\}$ is the sequence of the second-kind Bernoulli number[4].

2.4.3 Parallel Transport

2.4.4 Accelerating Convergences

[4] 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.

Chapter 3

Numerical computation of the Log-composition for SE(2) and SVF

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*

3.1 The Group of Rigid Body Transformations

In the previous chapter we have introduced some fundamental tools from differential geometry commonly utilized in computational anatomy. The core object of the theory, exponential and logarithmic map, as well as pole ladder, are obviously different for each manifold and each metric, if there is one: practical implementations have to be determined case by case. This chapter is aimed to go through the details of the generalized theory for the cases of SE(2) and the Lie group of diffeomorphisms parametrized with stationary velocity fields.

3.1.1 Lie Logarithm, Lie Exponential and Log-compositions for SE(2)

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, and forms the euclidean group $E(2)$. For 2d rigid registration usually reflections are not required and so we restrict our attention to the special euclidean group $SE(2)$. We are interested in two things about them: their expression in matrix form, and the Lie group and the Lie algebra structures involved.
We denote denoted with $M_3(\mathbb{R})$ the set of all of the 3×3 matrices with real entries . 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 proper or improper subgroup of $GL_3(\mathbb{R})$. The group of 2d 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:

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

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:

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

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:

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

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.
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:

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

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:

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

These are still valid for its restriction

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

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.

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

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

3.1.2 Numerical Computations of the Log-composition for SE(n)

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].

Parallel Transport to compute the Log-composition

Log and Exp Approximations for Small 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!

Truncated BCH formula

Taylor expansion

Parallel transport

Accelerating convergences

3.2 The Set of Stationary Velocity Fields

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 -

[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.

3.2.1 Why Only Set

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.

– The starting point is in the difference between steady state and time dependent, at the core of the definition of the diffeomorphisms as solution to a class of ODE

Finite Difference Methods for Ordinary and Partial Differential Equations

3.2.2 Some Tools for the Infinite Dimensional Case

It can be proved that the Lie algebra of 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 Lie(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 $\mathrm{Diff}(\mathbb{R}^d)$.

Proof. Investigate a proof to define δ. □

Lemma 3.2.2 (identity lemma)**.** Be $p \in \mathrm{Diff}(\mathbb{R}^d)$, 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 $\mathrm{Diff}(\Omega)$. 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 $\text{Diff}(\Omega)$ 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 Lie Logarithm and Lie Exponential and Log-composition for SVF

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!

3.2.4 Numerical Computations of the Log-composition for SVF

Truncated BCH formula

Taylor expansion

Parallel transport

Accelerating convergences

Chapter 4

Log-computation Algorithm using Log-composition

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.

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.

Before we need to introduce the space of the approximation of the Lie algebra and the Lie group.

4.1 Space of approximations

xxx Space of the approximations, in connection with the small rotation chapter.
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-composition $\mathbf{v}_1 \oplus \mathbf{v}_2 = \log(\exp(\mathbf{v}_1) \circ \exp(\mathbf{v}_2))$ we define two approximating functions:

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

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

Where $\mathfrak{g}^{\sim}$ is 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

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

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

lead us to consider the algorithm presented in [BO08], here called log-computation, with a new reformulation.

4.2 The Log-computation Algorithm using Log-composition

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

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

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

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

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

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

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

4.2.1 BCH Strategy

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

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

thus, for the first degree we have

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

For the second degree we have:

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

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

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

4.2.2 Parallel Transport Strategy

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

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

4.2.3 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

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

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!)

Chapter 5

Experimental Results

There, I've done my best! If this won't suit I shall have to wait till I can do better.
-Jo

zz write chapter intro, how we did it!

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).

5.1 Group composition at the Service of Image Registration

xxx Sum up: log-composition in diffeomorphic image registration.

5.2 Experimental Results

5.2.1 Log-Compositions Performance for Toy examples and Patient Images

5.2.2 Logarithm Computation using Log-composition

5.3 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

[ACPA06] 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.

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

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

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

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

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

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

[BHO07] Matias Bossa, Monica Hernandez, and Salvador Olmos. Contributions to 3D diffeomorphic atlas estimation: application to brain images. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 10:667–674, 2007.

[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.

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

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

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

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

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

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

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

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

[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.

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

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

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

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

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

[LP14] Marco Lorenzi and Xavier Pennec. Discrete Ladders for Parallel Transport in Transformation Groups with an Affine Connection Structure. In Frank Nielsen, editor, *Geometric Theory of Information*, Signals and Communication Technology, pages 243–271. Springer, 2014.

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

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

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

[Mil84] J Milnor. Remarks on infinite-dimensional lie groups, in 'relativity, groups and topology, ii'(les houches, 1983), 1007–1057, 1984.

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

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

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

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

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

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

[SS09] Elias M Stein and Rami Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces*. Princeton University Press, 2009.

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

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

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

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

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

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

[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.

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