

Ph.D. Thesis

Computational anatomy and brain morphometry



Ernesto Zacur

Instituto Universitario de Investigación
en Ingeniería de Aragón
Universidad Zaragoza

Supervisor: Salvador Olmos Gassó

June 2014



**Instituto Universitario de Investigación
en Ingeniería de Aragón
Universidad Zaragoza**

Ph.D. Thesis

Computational anatomy and brain morphometry

Anatomía computacional y morfometría cerebral

Ernesto Zacur

Supervisor:
Salvador Olmos Gassó

June 2014

Gracias a Anita, para Anita.

Abstract

The main objective of this thesis is to develop a mathematical framework for the statistical analysis of spatial transformations. With the basis on Lie groups and differential geometry, the set of spatial transformations is endowed with a metric structure. Although standard statistical tools are developed for elements belonging to a vector space, they cannot be directly used on spatial transformations as in general they do not belong to a vectorial space. On the other hand, there exist several extensions of classical statistical tools to analyze elements belonging to sets with a simpler structure, such a metric space. In this thesis it is studied how to endow with a Riemannian metric structure to the set of spatial transformations, which are considered as elements of a Lie group and, therefore, have an associated differentiable manifold structure. Special relevance is given to invariant metrics because they provide symmetries under either a change of coordinates or the use of different measurement devices. Several methods are given and analyzed to compute geodesics in the space of spatial transformations. Once geodesics are available, distances between spatial transformations can be straightforwardly defined as the length of the shortest geodesic between them.

Several applications of interest in computer vision and medical image analysis are given as illustrative examples of the proposed methodologies. Special emphasis is given to the problem of Computational Anatomy and brain morphometry, motivated by the need of finding anatomical changes in the brain induced by Alzheimer's disease. In this field, particular contributions are geodesic regression of pose information from subcortical nuclei as well as Tensor-based morphometry (TBM) using a multivariate description of the full Jacobian matrix of the deformation field.

Even though the thesis is focused on the development of tools for the morphometric study of brain structures, the proposed methodologies are also of great interest in other areas: or for studies of other organs; or in other disciplines such as computer vision.

Resumen

El principal objetivo de esta tesis es desarrollar un marco formal para el estudio, cuantificación y análisis estadístico de transformaciones espaciales o geométricas. El marco propone dotar de una estructura métrica al espacio de las transformaciones espaciales basado en la geometría riemanniana y los grupos de Lie. Las herramientas de la estadística clásica están desarrolladas para su uso sobre elementos con estructura de espacio vectorial. Sin embargo las transformaciones espaciales, en general, no se pueden considerar como espacios vectoriales. Por otro lado, existen extensiones de la estadística clásica que permiten el análisis de elementos pertenecientes a espacios con estructuras más simples, como la de espacio métrico. Considerando las transformaciones como elementos de un grupo de Lie y haciendo uso de la variedad diferencial asociada, en esta tesis se estudian los pasos necesarios para dotar a estos grupos de una métrica riemanniana y ser capaces de calcular distancias entre transformaciones espaciales. Entre las métricas más relevantes se propone el uso de aquellas que resultan invariantes ante la acción de los elementos del grupo y que proporcionan simetrías ante cambios de coordenadas o sistemas de medición. Después de una revisión de estos conceptos, se estudian y proponen distintos métodos para calcular geodésicas en un espacio de transformaciones espaciales. Dotados de estas geodésicas el objetivo de calcular distancias entre transformaciones espaciales queda reducido a encontrar la geodésica más corta entre dos transformaciones.

El uso de estas herramientas de análisis se ejemplifica en algunas aplicaciones de interés en el área de visión por ordenador. Además, estudios más extensos en el área de imágenes médicas se han llevado a cabo para el análisis de cambios anatómicos producidos por la enfermedad de Alzheimer. En esta área se propone la extensión de estudios clásicos, basados en la volumetría de estructuras, a descriptores multivariados más completos de la anatomía de los individuos. También se estudia la técnica de Tensor-based morphometry (TBM) y se propone su extensión utilizando un descriptor multivariado de la deformación local.

Los descriptores multivariados propuestos se evalúan en un estudio de regresión con respecto a la escala cognitiva ADAS-cog y en un estudio de test de hipótesis entre un grupo de control y un grupo patológico. Debido a que las entidades matemáticas que se analizan no pertenecen a un espacio vectorial sino que a un grupo de transformaciones, para llevar a cabo estos estudios se emplean las metodologías aprendidas en la investigación de la geometría riemanniana.

Si bien la tesis está enfocada al desarrollo de herramientas para el estudio morfométrico de las estructuras cerebrales, las metodologías propuestas son también de gran interés en otras áreas: o bien para estudios de otros órganos; o bien en análisis propios de otras disciplinas como visión por ordenador.

Contents

Abstract	i
1 Introduction	1
1.1 Introduction	2
1.1.1 Computational anatomy	3
1.1.2 Neuroanatomy and neuroimaging	4
1.2 Tools from computational anatomy	7
1.2.1 Registration of anatomical structures	7
1.2.2 Geometrical description of the anatomy	7
1.2.3 Spatial transformations	8
1.2.4 Similarity measures and regularization	9
1.3 Mathematical bases for the statistical analysis of spatial transformations	11
1.3.1 Non-classical formulation of statistical tools	11
1.3.2 Spatial transformation sets with differentiable structure	12
1.3.3 Riemannian distances and exponential mappings	12
1.3.4 Invariance under spatial transformations	13
1.4 Morphometry in neuroimaging	14
1.4.1 Alzheimer's disease	14
1.4.2 The role of structural neuroimaging in AD	16
1.4.3 Data for clinical validation	19
1.5 Outline of the thesis	20
2 Background	23
2.1 Geometrical objects and spatial transformations	24
2.2 Action of spatial transformations on geometrical objects	26
2.2.1 Subsets of the ambient space	26
2.2.2 Scalar functions	28
Practical aspects for representation of images	28
Image gradients	30
2.2.3 Examples of actions which modify attribute values	31
2.2.4 Combined objects	32
2.2.5 Spatial transformations as geometrical objects	32
2.3 Parametrization of the action of spatial transformations	33
2.4 Spatial transformation groups	35
2.4.1 Matrix groups and matrix representations	35
2.4.2 The action of matrix groups on spatial points	37
2.5 Spatial transformation models	38
2.5.1 Discrete spatial transformations models	38
2.5.2 Finite dimensional models	38

2.5.3	Object-dependent transformation models	43
2.5.4	Infinite dimensional spatial transformation models	43
3	Image registration	49
3.1	Registration of geometrical objects	50
3.2	Similarity and dissimilarity measures between geometrical objects	51
3.2.1	Distances between landmarks	51
3.2.2	Distance between unlabeled points	52
3.2.3	Distance between points and curves	52
3.2.4	Measures of the overlapping between regions	52
3.2.5	Dissimilarity measures between scalar images	53
3.3	Regularity measures	54
3.4	Registration using parametric spatial transformation models	55
3.4.1	Registration of scalar functions	58
3.5	Registration using object-dependent spatial transformation models	63
3.5.1	Exact matching of landmarks	64
3.5.2	Inexact matching formulation	66
3.5.3	Inexact matching for registering scalar functions	68
3.6	Registration using non-parametric spatial transformation models	69
4	Left-invariant geodesics	75
4.1	Introduction	76
4.2	Background	78
4.2.1	Charts and manifolds, Lie groups, matrix groups	78
4.2.2	Tangent vectors, tangent space and vector fields	80
4.2.3	Translating elements, curves and velocities	80
4.2.4	Conjugation, Adjoint and adjoint	81
4.2.5	Group exponential	82
4.2.6	Riemannian metrics, lengths, geodesics and $\text{Exp}(\cdot)$	83
4.2.7	Isometries	85
4.2.8	Invariant Riemannian metrics	86
4.2.9	Totally geodesic subgroups	87
4.2.10	Bi-invariant metrics	87
4.2.11	Geodesics as “straightest” curves	88
4.3	Computing left-invariant Riemannian geodesics	92
4.3.1	Computing geodesic in coordinate charts	93
4.3.2	Solving geodesics in the group exponential chart	95
4.3.3	Evolution of geodesics described in the algebra	97
4.3.4	A coordinate-free solution based on (E-P)	100
4.3.5	Geodesics using Optimal Control	103
4.3.6	Symmetric representation of the geodesics evolution	105
4.4	Conserved quantities along geodesics	105
4.5	Reducing to a first-order equation	106
4.6	Sensitivity with respect to initial velocity	108
4.7	$\text{Exp}(\cdot)$ under different metrics	109
4.8	Expressions for the right-invariant metric case	110

5 Solutions for left-invariant geodesics and applications	113
5.1 Closed-form solutions of $\text{Exp}(\cdot)$	114
5.1.1 Translation group, $\mathcal{T}(d)$	114
5.1.2 Isotropic Scale group, $\mathcal{S}^+(1)$	115
5.1.3 Non-Isotropic Scale group, $\mathcal{S}^+(d)$	115
5.1.4 Isotropic Scale + Translation group, $\mathcal{ST}(d)$	115
5.1.5 Rotation group, $\mathcal{SO}(d)$	116
5.1.6 Special Euclidean group, $\mathcal{SE}(d)$	116
5.1.7 Similarity group, $\mathcal{SIM}(d)$	117
5.1.8 Orientation preserving General Linear group, $\mathcal{GL}^+(d)$	117
5.1.9 Special Linear group, $\mathcal{SL}(d)$	117
5.1.10 Centered Transformation group, $(\mathcal{G}(d) \times \mathcal{T}(d))$	118
5.1.11 Projective group, $\mathcal{PGL}(d)$	118
5.1.12 Möbius Transformation group, $\mathcal{MO}(d)$	119
5.2 Numerical solutions for the $\text{Exp}()$ function	121
5.2.1 Explicit additive Euler method	121
5.2.2 Explicit Lie–Euler method	122
5.2.3 Taylor series	122
5.2.4 Fixed step-size Taylor method	125
5.2.5 Adaptive Taylor method	125
5.2.6 Computing the Riemannian logarithm function	126
5.3 Results	128
5.3.1 Accuracy	128
5.3.2 Timing performance of Adaptive Taylor method	130
5.3.3 Departures from the group set	131
5.4 Application examples	131
5.4.1 Interpolating spatial transformations	133
5.4.2 Intrinsic sample mean of spatial transformations	136
5.4.3 Intrinsic filtering	140
5.4.4 Extension to any metric for $\mathcal{ST}(1)$ group	142
6 Applications to brain morphometry	145
6.1 Statistical analysis of <i>pose</i> information from subcortical nuclei	146
6.1.1 Subject and group selection	147
6.1.2 Subcortical nuclei delineation	148
6.1.3 Spatial transformation model and mean shape for each subcortical object	148
6.1.4 Pose feature and intrinsic mean	150
6.1.5 Statistical analysis of pose features. A regression study	152
6.1.6 Hypothesis testing and statistical assessment	157
6.2 Tensor-Based Morphometry	159
6.2.1 Subject groups and template construction	160
6.2.2 Tensor-based morphometry (TBM) quantification	161
6.2.3 Statistical analysis	164
6.2.4 Assessment of statistical significance	166
6.3 Multivariate tensor-based morphometry	170
6.3.1 Invariance with respect to the template	171
Jacobian determinant	173
Deformation tensor	173
A proper distance function on $\mathbf{GL}^+(d)$	174

Details for computing \mathbf{d}_{RI}	175
6.3.2 Cramér test	177
6.3.3 Synthetic study	178
6.3.4 Results on ADNI dataset	181
6.4 Discussion	182
7 Conclusion and future extensions	185
Publications derived from the thesis	193
Bibliography	195

CHAPTER

1

Introduction

1.1	Introduction	2
1.1.1	Computational anatomy	3
1.1.2	Neuroanatomy and neuroimaging	4
1.2	Tools from computational anatomy	7
1.2.1	Registration of anatomical structures	7
1.2.2	Geometrical description of the anatomy	7
1.2.3	Spatial transformations	8
1.2.4	Similarity measures and regularization	9
1.3	Mathematical bases for the statistical analysis of spatial transformations	11
1.3.1	Non-classical formulation of statistical tools	11
1.3.2	Spatial transformation sets with differentiable structure	12
1.3.3	Riemannian distances and exponential mappings	12
1.3.4	Invariance under spatial transformations	13
1.4	Morphometry in neuroimaging	14
1.4.1	Alzheimer's disease	14
1.4.2	The role of structural neuroimaging in AD	16
1.4.3	Data for clinical validation	19
1.5	Outline of the thesis	20

1.1 Introduction

Anatomy is a descriptive discipline from the biology which focus on the study of shapes or forms, changes of shapes, structure and distributions of the components of living organisms. Anatomical studies can be categorized into studies at the macroscopic level (topographic anatomy or gross anatomy) and microscopic anatomy (including histology and cytology) [Vargas 2002]. Topographic anatomy uses invasive and noninvasive tools in order to obtain information about the structures and their relationships at macroscopic level. Although anatomy dispenses the functional and physiological aspects of the organs, morphometric techniques (quantitative analysis of the shape) are often used for prognosis, diagnosis, differential diagnosis, treatment or monitoring of some diseases. In addition to these clinical purposes, morphometric techniques are also used in the medical science for basic research on diseases and for the development and assessment of drugs.

In the early days of medicine, topographic pathological anatomy (the study, through morphological techniques of the causes, development and consequences of diseases) was confined to superficial studies or to the use of dissection techniques to analyze *ex vivo* organs. This situation have changed radically in the late nineteenth century with the discovery of X rays [Röntgen 1895] [Doby 1997] [Bradley 2008] and its use to observe the interior of living organisms with minimally invasive techniques. From this key milestone in the history of the biomedical sciences to date a large number of imaging modalities for the observation and study of both the function and the structure of organisms have been developed. This gave rise to a new discipline called *medical imaging*. It makes use of various imaging techniques such as X-rays (RX), ultrasound (US), magnetic resonance (MR), computed tomography (CT), positron emission tomography (PET), single photon emission computed tomography (SPECT), angiography, endoscopy, thermography, microscopy, fluoroscopy, medical photography, among others. These techniques make possible the study of the anatomy at wide range of spatial scales (from cells images to whole body imaging) and also at different time scales (from milliseconds to capture heartbeats, years for aging studies, even up to eras for comparative anatomy and evolutionary biology studies). Moreover, different modalities provide information about distinct aspects of anatomy and in some cases about the physiology of organs, often being complementary in the understanding of the anatomical structures and their functions.



Evolution of topographic anatomy techniques. From left to right: illustration of Andreas Vesalius, “*De humani corporis fabrica*” (1543) [Vesalius 1543] [Garrison and Hast 2013]; first RX image in history by Wilhelm Conrad Röntgen (1895) [Röntgen 1895]; modern digital radiography; angiography; “three-dimensional visualization” of a computed tomography; magnetic resonance image.

Besides from the obvious implications in the clinical practice, one of the most important aspects of the revolution generated by *in vivo* observational techniques has been the strengthening of the anatomy as a discipline that describes the structure and configuration of the organs at a *population level* instead of at a *representative individuals level* [Pennec and Fillard 2014]. These

population descriptions of the anatomy are called *body plans* and refer to the “assemblage of morphological features shared by most members of a species”. Also, thanks to new imaging techniques are obviously non-destructive, it is possible to track anatomical changes over time in a single individual. For example, nowadays it is possible to follow the movement of a heart during a beat or the evolution and changes of a brain structure along decades. The studies devoted to the comparison of anatomies between different individuals on a population are called *cross-sectional studies* while those dedicated to tracking changes along time from a single individual are called *longitudinal studies*.

During last two decades there has been a large increasing in the use of diagnostic imaging. The use and application of medical imaging analysis techniques is continuously growing thanks to image quality improvements, the appearance of new modalities, and the price reduction in acquisition systems. These facts have motivated the development of several techniques for the quantification and analysis of medical imaging data. Analytical techniques typically require the use of computer algorithms, and due to the large amount of data contained in the medical images there is usually special efforts put in the efficiency of the algorithms. To give an idea, currently a brain MR image contains about 10 millions of intensity values encoded in 12 bits per value (according to standard [Pianykh 2012]) which results in 15 MB; a computed tomography image of a beating heart sampled at 20 frames per heartbeat contains about 650 million values.

In addition to the technical challenges proposed from new medical imaging modalities, scientific challenges are continuously been proposed from the analysis of data. Challenges from the medical imaging data analysis are very diverse, among others, data visualization, feature extraction, segmentation (localize and delineate an organ from the acquired data), image registration in both cross-sectional or longitudinal studies, quantification of anatomical shapes, measurement of movements, statistical analysis. A common goal of these research lines is to propose new techniques or improve existing aspects allowing to extract more, and more complete, accurate and precise *in vivo* information about the anatomy and physiology of the organs of the human body. Also, the scope of medical image analysis is very broad including from the assistance for radiologists in individual clinical diagnosis to fully automatic analyses used in medical research on populations with hundreds or thousands subjects.

The studied methodologies in this thesis can be used in a wide variety of applications from computer graphics, image processing, image analysis, computer vision. However, we will restrict ourselves to applications in the field of medical imaging and computational anatomy, in particularly to neuroimaging applications related with the study of Alzheimer’s disease (AD).

1.1.1 Computational anatomy

Currently, for most cases of medical research, it is not sufficient with qualitative descriptions to obtain reliable conclusions. Thus, it is necessary a mathematical modeling which allows a quantification setting. In this scenario, the anatomy has been evolved from qualitative to formal and rigorous descriptions. At the beginning, the quantification was proposed by the use of classical morphometry techniques based on magnitudes such as lengths, areas, volumes, angles. Modern anatomical analyses have been enriched from several areas of “advanced mathematics” and they are currently defined in terms of modern and differential geometry concepts, statistics on manifolds, statistical generative models, allowing to obtain useful mathematical descriptions for the medical research.

The combination of computational methods and statistical models to describe the anatomy of populations have led to a discipline called *computational anatomy*. It was born for modeling, analyze and quantify variabilities among anatomies. Its objectives include the estimation of representative or “averaged” anatomies from populations (also called *anatomical atlas* or

(templates) and the study and modeling of the anatomical changes across individuals, diseases, environmental conditions, age, race or even between different biological species. Computational anatomy also addresses challenges in the area of computed aided surgery. One of its most important uses is to describe the relationships between anatomies with respect to diseases, allowing to detect diseases from medical images and to give a quantifiable value for diagnostic imaging. Interestingly, modeling and quantification in diagnosis have been made long in the biochemistry area while in the area of clinical radiology is developed at a lesser extent, leaving several times the diagnostic to the criterion and appreciation of radiologists (or other medical professionals). Computational anatomy also provides a quantification level which is helpful to understand, predict, detect, correct and control pathological dysfunctions at both individual and population levels.

The computational anatomy framework have started by D'Arcy Thompson [Thompson 1917] in the early twentieth century whose used transformations of the underlying space to equalize and compare the shape of biological objects. Later, the framework was formalized by Grenander in the 70's [Grenander 1970] [Grenander and Miller 1998]. In this framework the anatomical characteristics of each individual are characterized by geometric deformations which presents with respect to an "average" anatomy or anatomical atlas [Miller et al. 1997] [Miller 2004]. This characterization is supported by the evidence that the anatomies of all individuals (at least at the topographical level) share the same body plan. In the absence of a physical or structural generative model that relates the anatomy of the individuals, the more general framework to model the morphological variations lies in a purely geometrical scenario. Therefore, the key entities for the quantification of anatomical variations are geometric transformations (which we call *spatial transformations*) that deform the atlas to "look like" individuals.

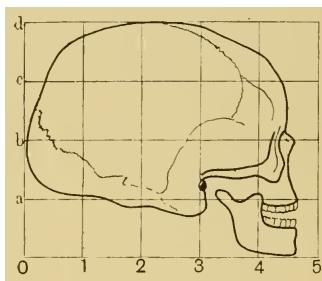


Fig. 404. Human skull.

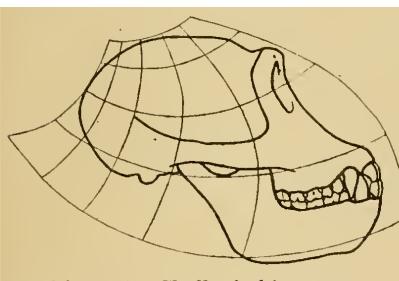


Fig. 406. Skull of chimpanzee.

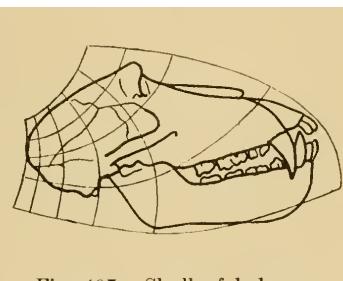


Fig. 407. Skull of baboon.

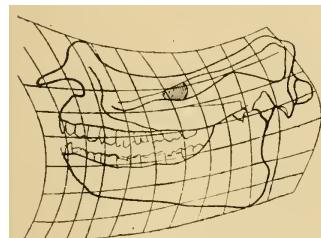


Fig. 397. *Titanotherium robustum*.

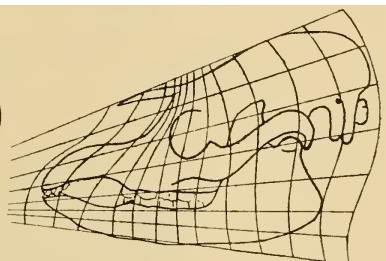


Fig. 398. Tapir's skull.

Images from D'Arcy Thompson, "On Growth and Form" (1917) [Thompson 1917]. (Chapter XVII, *On the theory of transformations, or the comparison of related forms*).

1.1.2 Neuroanatomy and neuroimaging

Within the anatomy discipline, the neuroanatomy branch is focused on the study of the brain and the central nervous system. At topographic level, the brain is formed by multiple well

differentiated structures and it is attached to the spinal cord by the brainstem. Joined to the brain stem is the cerebellum which is a separate structure under the two cerebral hemispheres. The hemispheres are divided by a deep groove called medial or longitudinal fissure and at its base can be found the corpus callosum which is a large bundle of nerve fibers interconnecting both hemispheres. Besides the interhemispheric connections passing through the corpus callosum, there also exist two interhemispheric fiber tracts called *anterior and posterior commissures*. Both commissures are commonly considered key fiducial points and have been used by Talairach and Tournoux to define a standard stereotactic coordinate system useful to represent brain anatomies [Talairach and Tournoux 1988]. Each hemisphere is mainly composed of grey matter, white matter and other well-distinguishable structures. The outermost layer of tissue is called cerebral cortex. It is a very convoluted structure folded into ridges (called gyri) and grooves (called sulci). Among the deepest and more consistently present sulci in all individuals are: *lateral sulcus*; *central sulcus*; *occipital sulcus*; *parieto-occipital sulcus*; and *calcarine sulcus*. These sulci usually subdivide the cortex into distinguished regions with well defined functionalities that are called brain lobes: *frontal lobe*, responsible of conscious thought; *parietal lobe*, which integrates sensory information, plays important roles in the manipulation of objects, and is involved in visuospatial processing; *occipital lobe*, mostly involved in the sense of sight; *temporal lobe*, senses of smell and sound, as well as processing of complex stimuli like faces and scenes; *limbic lobe*, involved in emotions and memory; and *insular cortex*, pain and sensual touch. Histologically, the cortex consists of grey matter, mostly composed of neuronal bodies. The cortex overlies the white matter which is mostly composed of neural axons whose myelin covers give the characteristic whitish color. In the inner part of the brain the subcortical nuclei and the basal ganglia can be localized. They are well delimited masses of grey matter with well defined functional purposes. Among others, they include the *caudate nucleus*, *thalamus*, *putamen*, *nucleus accumbens*, *globus pallidus*, *amygdala* and *hippocampus*. In addition, within the brain there are four anatomical cavities called ventricles by which cerebrospinal fluid circulates. To conclude this very brief anatomical description, the brain is covered by three membranes of connective tissue called meninges. The innermost is called pia mater, the middle layer is called arachnoid mater and the external is called dura mater. Finally, the bone structure widely known as skull surrounds and protects all this assemblage [Ono et al. 1990] [Vargas 2002] [Snell 2007].

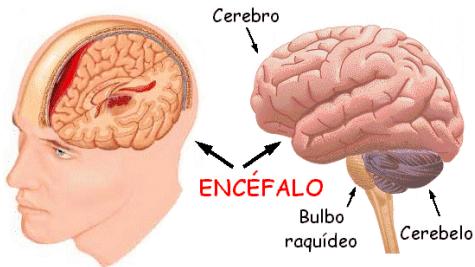
There are numerous medical imaging techniques to “observe” within the brain in a minimally invasive manner. Among its major purposes are remarkable the detection and characterization of central nervous system irregularities and medical research of psychiatric and neurodegenerative diseases.

Most used imaging techniques fall into two main categories:

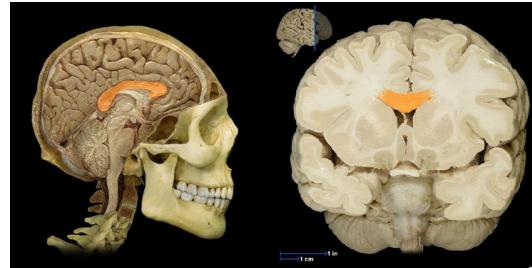
- **Structural image:** which provides a static anatomical information.
- **Functional image:** which provides information about metabolic and physiological processes that take place in different regions. Often the location of responses is not determined with sufficient spatial accuracy as to be considered an anatomical image.

For some imaging techniques this division is not so clear because structure and function are often intertwined in neuroimaging and therefore this classification is somewhat arbitrary [Symms et al. 2004].

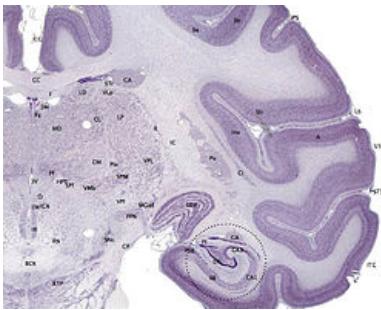
The most employed imaging technique to characterize neuroanatomy is, so far, magnetic resonance imaging (MRI). MRI exploits the phenomenon of nuclear magnetic resonance to produce structural images of the internal organs and other tissues. Different tissues can be distinguished by the characteristic properties of their emitted resonance signals. The decay



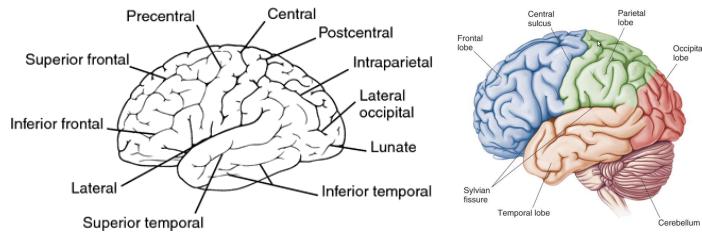
(a) Brain, brainstem, cerebellum, hemispheres



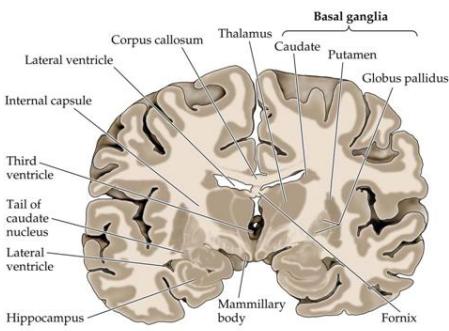
(b) Medial fissure and corpus callosum



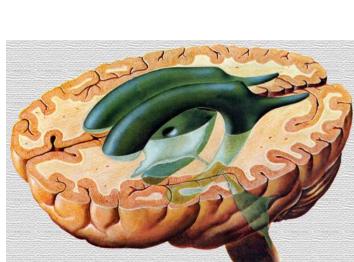
(c) Histological section distinguishing grey and white matter



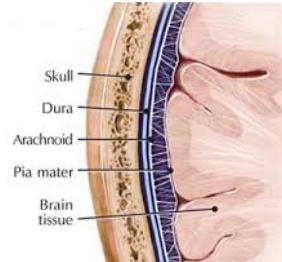
(d) Major sulci and brain lobes



(e) Subcortical nuclei



(f) Ventricles

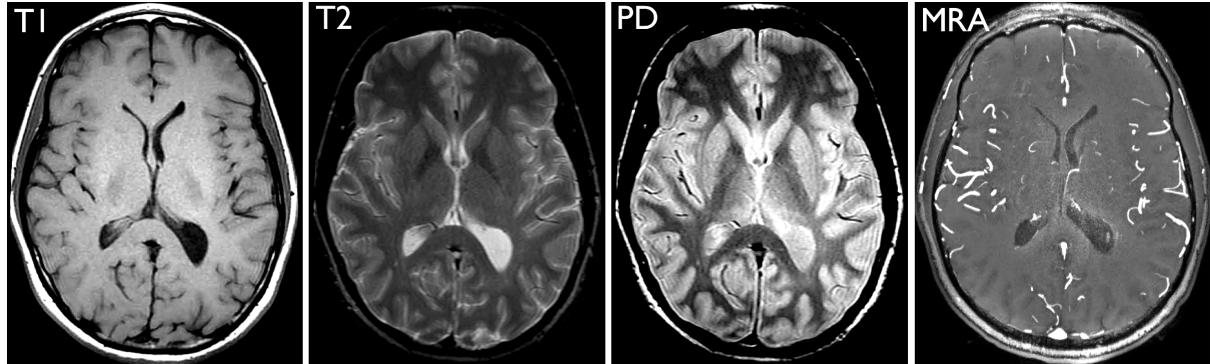


(g) Meninges

of the resonance signal as the nuclear spins return to a resting state is associated with time constants describing its longitudinal and transverse components. Recovery of the longitudinal magnetisation as the spins realign with the static magnetic field is known as spin-lattice or T1 relaxation. Decay of the transverse magnetisation as the spins dephase is known as spin-spin or T2 relaxation. The environment of the protons under consideration influences these time constants, and therefore the decay properties of the resonance signal. MRI contrast is dependent on the differing T1 and T2 relaxation properties of various biological tissues [Bushong 2003] [Weishaupt et al. 2003] [Gibby 2005] [Constantinides 2014].

Structural MRIs provide excellent contrast between soft tissues and have a high spatial resolution. Moreover, since the brain is a static organ, the acquisition process usually results in images with high quality as opposed to, for example, cardiac imaging. MRIs include many modalities which, generally, provide complementary information about the tissue.

Throughout this thesis, T1-weighted images are used which are the most popular general purpose 3D MR acquisition modality. They provide good contrast between grey and white matter, and a good balance between resolution, signal-to-noise, and scanning time. On the other hand, T2-weighted imaging allows to visualize and quantify vascular lesions in the white



Images of the same subject with different MRI modalities. From left to right: T1-weighted image; T2-weighted image; proton density; MR angiography.

matter [Fazekas et al. 1993].

1.2 Tools from computational anatomy

1.2.1 Registration of anatomical structures

A key tool in morphometric studies and computational anatomy is the determination of the geometric transformation that relates an anatomical atlas with an individual. This alignment is called registration and is performed by finding the geometric transformation which best overlaps the deformed structures of the atlas to the structures of the individual. In the same way the individual may be deformed to be aligned with the atlas. The registration process is usually posed as an optimization of a measure about how much “overlapped” or “similar” are two geometrical structures [Modersitzki 2004] [Hajnal and Hill 2010] [Goshtasby 2012] [Sotiras et al. 2013]. The optimization is conducted over a set of spatial transformations which is typically selected in accordance with the purposes of the registration. When two anatomical structures are registered, their shapes, appearances, measurable magnitudes in them, or in general its visual characteristics should be spatially aligned. The registration process establishes a spatial correspondence, such that each point in one of the images, representing a localized measurement of a portion of tissue, is spatially aligned with the same tissue in the deformed image [Rohlfing 2012] [Rohlfing and Avants 2012].

After the registration process, the morphological variability between the atlas and individuals are removed and such variability stays encoded in the resulting deformation. This scenario also allows to represent the anatomical variabilities in a stereotyped or standardized reference system [Talairach and Tournoux 1988].

1.2.2 Geometrical description of the anatomy

Anatomical structures at the topographical level will be considered throughout this thesis as the information they provide about their geometry. Therefore, the anatomical structures will be described as *geometrical objects* in the space. The space where these objects are defined will be called *ambient space* and will be denoted by \mathbb{A} . It will be considered that \mathbb{A} is the Euclidean space \mathbb{R}^d with d typically equal to 2 or 3. The space \mathbb{A} can be set up with a coordinate system, such as a Cartesian system, so that each point $y \in \mathbb{A}$ can be described by the coordinates where it is placed.

The geometrical information that can be obtained from geometrical objects usually have different levels of conceptual interpretation. Some contents are only visual characteristic about

the appearance of the objects which can be easily distinguished with some computational processing techniques. Some others contents are high-level interpretations which usually have to be performed by experts.

Among the more common types of geometrical objects most often considered in medical imaging are:

- Significative points: they are isolated points with distinguishable characteristics, such as corners, intersection, boundary points, and some other points with a relevant characteristic or visual information. They can be obtained with detectors of significant features on the acquired images.
- Landmarks (also known as fiducial points): are the points that can be found in a structure and homologously in all (or almost all) the structures of a population. They are also significant points but, in addition, landmarks are “named” or “labelled” in such a way that along a population same names identify *corresponding* locations. Their identification or detection in the images is usually done manually and consists in the identification, by an expert, of relevant positions that can be purely structural or also with a physiological meaning.
- Significant curves or surfaces: as an extension of significant points, are a continuous collection of points that occupy the site of some characteristic of the image. They are usually distinguishable edges from the images, like sudden changes of intensity. Similar to the landmarks case, these structures can be labelled if there exist a fair correspondence along the images of a population.
- Regions of interest (ROI): most anatomical structures are often identified as a specific area or volume of tissues with similar characteristics. The delineation of structures is called segmentation. The regions are commonly described by their boundaries which are closed curves in 2D or closed surfaces in 3D.
- Images as scalar functions: acquired medical images can be considered as geometric information since each point of the ambient space have assigned a color or an intensity value. Instead of a description based on pixels or voxels, as it is usually described by the acquisition device, we shall consider images as continuous scalar functions defined over the entire ambient space. To convert from the discrete nature of the acquired images to a continuous representation, an interpolation mechanism have to be specified.

1.2.3 Spatial transformations

Previously it was emphasized that the description of geometrical objects will be as an embedded object in a continuous ambient space \mathbb{A} . Spatial transformations are deformations of the ambient space and are accordingly applied on the geometrical objects defined on it such that the transformed object is the collection of its transformed points.

In the modeling of the variability of anatomical structures, it is desirable (and even mandatory in most cases) that the geometric transformation does not destroy the integrity of the anatomical structure. Therefore, spatial transformations will be homeomorphisms of the ambient space, *i.e.* continuous and bijective mappings from \mathbb{A} to itself. An immediate consequence is that any spatial transformation has an inverse transformation, which is also a spatial transformation. This invertibility property can be justified in the registration process due to several reasons: there is usually no preference about the order in which objects have to be registered and it is expected that the registration of an object P towards an object R results in a process as valid as the registration of R towards P ; registration process only makes sense when both objects describe the

same content, with the same structures or field of view, and therefore the appearance of holes in the deformation mapping would suppose that one of the objects contains more spatial information than the other one; finally, the registration is a process whose major goal is to determine spatial correspondences between both objects and the appearance of folds in the deformation mapping means that some points in one object have more than one correspondent points in the other object.

The application of a spatial transformation \mathfrak{h} on a point y of the ambient space will be denoted as $\mathfrak{h} \star y$ and this operation will be called *action*. When the spatial transformation is applied on a geometrical object P , it will be also denoted by $\mathfrak{h} \star P$ and also called *action of \mathfrak{h} on the object P* .

As spatial transformations are homeomorphism of the ambient space, it is straightforward to notice that successive actions of several spatial transformations can be described by the application of a unique spatial transformation, *i.e.* for any two spatial transformations \mathfrak{h} and q there exist a spatial transformation r such that

$$q \star (\mathfrak{h} \star y) = r \star y$$

for all points y of \mathbb{A} (and also for all objects). The spatial transformation r is called the *composition of q and \mathfrak{h}* and is denoted by $r = q \bullet \mathfrak{h}$. This property together with the existence of the inverse spatial transformation \mathfrak{h}^{-1} indicate that spatial transformation have a (*transformation*) *group structure*.

According to Arnold [Arnold 1992, p.58]: “The concept of a transformation group is one of the most fundamental in all of mathematics and at the same time one of the simplest: the human mind naturally thinks in terms of invariants of transformation groups (this is connected with both the visual apparatus and our power of abstraction)” (see also [Dodwell 1983] [Miao and Rao 2007]).

The action of transformation groups on geometrical objects usually preserves interesting high level characteristics, magnitudes and relations between the structures that compose the objects. Transformation groups are also the language to define symmetries both at theoretical level and in application problems.

1.2.4 Similarity measures and regularization

A similarity measure (denoted by **sim**) is a tool to calculate how much of spatially aligned are the points of a geometrical object with respect to the corresponding points in another object [Modersitzki 2004]. Conversely, a measure of dissimilarity (denoted as **dis**) quantifies how much distant are the geometrical objects [Goshtasby 2012].

When spatial correspondences are known, the task is easily accomplished by measuring the distances in \mathbb{A} between corresponding points. For example, let us start by considering the case of two geometrical objects composed of landmarks $P = (p_1, p_2, \dots, p_L)$ and $Q = (q_1, q_2, \dots, q_L)$ with the assurance that each point p_l is homologous to q_l . By deforming the object P with a spatial transformation \mathfrak{h} it can be measure the “quality of the alignment” with respect to the object Q by the following dissimilarity energy

$$\mathbf{dis}(\mathfrak{h} \star P, Q) = \sum_{l=1}^L \|\mathfrak{h} \star p_l - q_l\|^2,$$

where $\|\cdot\|$ is the Euclidean norm in \mathbb{A} . The lower the dissimilarity energy the better aligned will be $\mathfrak{h} \star P$ with Q .

In most practical cases, point-to-point correspondences between images are unknown. As it was stated in [Rohlfing and Avants 2012], the aim of the registration process is to determine

the correspondence between the locations of both images and a by-product of the obtained correspondence is the maximization of a similarity (or the minimization of a dissimilarity) measure between the images. Roughly, the higher the similarity the more closer the correspondent points and better achieved the correspondence. Then, the idea to obtain the correspondences is to deform one of the image such that both images look spatially as much similar as possible.

Two main cases should be considered when defining a similarity (or dissimilarity) measure between scalar functions or images: when images come from the same acquisition process; and, when they are images from different modalities. In the first case, called mono-modality registration, it is reasonable to assume that the appearance (color and texture) of homologous structures is the same. Under this model, a widely used dissimilarity measure is the *sum of squared differences* (SSD), which is defined as

$$\mathbf{dis}(\mathbf{A}, \mathbf{B}) = \int_{y \in \mathbb{A}} (\mathbf{A}_{(y)} - \mathbf{B}_{(y)})^2,$$

where \mathbf{A} and \mathbf{B} are two images. The assumption that appearances are invariant along images is not always true and there may be, for example in MRI, field inhomogeneity factors that modify the intensity value of images [Hou 2006]. Thus, other mono-modality measures have been proposed in order to provide more robust similarity measures, for example global or local correlation measures. For the registration of images from different modalities, classical similarity measures have been proposed based on information theory [Pluim et al. 2003], but they will not be considered in this thesis.

Once a dissimilarity measure is chosen, the registration problem is typically posed as finding the spatial transformation $h^* \in \mathcal{S}$ such that $\mathbf{dis}(h^* \star P, Q)$ be a minimum (or a maximum in the case of a similarity measure). The space of possible spatial transformations \mathcal{S} , where to find h^* , is typically defined in terms of a *spatial transformation model* which provides a parametrization for the spatial transformations of interest. The dimension of the parameters defines the degrees of freedom of \mathcal{S} .

When \mathcal{S} has low degrees of freedom the minimization problem is usually well behaved in terms of existence and uniqueness of its solution. Nevertheless, the minimization problem becomes ill-posed when high dimensional parameterizations are needed to describe \mathcal{S} . This issue can be compared with an undetermined linear system where more variables than equations are specified. For example, in the case of registering L landmark positions in \mathbb{R}^d the problem results in a least square problem determined by $L \cdot d$ equations and when the set \mathcal{S} has more than $L \cdot d$ degrees of freedom it is usual that more than one spatial transformation from \mathcal{S} solve the problem. The loss of uniqueness of solutions also affects to the registration of scalar functions or images. This issue is called *aperture problem* and refers to the fact that perturbing deformations along level sets (isophotos) or in areas with a constant intensity value do not change the dissimilarity measure.

To face with registration problem under deformations with high degrees of freedoms it is common to include a regularization term of the spatial transformation. This term will be denoted as $\mathbf{reg}(\cdot) : \mathcal{H}_\mathbb{A} \rightarrow \mathbb{R}_{\geq 0}$. Finally, the registration problem is given by the following optimization problem:

$$\underset{h \in \mathcal{S}}{\text{minimize}} \quad \mathbf{reg}(h) + \lambda \mathbf{dis}(h \star P, Q),$$

where λ is a tradeoff between the dissimilarity and the regularization.

1.3 Mathematical bases for the statistical analysis of spatial transformations

Both, in the registration process and in the subsequent quantification of the resulting deformations, the mainly involved mathematical entities are spatial transformations. They are elements belonging to several spatial transformation models used throughout the thesis. At the time of performing statistical analysis on the transformations, the first idea that comes to mind is to use classical statistical concepts, for example over they parameterizations. However, classical statistical tools are typical defined only for sets equipped with a vector space structure.

In Section 1.2.3 it was stated that most interesting spatial transformation sets possess a group structure under the composition and the inversion operations. In general, taken into account these operations only, spatial transformation sets cannot be endowed with a vector space structure. Moreover, it is not possible to isomorphically identify spatial transformations groups with a vector space as was performed for other geometrical entities such as positive definite tensors in what is called the log-Euclidean framework [Arsigny et al. 2007].

In the following, it will be sketched an intrinsic formulation where to define statistical tools. It will be highlighted how to develop the framework by adding the minimal ingredients over the existent algebraic group structure that transformation sets usually have.

1.3.1 Non-classical formulation of statistical tools

The development of intrinsic statistical tools must be consistent with the classical statistical tools in the sense that by performing an intrinsic analysis over vector spaces the classical result is recovered. Let us start by considering the mean estimator of a distribution from a finite sample $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N\}$ of vectors from some Euclidean space \mathbb{E} . It is widely known that the average of these elements

$$\hat{m} = \frac{1}{N} \sum_{i=1}^N \bar{x}_i.$$

provides an empirical estimator for the mean of the distribution from where they were drawn. Also, let us consider the empirical variance relative to a point $\bar{y} \in \mathbb{E}$ defined by

$$\hat{\lambda}(\bar{y}; \{\bar{x}_i\}_{i=1}^N) = \frac{1}{N} \sum_{i=1}^N \|\bar{x}_i - \bar{y}\|^2,$$

where $\|\cdot\|$ is the Euclidean norm in \mathbb{E} . Noticing that $\hat{\lambda}$ is a convex function and that $\partial_{\bar{y}} \hat{\lambda}(\bar{y}; \{\bar{x}_i\}) = 2 \sum_i (\bar{x}_i - \bar{y})/N$, it is straightforward to check that \hat{m} (the average of the samples $\{\bar{x}_i\}$) is the global minimum of the function $\hat{\lambda}$. The function $\hat{\lambda}(\bar{y})$ can be interpreted as dispersion of the sample $\{\bar{x}_i\}$ around the element \bar{y} , where each sample \bar{x}_i contributes with the square of the distance between \bar{x}_i and \bar{y} . Thus, by replacing $\|\bar{x}_i - \bar{y}\|$ with a distance function **distance**(\bar{x}_i, \bar{y}) in the space \mathbb{E} , a variational definition of the “central point” of a sample is given by

$$\hat{m} = \underset{\bar{y} \in \mathbb{E}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N \operatorname{distance}(\bar{x}_i, \bar{y})^2.$$

In this variational definition the requirement that \mathbb{E} must be a vector space was relaxed and only a metric structure is needed on \mathbb{E} , *i.e.* it is just required a distance function **distance**: $\mathbb{E} \times \mathbb{E} \rightarrow \mathbb{R}_{\geq 0}$ fulfilling the conditions of positive-definiteness, symmetry and satisfying the triangle inequality.

Similarly, another “central elements” can be defined, for example a generalization of the median (which is classically defined on an ordered space) is given by minimizing $\sum \operatorname{distance}(x_i, y)$ and,

in general, it is called *central element of degree α* to the minimizers of $\sum \text{distance}(\chi_i, y)^\alpha$, for $\alpha \geq 0$. This formulation looks as an interesting alternative to develop the statistical tools to deal with abstract entities such as spatial transformations. In this thesis, this “distance based” generalization of the classical statistics will be used to generalize several statistical concepts and analytical tools such as regressions and hypothesis tests.

However, all the necessary ingredients to make statistics on the spatial transformations are not yet defined. A metric structure have to be assigned to the spatial transformations, *i.e.* one should be able to answer the following question: How distant are two spatial transformations?. This will help to answer: how much two geometrical objects differ?. Once we are able to answer those questions, we are ready to known how much several anatomies differ.

1.3.2 Spatial transformation sets with differentiable structure

Given a set of spatial transformations $\mathcal{S} \subset \mathcal{H}_{\mathbb{A}}$ (where $\mathcal{H}_{\mathbb{A}}$ is the set of all homeomorphisms of \mathbb{A}), the goal to assign a metric structure begins by imposing a topology on \mathcal{S} , *i.e.* to be able to determine open neighborhoods around every element of \mathcal{S} . There can exist many different ways to assign a topology on \mathcal{S} , for example a topology can be inherited from a bijective parametrization from an open set of \mathbb{R}^m , for some m . However, this thesis is focused on the cases when \mathcal{S} is a *Lie group*. A Lie group is a set which besides being a group, it is a differentiable manifold where the composition operation is a differentiable operation within the manifold structure. It is remarkable that most of the transformation sets used in medical imaging and computer vision applications are Lie groups. Then, let us start from this structure to endow a metric structure to the set.

Once a (differentiable) manifold structure on \mathcal{S} is assigned, the set \mathcal{S} can be realized at the intuitive level as a “(smooth) curved surface” embedded in a higher dimensional Euclidean space. According to Whitney’s Theorem [Burns and Gidea 2005], every smooth k -dimensional manifold can be isomorphically embedded in \mathbb{R}^{2k} . This “visual notion” is useful in the description of many concepts introduced throughout the thesis.

Almost every Lie group that we will encounter (or at least many of most studied Lie groups) are isomorphically equivalent to a matrix group. These include the general linear groups $\mathbf{GL}(n)$ of non-singular matrices from \mathbb{M}_n (the set of all $n \times n$ real matrices), and various of its subgroups obtained by imposing linear, bilinear or polynomial constraints over the entries of the matrices. Algebraically, the group operations become concrete and numerically representable by the matrix multiplication and matrix inversion operations. This identification provides a huge simplification in the study of the topological, algebraic, and analytical properties of Lie groups. Furthermore, the identification of a transformation group with a matrix group allows to deploy in a simple manner the entire machinery of the computational analytical tools in programming environments such as MATLAB, Mathematica, or several other scientific libraries that incorporate matrix algebra operations natively and in a highly efficient manner.

1.3.3 Riemannian distances and exponential mappings

Once a set of spatial transformations \mathcal{S} is endowed with a differentiable manifold structure, it is possible to calculate derivatives of curves defined on \mathcal{S} . These derivative are called *velocities of the curve* and their norms are referred as *speed*. The length of a curve segment is achieved by integrating speeds along the curve. Notice that to be able to quantify the instantaneous speed of a curve, velocities have to be considered as elements belonging to a normed vector space. The most typical procedure is by smoothly assign an *inner product* to every *tangent space* of \mathcal{S} .

This assignment is called *Riemannian metric* and it is said that the manifold \mathcal{S} is a *Riemannian manifold* [Do Carmo 1992] [Gallot et al. 2004].

Considering all possible curves that connect two given fixed elements of \mathcal{S} , the length of the shortest curve defines the *Riemannian distance*. Endowing \mathcal{S} with this distance, the set \mathcal{S} acquires a *metric space* structure. By means of variational calculus techniques, it can be proved that curves between two fixed elements and with an extreme length are described by a second order system of differential equations which will depend on the considered Riemannian metric on \mathcal{S} . Using results from the theory of differential equations, it can be seen that there is a unique curve passing through an element $q \in \mathcal{S}$ with a given velocity \mathcal{V} and with an extremal length. These curves are called *geodesics* and are described by the *Riemannian exponential function* which is denoted by $\text{Exp}(\cdot)$, such that $\gamma_{(t)} = \text{Exp}_q(t\mathcal{V})$ is a geodesic that at $t = 0$ passes through q with velocity \mathcal{V} . The Riemannian exponential is denoted by capitalized letters to distinguish it from other exponential concepts that will be discussed along the thesis. The inverse function of $\text{Exp}(\cdot)$ is called *Riemannian logarithm*, and determines the velocity that a geodesics passing through q must follow to arrive at $t = 1$ to a given element $p \in \mathcal{S}$, i.e. $\text{Log}_q(p) = \mathcal{V} \Rightarrow \text{Exp}_q(\mathcal{V}) = p$. The Riemannian logarithm function is closely related to the Riemannian distance resulting $\text{distance}(q, p) = \|\text{Log}_q(p)\|_q$, where $\|\cdot\|_q$ is the norm in $T_q\mathcal{S}$.

This thesis will discuss in detail these ideas to reach to computational methods or algorithms allowing the calculation of $\text{distance}(q, p)$.

1.3.4 Invariance under spatial transformations

One of the most important aspects of the modern geometry is its description in terms of transformation groups. F. Klein formalized this idea in his “Erlangen program” [Klein 1893] [Yaglom et al. 1988] by defining the geometry of a given space as the set of properties or magnitudes that remains invariant under the action of a particular transformation group \mathcal{G} . In this way, the geometry of the ambient space \mathbb{A} is not only determined by \mathbb{A} but also by the transformation group which is considered to act on \mathbb{A} . In this setting, the statistical analysis on \mathcal{S} must consider that distances between spatial transformations do not change under the action of some selected group of transformations. This invariance has a great impact on practical applications where it is desirable that the outcome of the quantification is invariant to the symmetries of the problem. For example, differences between two shapes should take into account the equivalence class that defines shapes. That is, if a shape is defined by all the features that are invariant under a change of the coordinate system, this symmetry must be included in the quantitative analysis of shape variations [Pennec and Ayache 1998].

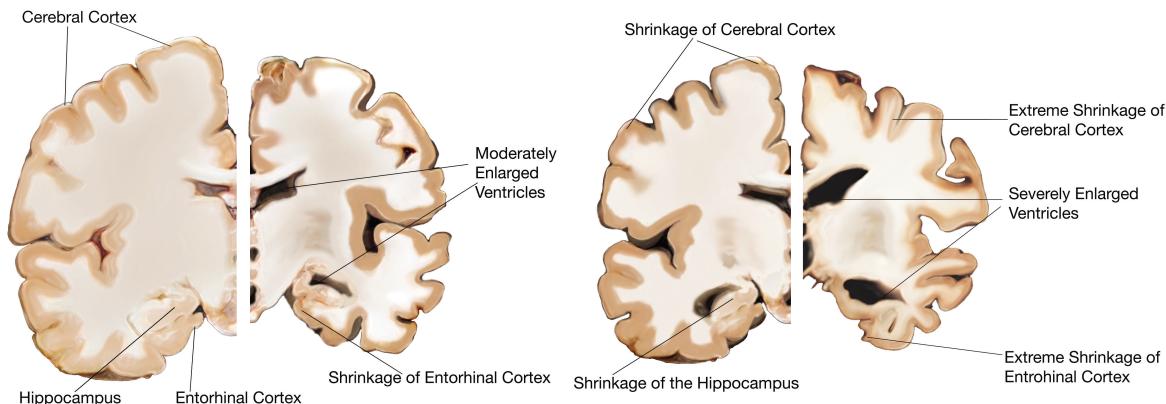
For the case of quantify spatial transformations, it is interesting to consider those distances that are preserved under the symmetries given by the same transformation group. Thus, we will consider in the statistical analyses *invariant Riemannian distances*, which are those distances holding $\text{distance}(q, p) = \text{distance}(r \bullet q, r \bullet p)$ for any element r from the considered transformation group. In the same way, it can be considered distances holding $\text{distance}(q, p) = \text{distance}(q \bullet r, p \bullet r)$. While the former are called *left-invariant distances*, the latter are called *right-invariant distances* and when both invariances hold simultaneously it is called *bi-invariant distance*. Different applications can lead to symmetries under left-actions, right-actions or from both sides. Left-invariant distances are usually preferred for applications where measurements depend on an external coordinate system [Park 1995] [Pennec and Ayache 1998], and right-invariant distances are usually appropriate for applications where a reference system is used to normalize the results [Zacur et al. 2013]. However, the definition of the symmetries is completely application-dependent and must be studied carefully for each application.

1.4 Morphometry in neuroimaging

Neurodegenerative diseases cause impairment and neuronal death which over time results in a loss of cognitive functions, memory loss or a state of motor disability. Patients with neurodegenerative diseases tend to fall over time in a *dementia* state which is the loss of cognitive functions more severely than the natural decline caused by the effect of age. Cognitive changes affect the ability of patients in performing everyday tasks. They express as a degradation of functions associated to memory, language, attention, orientation, problem solving areas. Neurodegenerative dementias include Alzheimer's disease (AD), dementia with Lewy bodies, frontotemporal dementia, Huntington's disease, and Creutzfeldt–Jacob disease [Horner et al. 2007].

Neuronal death caused by neurodegeneration results in a decrease in the size of organs and tissues of the nervous system. This morphological change is called *atrophy* and it is considered as an expression at the topographical level of a functional and physiological process at the cellular level. From acquired brain images, specially from MRI, morphometric techniques are frequently used to study and quantify these atrophies and other anatomical changes caused by neurodegenerative diseases.

Scientifically interesting and clinically important patterns of brain changes in dementia can be detected from the analysis of MRIs. Differences in atrophy patterns between groups of subjects may be relevant to diagnosis, tracking of disease progression, and monitoring of potential disease-modifying treatments.



Neurodegenerative and progressive atrophy in Alzheimer's disease. Left: comparison of *ex vivo* brain morphologies from a healthy subject and a Alzheimer's disease patient in a moderate stage. Right: comparison between Alzheimer's disease at moderate stage and at a very advanced stage.

1.4.1 Alzheimer's disease

Alzheimer's disease (AD) is the most common dementia with a prevalence about 50–60 % compared to other dementias. AD cases are usually divided into three main types which share the same pathological features: sporadic late-onset AD which tends to manifest after age 60; early-onset AD (accounting for 5–10 % of all AD cases) mostly associated to genetic defects; and the extremely rare (less than 1 % of all AD cases) familial AD which typically has an earlier onset around the age 40 [Dawbarn and Allen 2007]. Age is the most significant risk factor associated with the development of sporadic AD, although genetic, environmental, and other factors are also relevant. Regarding genetical factors, the Apolipoprotein E gene (ApoE) is so far the most associated with the development of sporadic AD. There are three major isoforms of the ApoE: $\varepsilon 2$,

$\varepsilon 3$ and $\varepsilon 4$. The most common allele is $\varepsilon 3$, which is present in 70–80 % of population [Zannis et al. 1981]. The $\varepsilon 4$ allele is associated with an increased risk of developing AD, while $\varepsilon 2$ has a neuroprotective effect [Corder et al. 1993] [Farrer et al. 1997] [Graff-Radford et al. 2002].

Despite AD is a progressive disease which can usually develop or manifest in different ways, it is possible to identify some common symptoms. These symptoms are memory loss, mood disturbances and early difficulties in carrying out daily activities. As the disease progresses, impairments in memory, communication and mobility are increased and patients gradually become dependents (usually under the care of their families). United Nations estimates a current prevalence about 25 million patients. Considering current conditions, about 115 million worldwide patients are expected for year 2050 [Organization and International 2012]. Within the Spanish territory the current prevalence is estimated at around 400 000 patients (at 2008) [de Pedro-Cuesta et al. 2009]. These high prevalence rates are mainly due to the increase in the life expectancy and longevity that people experience today thanks to improvements in life quality and medical procedures. At present, no treatment has been shown convincingly cure, prevention or reduction in the progression of the disease. AD patients generate high health costs and also indirect social costs. Also, AD patients present a higher lifetime with disabilities than other common illness such as stroke, musculoskeletal disorders, cardiovascular diseases or cancer. Even modest progresses in reducing degeneration progress may result in significant benefits for patients, their families and healthcare systems in general.

Clinical diagnosis of AD are made according to different consensual criteria [McKhann et al. 1984] [Hampel et al. 2013], such as the NINCDS-ADRDA criterion, which provides guidelines for the classification of patients. People found to have memory problems incommensurate with normal aging, but without sufficient cognitive problems to fulfil criteria for dementia, are said to have Mild Cognitive Impairment (MCI) [Morris et al. 2001]. Some such people later go on to suffer AD, while others not. The conversion rate is approximately 10–15 % per year [Petersen et al. 1999].

Diagnostic criteria of AD mainly involves clinical examination, consideration of medical history and several behavioral and cognitive test, which is assessed based on direct interviews with the patient and relatives. Test for the evaluation of the severity of dementia include the *Mini-Mental State Examination* (MMSE) [Folstein et al. 1975] [Cockrell and Folstein 2002], *Alzheimer's Disease Assessment Scale Cognitive SubScale* (ADAS-cog) [Mohs and Cohen 1988] [Mohs 1994] [Manzano et al. 1994] [Peña-Casanova 1997], and *Clinical Dementia Rating-Sum of Boxes* (CDRSB) [Morris 1993].

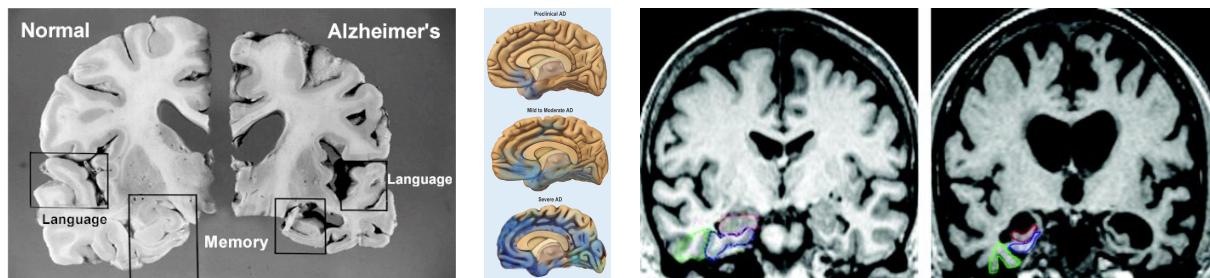
At histopathological level, AD has two fundamental characteristics observed more than 100 years ago by the German psychiatrist and neuropathologist Alois Alzheimer (1864-1915). AD is characterized by the co-occurrence of different phenomena, starting from the deposition of amyloid plaques and neuro-brillary tangles, to the progressive synaptic, neuronal and axonal damage. These findings reported by several independent scientific studies led to the formulation of the so called amyloid cascade hypothesis [Jack Jr et al. 2010], in which the authors proposed a temporal model for the progression of the neurodegenerative biomarkers in AD. According to this model, the pathology is initiated by the accumulation of amyloid deposits in the brain, which can occur decades before the development of the clinical symptoms. The amyloid deposition induces a cascade of pathological processes which lead to the endpoint of stereotypical patterns of neuronal loss in the temporal and cortical areas. Above a certain threshold of structural damage, the amnestic symptoms progressively increase from the initial appearance of memory disturbance, which characterizes the clinical intermediate stage commonly defined as amnestic mild cognitive impairment (MCI), to the widespread cognitive deficits in multiple domains. More recently, it has been identified that AD state is also correlated with low levels of A β and higher levels of Tau protein measured in Cerebro Spinal Fluid (CSF) from lumbar punctures. Also,

it have been developed methods for pre-symptomatic AD detection from Positron Emission Tomography (PET) modalities with radio-tracer markers such as PiB and FDG, being some of the most promising potential AD biomarkers.

Although there are some other etiological and progression models [Drachman 2014], all of them coincide that neuronal damage begins several years before the onset of clinical symptoms. Currently, there is a growing interest in moving pharmacological research and palliative treatments towards the pre-symptomatic stages of AD. Then, it is necessary to do research in the field of early diagnosis to determine biomarkers of the disease as early as possible. Reliable and well-validated biomarkers of AD are needed for epidemiological screening, monitoring disease progression and in a secure future, for measuring the response of the future treatment. Early markers would have also a great value in the development of drugs. Markers would assist in the selection of optimal candidates at the time of undertaking lengthy and expensive clinical trials.

1.4.2 The role of structural neuroimaging in AD

Structural atrophy of the whole brain characterizes the transition from MCI to AD and it was shown to strongly correlate with cognitive performance and neuropsychological scores [Small 2002] [Frisoni et al. 2010] [Ewers et al. 2011] [Johnson et al. 2012]. Furthermore, typical patterns of neurodegeneration in cortex and temporal areas are measurable at early and pre-symptomatic stages of the disease. Among the most accepted methods of analysis is the measurement of neuronal atrophy and the consequent decrease in volume in regions of interest such as the hippocampus and the temporal lobe. At cortical level, areas from the temporal lobe reduce their thickness. Indirectly, due to intracranial volume must be preserved, it can be seen significant increasing of the ventricles and cortical sulci. Atrophies can be measured by typical MRI study with a standard protocol such as T1 weighting. Thus, structural MRI has become an essential tool to support clinical diagnosis and medical research [Whitwell et al. 2007].



Typical atrophic patterns of neurodegeneration in Alzheimer's disease. Left: histological sections. Central: illustration of the evolution of the atrophied areas in different AD stages. Right: volumetry quantification on brain MRI.

Numerous methodologies for the characterization and analysis of the anatomy of the brain have been proposed. They can be classified into two main families: local and global characterization techniques. Local methods provide a characterization of the geometry of a part of the brain, usually a structure or a delimited region of interest. On the contrary, global characterization methods provide a characterization of the entire brain without having to specify or decide *a priori* the structure or region to study. Below are some of the methods used most often.

Volumetry

Volumetry consists in determining and analyzing the volume that a structure or a delimited region of the brain occupies. Volume is a very intuitive descriptor and it is relatively easy to calculate.

Furthermore, it is a simple magnitude and classical statistical tools can be easily used for its analysis. Volume is used as biomarker for the presence or severity of the AD [Pruessner et al. 2000] [Jack et al. 2008b] [Colliot et al. 2008] [Leung et al. 2010a]. It is also used in other diseases such as schizophrenia [Vogeley et al. 1998], autism [Craig et al. 2007] [Brieber et al. 2007], and for monitoring the effect of therapies.

Volumetry is specific to an area of study, region of interest (ROI), that must be specified *a priori*. Determination of the volume requires the delineation or segmentation of the structure under study. In recent years, computational segmentation techniques, both automatic and supervised, have improved greatly, in part favored by the improvement in the quality of acquired images.

Boundary Shift Integral (BSI)

It is a technique that allows to measure the volume change of a structure from two longitudinal MRIs.

The images are first aligned, then structure is delineated in one of the images and the volume change is calculated directly from the change of intensity values from the other image along the boundary of the structure [Freeborough and Fox 1997] [Leung et al. 2010b]. BSI is a technique relatively simple to implement and allows estimates of the atrophy with sub-voxel precision which increases the significance of the results in comparison with a traditional volumetry. Also, robustness improvements are obtained since segmentation is performed in a single image decreasing the segmentation errors.

Several applications of BSI on the AD can be found [Fox and Freeborough 1997] [Schott et al. 2005] and on other pathologies such as Parkinson's disease [Wild and Fox 2009], or progressive supranuclear palsy [Paviour et al. 2006].

Morphometry of a single structure

As the volume is a nonspecific descriptor of the geometrical aspects of a structure, several methods have been proposed to characterize the geometry in a more complete manner. One of the most used descriptors is by means of anatomical landmarks and the analysis of their coordinates [Cootes et al. 1995] [Bookstein 1997]. Landmarks description depends on the selected coordinate system and on the location of the subjects, therefore coordinates do not seem to be an intrinsic descriptor of the morphology. To resolve this limitation, typically a prior alignment is performed to remove postural effects [Small 1996] [Dryden and Mardia 1998].

In this methodology the correspondence between the anatomical landmarks is essential. The determination of landmarks and correspondences are usually obtained manually, but it is a very tedious task and prone to intra- and inter-operator inaccuracies. Alternatively, automatic or supervised methods have been proposed to estimate correspondence between points based on proximity criteria [Chui and Rangarajan 2003] [Datar et al. 2013] or methods based on images registration to propagate landmarks positions [Frangi et al. 2002].

Besides, other works proposed to characterize the geometry of an object by its medial representation or skeleton and the thickness [Pizer et al. 2003] [Fletcher et al. 2004].

Studies on the cortex thickness

Another methodology of analysis is based on the quantification of the cortical thickness. The methodology starts by outlining the cortical surfaces, both the external or pial surface and the internal surface at the grey/white matter interface. At each point of the surface the cortex thickness is geometrically measured. Finally, after a registration process of the surfaces it

is possible to compare the cortex thickness for the different cortical lobes [Dale et al. 1999] [Fischl et al. 1999].

This studies have been shown significant differences on the cortex thickness in pathologies such as schizophrenia [Nesvåg et al. 2008], autism [Ecker et al. 2010], AD [Fortea et al. 2010].

Whole brain morphometry techniques

Morphometric techniques described above require to specify a structure or a set of structures of interest, as well as a segmentation step performed manually or automatically. The main motivation for using global morphometry techniques is to allow the analysis of anatomies without the *a priori* wisdom about the location of the findings.

Voxel-based morphometry (VBM)

This technique consists in the quantification of the change in the density of tissues. First, grey and white matter regions are probabilistically identified on several subjects. Latter, every subject brain is registered to a template, and differences in the probability of the tissues are compared across the registered images [Ashburner and Friston 2000].

VBM technique has permitted the detection of structural differences of grey/white matter in a number of central nervous system diseases, such as schizophrenia [Honea et al. 2005] [Hulshoff Pol et al. 2006], epilepsy [Bernasconi et al. 2004], depressive disorder [Lyoo et al. 2006] [Nugent et al. 2006], Parkinson's disease [Ramírez-Ruiz et al. 2005], personality disorder [Rüsch et al. 2003], Huntington's disease [Kassubek et al. 2004], Down syndrome, among others.

Deformation-based morphometry (DBM)

This techniques are developed in the computational anatomy framework where anatomical information is encoded in the deformations between a template and each subject. Each deformation is characterized by specifying a displacement for each voxel. On these displacements classical statistical methods are applied to quantify differences between groups. DBM can be performed globally [Davatzikos et al. 2001], *i.e.* using all the displacement field as a single observation, or as independent measures for each voxel allowing scenarios for local hypothesis contrasts [Chung et al. 2001] [Gaser et al. 2001]. It should be noted that the displacement field also includes information about position and size which is usually irrelevant for group differentiation. It is a common practice to remove this information by previously performing a rigid registration.

Tensor-based morphometry (TBM)

One way to solve the limitations of the DBM technique is by using local descriptors of the deformation instead of global displacements. Deformation mappings are approximated at each voxel by its Jacobian matrix and it is used as local feature of the deformation. The Jacobian matrix at each voxel describes the relative relation of the voxel with respect to its neighbors. Instead of using multivariate analyses of the whole Jacobian matrix, typically its determinant is used as scalar feature. This determinant has an intuitive and straightforward interpretation because it reflects the local volume change of the anatomy with respect to the anatomy of the template [Thompson et al. 2000a] [Fox et al. 2001].

TBM technique has been recently used in the study of diseases such as AD [Hua et al. 2009, 2008] [Bossa et al. 2009, 2010b] or in AIDS [Lepore et al. 2008].

1.4.3 Data for clinical validation

There exist several publicly available databases of neuroimaging for studies of morphometry in AD. Probably, one of the most complete and extensive is the project *Alzheimer's Disease Neuroimaging Initiative* (ADNI)¹ [Mueller et al. 2005] funded primarily by NIH (National Institutes of Health) belonging to U.S. Department of Health and Human Services. ADNI is a multicenter study with the aim of revealing clinical, genetic and neuroimaging aspects in early detection and monitoring of AD.

ADNI project began in October 2004 and includes 57 medical centers in U.S. and Canada. Its design covers a recruiting over 800 subjects aged between 55 and 90 years, of which 200 are AD diagnosed patients, 400 mild cognitive impairment (MCI) patients, and 200 are control subjects without neuropsychiatric diseases diagnosed. Regarding neuroimaging data, ADNI includes for each subject PET and MR images that are repeated approximately every 6 months.

In the ambitions of ADNI project is the design of standard acquisition neuroimaging and other biomarkers in blood or cerebrospinal fluid for the study of AD. Data collected by ADNI have allowed substantial progresses about the understanding of the disease.

Currently the project is ongoing (through its successors ADNI2 and ADNI-GO) and arriving at a total of about 1300 enrolled subjects.

Collected data are offered to scientific communities for the performance of studies in the following areas:

- Identify and define markers and biomarkers that best predict future cognitive decline and the conversion MCI/AD. These biomarkers shall serve as tools in early diagnosis and may be also used for patient selection in trials of therapeutic and preventive treatments.
- Identify most relevant genetic patterns in AD by performing sequencing analysis of genome-wide association.
- Perform hypothesis tests in structural MR images and research in increasing the statistical power of current methodologies.
- Perform hypothesis tests and regression models from structural changes on MRI and functional PET with evolving cognition states.
- Improve clinical trials process by developing simulations of several scenarios taking into account cognitive scores and specific markers for the selection criteria. These scenarios should allow the evaluation of *a posteriori* measurements and identify key statistical methodologies to assess effects of treatments.

In this thesis, collected ADNI data will be used to perform statistical analysis of brain images accompanied by relevant clinical variables and the diagnosis states.

Regarding technical characteristic of MR images, they were acquired with 1.5 Tesla scanner by using the MRI protocol developed for ADNI [Jack et al. 2008a]. For each subject, T1-weighted MRI scans were collected using a sagittal 3D magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence with voxel size $0.94 \text{ mm} \times 0.94 \text{ mm} \times 1.2 \text{ mm}$. Additionally, images are preprocessing including geometric distortion correction, bias field correction and geometrical scaling. The images were calibrated with phantom-based geometric corrections to ensure consistency among scans acquired at different sites. The pre-processed images can be downloaded from ADNI website.

¹ adni.loni.usc.edu

1.5 Outline of the thesis

The thesis can be structured into three main parts: the first one consisting of Chapters 2 and 3 lays the foundations of image registration; the second part consisting of Chapters 4 and 5 explores the formal mathematical setting of Riemannian geometry for the analysis of the different groups of spatial transformations and also gives some practical aspects and illustrative examples; the third part consisting of Chapter 6 makes use of the two previous pillars and proposes and performs several methodologies in the study of anatomical changes occurring in Alzheimer's disease.

Chapter 2 introduces the notation and concepts used through the rest of the thesis. The chapter begins with a formalization of the concepts of geometrical objects and spatial transformations. Section 2.2 put the focus on the definition of the geometrical objects with relevant interest in medical imaging namely unlabeled set of points, curves, surface, landmarks, scalar intensity fields. The action of spatial transformations is made explicit for each type of geometrical objects. Furthermore, practical aspects of the required numerical computations are given, in particular for scalar intensity fields. Section 2.3 introduces the concept of spatial transformation models, establishes their uses for futures optimizations and defines three categories of spatial transformation models. Section 2.4 considers the sets of spatial transformation with a group structure. Most of the commonly used spatial transformations groups for medical imaging purposes are linear groups and can be represented by matrices. The matrix representation provides a great simplification in the analysis of interesting groups. Chapter 2 finish by giving a taxonomy of several spatial transformation models which provides some properties and makes their actions explicit.

Chapter 3 is completely devoted to the image registration process. The chapter includes a revision of several similarity and regularity measures commonly used in medical imaging. Registration procedure for different categories of spatial transformation models are gathered. Section 3.4 considers parametric models such as, similarity transformations or special linear transformations. Mostly, two types of objects are considered, landmarks and images. While for simple transformation models algebraic expressions can be derived to compute the transformation, most of the procedures must be carried out by iterative numerical optimization techniques. The concept of spatial forces is introduced and illustrated on sparse landmarks as well as on dense scalar images. Section 3.5 treats the case of object-dependent models where a regularization technique must be usually considered. The regularization is presented in two settings: the exact and the inexact matching formulations. For the case of sparse parametrizations, the technique of Green's function, which can be applied to both formulations, is presented. In Section 3.6, a diffeomorphic registration with stationary velocity fields is presented as an example of non-parametric spatial transformation models. The computation of derivatives of the matching functional is derived and some practical aspects are covered.

In Chapter 4, the mathematical Riemannian setting is overviewed with special interest in the formal definition of invariant geodesics and distances on Lie groups. In Section 4.2, a formal background is provided with the main purpose of bridging the gap between image analysis and differential geometry. Several approaches to compute left-invariant Riemannian geodesics on Lie groups are gathered in Section 4.3. Some other relevant aspects of these geodesics are considered through Sections 4.4 to 4.8.

Once the formal characterization of the Riemannian setting has been presented, Chapter 5 focus on the practical aspect for the computation of invariant Riemannian geodesics on Lie groups. The chapter starts by gathering closed-form solutions for several groups with practical interest. For the cases when a closed-form is not known, Section 5.2 proposes some simple numerical integrators valid for any matrix group and under any Riemannian metric. The Section 5.2 also

deals with the Riemannian logarithm function. Section 5.3 evaluates the performance of the proposed numerical integrators in terms of accuracy, computational timing and preservation of group structure. Finally, the chapter ends with some simple applications illustrating the use of the Riemannian setting for the analysis of spatial transformations: Section 5.4.1 shows invariant interpolations between spatial transformations; the intrinsic mean of spatial transformations is posed and analyzed in Section 5.4.2; Section 5.4.3 shows intrinsic filtering of spatial distributed Jacobian matrices; and, Section 5.4.4 illustrates the use of the conjugacy action which allows to extend closed-form solutions to some other metrics.

The image registration tools and the Riemannian setting on spatial transformation is exploited in practical applications from the neuroanatomy field. In particular, anatomical changes suffered during the Alzheimer’s disease progress are described and evaluated in Chapter 6. Based in the computational anatomy framework, three methodologies for the analysis of the geometry and anatomy are conducted. All studies are performed on images from the ADNI database. Section 6.1 proposed the extension of the classical volumetry on brain structures to a multivariate description of the geometrical changes. The *pose* of subcortical structures is analyzed by geodesic regressions with respect to ADAS-cog score and age. Section 6.2 revisits the TBM methodology for the study of the whole brain anatomy. Several aspects of the method are discussed, from the construction of an anatomical template to the statistical assessment of the results. Finally, Section 6.3 presents a new TBM methodology which analyzes the full description of the local deformation. The methodology is based on a distance-based statistic and its performance is shown over three features of Jacobian matrices namely the determinant, the Cauchy–Green deformation tensor and the full Jacobian matrix analyzed with an invariant Riemannian metric. For each of the studies analyzed in the chapter, several practical aspects are included when the methodologies are presented.

CHAPTER

2

Background

2.1	Geometrical objects and spatial transformations	24
2.2	Action of spatial transformations on geometrical objects	26
2.2.1	Subsets of the ambient space	26
2.2.2	Scalar functions	28
2.2.3	Examples of actions which modify attribute values	31
2.2.4	Combined objects	32
2.2.5	Spatial transformations as geometrical objects	32
2.3	Parametrization of the action of spatial transformations	33
2.4	Spatial transformation groups	35
2.4.1	Matrix groups and matrix representations	35
2.4.2	The action of matrix groups on spatial points	37
2.5	Spatial transformation models	38
2.5.1	Discrete spatial transformations models	38
2.5.2	Finite dimensional models	38
2.5.3	Object-dependent transformation models	43
2.5.4	Infinite dimensional spatial transformation models	43

This chapter has the intention to give a formal description of the key elements involved in the registration and analysis of geometrical objects. Furthermore, these concepts will be used through the rest of the thesis.

Let us start by explaining what a geometrical object is and later by defining how its geometry can be modified by spatial transformations. The action of a spatial transformation on several types of geometrical objects will be explicitly given. Later, concepts and properties of spatial transformations sets will be introduced. Spatial transformations play an important role in the fields of computer graphics, computer vision and image analysis. For example, the use of spatial transformation is a useful and flexible tool in object modelling [Barr 1984] [Mortenson 1997] [Angelidis and Singh 2006]. Spatial transformations can be used for object alignment as well as to define an interpolation scheme that provides intermediate object positions [Shoemake 1985] [Park and Ravani 1997] [Alexa 2002] [Li and Hao 2006]. In shape analysis, spatial transformations are key ingredients for the definition of pose and shape [Dryden and Mardia 1998]. Image registration and pose recovery are often defined as optimization problems on the space of a given spatial transformation model [Modersitzki 2004] [Berkels et al. 2010] [Pham et al. 2011]. In the field of computational anatomy [Thompson 1917] [Grenander and Miller 1998] [Miller et al. 2002] the anatomical information is encoded in the spatial transformation that maps a given atlas to a target image.

For practical and computational purposes, it is interesting to identify any member from a family of spatial transformations of interest. This identification will define spatial transformation models which are useful tools to parameterize the action of spatial transformations.

It is very common in applications to be interested on spatial transformations which preserved certain geometrical attributes of the objects. For example, it can be of interest to keep straight and parallel lines, the distances between points, or their neighboring and incidences relations. Those magnitudes and relations usually define *symmetries* in the objects. The set of all spatial transformations which preserve a symmetry in the objects usually can be endowed with a group algebraic structure. Moreover, these set of spatial transformations are very often Lie groups which allows the definition of topologies, curves and distances between spatial transformations and will be studied later in Chapters 4 and 5.

In order to describe a general framework to be capable to deal with different descriptions of geometrical contents, such as point sets, landmarks, curves, or images, a set-theoretic approach will be used. Thus, many of the concepts defined in the following will be defined using a set-theoretic notation. Whenever the order of the elements matters the notation $(\cdot, \cdot, \dots, \cdot, \dots)$ will be used to denote the set, *i.e.* elements enclosed with rounded braces. In the case that every element depends on an index, it and often its range will be made explicit, for example $(a_i)_{i \in \mathbb{N}}$ or $(a_i)_{i=2}^8$. In the cases where the sets are just a collection of elements, without a predetermined ordination, the notations $\{\cdot, \cdot, \dots\}$ or $\{a_i\}$ will be used, *i.e.* by the use of curly braces. As before, to remark the dependence with an index, $\{a_i\}_i$ will be used, and to specify the range of an index, $\{a_i\}_{i \in \mathbb{K}}$ or $\{a_i\}_{i=7}^9$ or $\{a_i \mid i \in \mathbb{K}\}$.

2.1 Geometrical objects and spatial transformations

Let us consider an *ambient space* \mathbb{A} where some geometrical information is described. A general characterization of a valid geometrical ambient space is beyond this thesis and throughout this thesis \mathbb{A} will be always considered as the d -dimensional Euclidean space \mathbb{R}^d , for d typically equal to 2 or 3. It could be of interest in a further research to consider different ambient spaces such as spheres [Bakircioglu et al. 1999] [Tao et al. 2002] [Glaunès et al. 2004b] [Zou et al. 2007] [Yeo et al. 2010] [Lombaert et al. 2013] or manifold spaces in general [Chefd'hotel 2007] [Hirsch 2011].

The elements of the ambient space are called *spatial points*, or just *point* for short, and are usually denoted by x, y, z . On the space \mathbb{A} a d -dimensional coordinate system can be defined and once it is selected each spatial point can be described by its coordinate in a d -dimensional column-array, such that $y = (y^1, y^2, \dots, y^d)^T \in \mathbb{A}$, where superscript \cdot^c refers to the c -th coordinate.

The ambient space \mathbb{A} is the support domain where geometrical objects are defined. In a general way, a *geometrical object* B describing some geometrical content is an assignment of an attribute value $b_y \in \mathbb{F}$ to each point of the ambient space, where \mathbb{F} is the set of allowed attributes. A geometrical object can be described as a collection of *location-attribute* pairs: $B = \{(y, b_y) \mid y \in \mathbb{A}\}$. Additionally, a *projection operator* ϖ_y can be defined such that when it is applied on B returns the attribute of the location-attribute pair having y as its first element, *i.e.* $\varpi_y(B) = b_y$. Using ϖ_y it is possible to compare two geometrical objects: let B_1 and B_2 be two geometrical objects, they describe the same geometrical content if both are the same collection of location-attribute pairs, or alternatively if $\varpi_y(B_1) = \varpi_y(B_2)$ for all $y \in \mathbb{A}$.

Geometrical objects can be *modified* by *spatial transformations*. A *spatial transformation* is a bijective mapping between the elements of the ambient space \mathbb{A} into itself, $h: \mathbb{A} \rightarrow \mathbb{A}: y \mapsto h(y)$. Eventually the spatial transformation h can be defined as a function from a smaller space $\mathbb{A}' \subset \mathbb{A}$ into itself and extended to \mathbb{A} by setting that $h(x) = x$ if $x \notin \mathbb{A}'$. In the presented framework, spatial transformations act by modifying the whole ambient space and the geometrical objects defined on \mathbb{A} are modified accordingly in such a way that the action of a spatial transformation h on an object consists in the action on each point of the object.

Spatial transformations must preserve the integrity of the anatomical structures. This is obtained by requiring to the spatial transformations that do not produce disconnections, holes, or folds in the ambient space. The first of these requirements is formalized by considering spatial transformations as continuous mapping from \mathbb{A} to \mathbb{A} . The second requirement is formalized by considering only surjective mappings and the third one by injective mappings. Therefore, spatial transformations are defined as continuous and bijective mappings from \mathbb{A} on itself. Under the requirements of continuity and bijectivity spatial transformations results in *homeomorphic mappings* of the ambient space and the set of all spatial transformations results in the set of all homeomorphisms which will be denoted by $\mathcal{H}_{\mathbb{A}} = \{h \mid \text{such that } h: \mathbb{A} \rightarrow \mathbb{A} \text{ is a homeomorphism}\}$.

Some basic properties of spatial transformations are described below. Obviously, the identity map of \mathbb{A} , $\text{id}_{\mathbb{A}}: \mathbb{A} \rightarrow \mathbb{A}: y \mapsto \text{id}_{\mathbb{A}}(y) = y$, is a spatial transformation. Because each spatial transformation h is a bijection, then its *inverse transformation*, which is denoted by h^{-1} , always exists and satisfies that if $x = h(y)$ then $y = h^{-1}(x)$. Finally, given two spatial transformations h and q over \mathbb{A} , then $q \circ h$ is also a spatial transformation over \mathbb{A} , where $(\cdot \circ \cdot)$ refers to the function composition. Using these properties it can be shown that the set of all spatial transformations over \mathbb{A} has a *group structure* under the composition (and inversion) of functions and its identity element is the identity map $\text{id}_{\mathbb{A}}$.

When a spatial transformation h is applied, every point $y \in \mathbb{A}$ of the ambient space “is moved” to the new location $h(y)$. In order to remark the geometrical nature of the formulation, this operation will be called *action* of h on y and will be denoted as $h \star y$. With the same aim of highlighting the geometrical point of view, the composition of two spatial transformation will be denoted as $h_2 \bullet h_1$ such that $h_2 \star (h_1 \star y) = (h_2 \bullet h_1) \star y$ for all y , instead of the classical functional viewpoint¹ $(h_2 \circ h_1)(y) = h_2(h_1(y))$. In the language of group theory, this operation is called *composition* of h_2 and h_1 (along different literature sources it is also called *product law*, *group law*, *group multiplication*). The identity element of the group structure will by

¹ Let $\mathcal{F}_{\mathbb{A}}$ be the space of functions from \mathbb{A} to \mathbb{A} , the function composition operation is defined as $\circ: \mathcal{F}_{\mathbb{A}} \times \mathcal{F}_{\mathbb{A}} \rightarrow \mathcal{F}_{\mathbb{A}}: f(\cdot) \circ g(\cdot) \mapsto f(g(\cdot))$. Instead, the operation $(\cdot \bullet \cdot)$ is defined as the restriction of $(\cdot \circ \cdot)$ to the set of homeomorphisms $\mathcal{H}_{\mathbb{A}}$ such that $\bullet = \circ|_{(\mathcal{H}_{\mathbb{A}} \times \mathcal{H}_{\mathbb{A}})}: \mathcal{H}_{\mathbb{A}} \times \mathcal{H}_{\mathbb{A}} \rightarrow \mathcal{H}_{\mathbb{A}}$.

denoted by e , which corresponds to $\text{id}_{\mathbb{A}}$, and the group inverse will be denoted by $\text{inv}(\mathfrak{h})$, which corresponds to the inverse function \mathfrak{h}^{-1} . This geometrical formulation infers to the spatial transformations an abstract mathematical nature over to which some algebraic operations will be defined irrespectively than the parameterizations and models chosen to describe it.

The geometrical information defined in a geometrical object changes when a spatial transformation is applied to the ambient space. It is fundamental to specify how a spatial transformation modifies a geometrical object defined in \mathbb{A} , such that the transformed object describes the same geometrical content and information in the transformed ambient space than the original object in its original coordinates. Given a geometrical object B defined in \mathbb{A} , the spatial transformation \mathfrak{h} acts on B and this operation is denoted by $\mathfrak{h} \star B$. This notation is extended from the action on points of the ambient space $\mathfrak{h} \star y = \mathfrak{h}(y)$.

For simplicity, it will be typically assumed that the action of a spatial transformation does not modify the value of the attributes. This condition is achieved in many practical cases that will be discussed along the thesis. However, a general definition of the action can be considered into account for more “complex” geometrical objects, such as spatial probability distributions (see for example in Section 2.2.3). There are two formulations to define the action of a spatial transformation on a geometrical object: the *Lagrangian formulation*; and the *Eulerian formulation*. Both are nicely defined in terms of how to transform a location-attribute pair (y, b_y) . In the Lagrangian formulation a spatial transformation \mathfrak{h} acts on the location keeping fixed the attribute value, $\mathfrak{h} \star (y, b_y) = (\mathfrak{h}(y), b_y)$. By contrast, the Eulerian formulation focuses on how the attribute value is changed in a fixed location of the ambient space, $\mathfrak{h} \star (y, b_y) = (y, b_{\mathfrak{h}^{-1}(y)})$.

2.2 Action of spatial transformations on geometrical objects

The action of a spatial transformation on a single point of the ambient space is $\mathfrak{h} \star y = \mathfrak{h}(y)$. However, we are interested in how spatial transformations act on different types of spatial objects commonly used in the medical imaging area. In the following, explicit expressions are given for the action on several interesting types of geometrical objects.

2.2.1 Subsets of the ambient space

The most simple type of geometrical object is *unlabeled point subsets* of the ambient space, $P = \{y_i\}_i$. The location-attribute description of these objects is given by considering $\mathbb{F} = \{0, 1\}$ (or equivalently $\mathbb{F} = \{\text{FALSE}, \text{TRUE}\}$) where 0 is assigned to points not belonging to the object and 1 to points belonging to the object. The action of a spatial transformation \mathfrak{h} on P is explicitly given by $\mathfrak{h} \star P = \{\mathfrak{h} \star y_i\}_i$.

The same framework can be used to represent curves² in the ambient space. In this case the set of points is continuous and can be indexed with a continuous parameter. For example, let $\gamma: [0, L] \subset \mathbb{R} \rightarrow \mathbb{A}$ be an injective smooth function, the associated geometrical object is $C = \{\gamma(t) | t \in [0, L]\}$. The action of a spatial transformation is naturally given by $\mathfrak{h} \star C = \{\mathfrak{h} \star \gamma(t) | t \in [0, L]\}$.

Some curves can be described by a mathematical expression or formula in terms of the variable t . However, to describe an arbitrary curve in a computer all points should be specified which would require an infinite amount of memory. Instead, a classical representation of a curve is to

² Throughout the thesis a *curve* will be considered as a continuous function from a non-empty connected subset of \mathbb{R} . Moreover, unless otherwise is stated, it will be assumed that the function is smooth, *i.e.* a function with well-defined derivatives up to any order (C^∞ functions). It is important to remark the distinction between a curve and its image, *i.e.* the set of points through which the curve passes. In the cases when we are referring to the image of an arbitrary curve, we will refer as the *path* or eventually as the *trajectory*.

define a sequence of ordered points called *nodes* (y_0, y_1, \dots, y_k) from where the curve pass through at increasing times $(t_0 = 0, t_1, \dots, t_k = L)$ together with an interpolation scheme to define the curve at an arbitrary time $t \in [t_i, t_{i+1}]$. The simplest example is a polygonal curve (called *polyline* in the area of computer graphics) joining consecutive nodes by straight segments. More flexible descriptions include polynomially expressed segments [Farin 2001], rational expressions, circular arcs [Meek and Walton 2004], or exponentially parametric curves [Kimia et al. 2003] [Zhou et al. 2012] joining consecutive nodes. For a sufficient large number of nodes sampling a given curve, a spatial transformation acts on the nodes and the transformed curve can be described by using the same representation. For example, let C be a polyline parameterized with nodes (y_0, y_1, \dots, y_k) at the times (t_0, t_1, \dots, t_k) . For a given t , such that $t_i \leq t \leq t_{i+1}$, the corresponding point of the curve is $C_{(t)} = \left(\frac{t_{i+1}-t}{t_{i+1}-t_i}\right) y_i + \left(\frac{t-t_i}{t_{i+1}-t_i}\right) y_{i+1}$. The action of a spatial transformation \hbar on $C_{(t)}$ is approximated by $\hbar \star C_{(t)} = (\hbar \star C)_{(t)} \approx \left(\frac{t_{i+1}-t}{t_{i+1}-t_i}\right) (\hbar \star y_i) + \left(\frac{t-t_i}{t_{i+1}-t_i}\right) (\hbar \star y_{i+1})$.

Using a similar approach also surfaces in a three dimensional space can be represented as a set of polygons that are connected by their edges and vertices. The most common surface representation is as a *triangle mesh* which comprises a set of ambient space points corresponding to the vertices of triangles and a list of triplets to specify the set of triangles that represent the surface [Dey and Levine 2008] [Smith 2012]. Triangle meshes are typical result of an isosurface process on a binary or gray valued image [Lorensen and Cline 1987] [Akkouche and Galin 2001] [Schreiner et al. 2006]. There also exist more complex and flexible computational descriptions of surfaces, some of them include polynomial or rational interpolations [Hamann et al. 1997] between vertices. A triangle mesh is given by the pair $S = (V, T)$, where $V = (y_i)_i$ is set a of vertex coordinates and $T = \{(v_t^1, v_t^2, v_t^3)\}_t$ is a set of triplets indicating an adjacency structure between the vertices. Regarding the action of a spatial transformation, it is usually enough to act on vertex set and preserve the adjacency. Thus, the action of a spatial transformation \hbar on $S = (V, T)$ is given by $\hbar \star S = ((\hbar \star y_i)_i, T)$.

Point subsets are also useful to describe a *geometrical region* which is a set of connected and disjoint subsets of the ambient space. Let R be a geometrical region defined in \mathbb{A} and let $\mathbf{1}_R: \mathbb{A} \rightarrow \{0, 1\}$ be its *indicator function* such that $\mathbf{1}_R(y)$ is equal to 1 if y belongs to the region R and 0 elsewhere. In the location-attribute description $R = \{(y, \mathbf{1}_R(y)) | y \in \mathbb{A}\}$ and therefore $\varpi_y R = \mathbf{1}_R(y)$. The action of a spatial transformation \hbar on R is naturally expressed by the action of \hbar on the indicator function, corresponding to the Eulerian formulation of the action. Therefore, $\mathbf{1}_{\hbar \star R}(y) = \mathbf{1}_R(\hbar^{-1} \star y)$. The previous identity is valid for any $y \in \mathbb{A}$, and therefore it can be expressed by the functional equality $\mathbf{1}_{\hbar \star R} = \mathbf{1}_R \circ \hbar^{-1}$.

A geometrical region can be alternatively defined by its boundary. For the case of $d = 2$ the boundary of a region results in a set of non-intersecting simple and closed planar curves. The action of a spatial transformation on a region directly acts on their boundary curves, corresponding to the Lagrangian formulation of the action. Therefore it is practical to parameterize regions by their boundaries. There also exist efficient algorithms to determine if a point is inside or outside a closed planar curve [Hormann and Agathos 2001]. Moreover, using an appropriate convention (for example given an orientation to the curves) this representation can also be used to define regions with holes [Liu et al. 2010]. The extension to a three dimensional region is relatively straightforward. An important property of triangle meshes is whether the mesh is closed, *i.e.* if the surface it represents completely encloses a volume. A simple connected region can be described by the interior of a closed surface. There exist efficient algorithms to evaluate at which side of closed triangle mesh a point lies [Baerentzen and Aanaes 2005]. Furthermore, by including a convention defining the normals of the surfaces [Borodin et al. 2004] it is possible to represent regions which are not simple connected.

In the previous examples, the points defining the geometrical objects are assigned with an attribute equal to 1 and 0 to the rest. However, a more complete information is obtained when each point can be assigned with a label representing some high-level knowledge about the anatomy. For example, the case of points subsets provided with anatomical labels is common in shapes analysis [Thompson 1917] [Small 1996] [Bookstein 1997] [Heo and Small 2006] and computational anatomy. These points are called *landmarks* and correspondences between points with the same label enrich the process to measure the geometrical differences. A geometrical object defined by landmarks is given in terms of location-attribute pairs by $L = \{(y_1, 1), (y_2, 2), \dots, (y_L, L)\} \cup \{(x, 0) | x \in \mathbb{A} - \{y_i\}_i\}$.

Another interesting example is given by structures of the brain cortex [Ono et al. 1990]. In the last years there have been important advances in the automatic delineation and labeling of the gyri and sulci structures in brain MRI [Mangin et al. 2004] [Perrot et al. 2009] [Perrot et al. 2011] [Fischl et al. 2002, 2004]. This labeling provides a rich information about the brain anatomy and have proved to be very useful to improve the geometrical matching between subjects [Khan et al. 2008] [Auzias et al. 2011] [Gori et al. 2013].

2.2.2 Scalar functions

Let $\mathbf{f}: \mathbb{A} \rightarrow \mathbb{R}$ be a scalar function defined on the ambient space (eventually defined on a smaller subset and somehow extended to the whole \mathbb{A}). The scalar function $\mathbf{f}(\cdot)$ can be considered as a geometrical object, $F = \{(y, \mathbf{f}(y)) | y \in \mathbb{A}\}$, where its attribute space $\mathbb{F} = \text{imag}(\mathbf{f})$. A query of its value at a location $y \in \mathbb{A}$ is given by $F_{(y)} = \varpi_y(F) = \mathbf{f}(y)$.

As in the case of the indicator function of a geometrical region, the action of a spatial transformation is easily described using the Eulerian formulation. Therefore, given a spatial transformation h , the value of the transformed function $h \star F$ at the point y can be obtained by $(h \star F)_{(y)} = \varpi_y(h \star F) = \mathbf{f}(h^{-1} \star y)$. As it is valid for every point in \mathbb{A} , while the original object F is associated with the original function \mathbf{f} , the transformed object $h \star F$ is associated with the function $\mathbf{f} \circ h^{-1}$.

Practical aspects for representation of images

Now some practical aspects of scalar functions as geometrical objects will be discussed. There are different ways to describe a scalar function \mathbf{f} but a finite representation or parametrization is needed to allow numerical computations with its values. A function in \mathbb{A} can be defined as a mathematical expression valid anywhere or via several mathematical expressions for different sub-domains. However, this description does not give much flexibility to describe real-world images. Also, a wide range of functions can be described in terms of their level sets [Caselles et al. 2008] [Caselles and Monasse 2010] where each level set is a connected domain which can be defined by the interior of a polygonal curve (see Section 2.2.1). Even though this representation deals well with synthetic or cartoon images [Mumford and Shah 1989] it is not convenient for photographs or acquired images.

The main interest of this thesis falls in the analysis of anatomical images acquired for medical purposes. As modern digital photographs which represent a two-dimensional spatial distribution of lighting intensities and are digitized into small and structured homogeneous regions of the space (widely known as pixels), medical images are typically described as a spatial distribution of intensity values in a three-dimensional structured partition of the space into homogeneous cells. These cells are commonly called *voxels* and their values typically quantify a physical magnitude measured in a small region of the physical space. Unlike photographs, it is common that medical images also contain information about the spatial location, orientation or attitude in a calibrated system of reference with respect to the acquisition device. It is also common that medical images

have non isotropic voxel sizes. The description as a structured grid partitions with scalar values associated to each cell must be understood as the measurement of a certain physical property of an underlying biological tissue at a small location of space.

For many purposes, such as image registration, it is useful to consider medical images as continuous spatial distributions of the measured properties of the underlying tissue, instead of their discrete data nature. A three-dimensional image G will be described as a structured grid given by the Cartesian product of three strictly monotone arrays $(X_i)_{i=1}^I$ spanning an horizontal direction, $(Y_j)_{j=1}^J$ for a vertical direction and $(Z_k)_{k=1}^K$ for an in-depth direction. The discrete locations defined by the grid correspond to spatial locations of the form $y_{ijk} = (X_i, Y_j, Z_k)^T$. To each location the acquired intensity value is assigned and these are stored in a $I \times J \times K$ array G such that $G_{(y_{ijk})} = G_{i,j,k}$. Alternatively, the intensity data array can be *vectorized*, *i.e.* stacked or linear-indexed, in a linear array \bar{G} of length IJK such that $G_{(y_{ijk})} = \bar{G}_{i+(j-1)I+(k-1)IJ}$. This storage in \bar{G} is commonly referred as *lexicographic ordering* [Modersitzki 2004].

In order to query for the image value at any spatial location, an interpolation function and an extrapolation rule must be specified. For an arbitrary point $y = (y^1, y^2, y^3)^T$, lying within the grid domain, first of all its neighboring grid locations must be determined by looking the corresponding interval i , j or k in the arrays (X_i) , (Y_j) and (Z_k) such that

$$\begin{aligned} X_i &\leq y^1 < X_{i+1} \\ Y_j &\leq y^2 < Y_{j+1} \\ Z_k &\leq y^3 < Z_{k+1}. \end{aligned}$$

This search can be performed in an efficient manner by a recursive binary search independently for each coordinate y^c . Then the value is computed via the interpolation function $G_{(y)} = \text{interpolation}(y; (\mathcal{N}_{ijk}))$, where (\mathcal{N}_{ijk}) is a list of the grid nodes neighboring the ijk -th cell.

There are several interpolation functions [Parker et al. 1983] [Maeland 1988] [Wahba 1990] [Unser et al. 1993] [Appledorn 1996] [Dodgson 1997] [Caselles et al. 1998] [Lehmann et al. 1999] [Thévenaz et al. 2000] [Blu et al. 2004] [Wang and Ward 2007] [Ramani et al. 2010] but in this thesis we will concentrate on linear combinations of basis functions

$$G_{(y)} = \sum_{\ell \in (\mathcal{N}_{ijk})} w_\ell(y) G_\ell,$$

where ℓ is a triplet identifying a grid node and $w_\ell(y)$ are the weights corresponding to each value G_ℓ . Each of the values $w_\ell(y)$ can be obtained by evaluating the corresponding basis function at point y .

For a point outside the grid domain an extrapolation rule should be specified. Examples of extrapolation rules are periodic boundary conditions, symmetric boundary conditions, a constant value, or some other specified function.

With the previous representation, an image results in a dense scalar function from the whole ambient space which inherits the properties of the basis functions. In the early days of volumetric data processing, the continuous description was achieved by nearest-neighbor interpolation where every sample point within a voxel is assigned the intensity value of that voxel. However, it is common to use continuous basis functions with or without continuous derivatives and more sophisticated interpolation schemes are typically used. The most popular scheme is by considering multi-linear basis functions such that, in 3D, eight neighboring grid locations define

a “cuboid” where the interpolated intensity value within the “cuboid” is

$$\begin{aligned}
 G_{(y)} = & (1 - \alpha_1) (1 - \alpha_2) (1 - \alpha_3) G_{i,j,k} \\
 & + \alpha_1 (1 - \alpha_2) (1 - \alpha_3) G_{i+1,j,k} \\
 & + (1 - \alpha_1) \alpha_2 (1 - \alpha_3) G_{i,j+1,k} \\
 & + \alpha_1 \alpha_2 (1 - \alpha_3) G_{i+1,j+1,k} \\
 & + (1 - \alpha_1) (1 - \alpha_2) \alpha_3 G_{i,j,k+1} \\
 & + \alpha_1 (1 - \alpha_2) \alpha_3 G_{i+1,j,k+1} \\
 & + (1 - \alpha_1) \alpha_2 \alpha_3 G_{i,j+1,k+1} \\
 & + \alpha_1 \alpha_2 \alpha_3 G_{i+1,j+1,k+1},
 \end{aligned} \tag{2.2.1}$$

where $\alpha_1 = \frac{y^1 - X_i}{X_{i+1} - X_i}$, $\alpha_2 = \frac{y^2 - Y_i}{Y_{i+1} - Y_i}$, and $\alpha_3 = \frac{y^3 - Z_i}{Z_{i+1} - Z_i}$, i.e. are the isoparametric coordinates within the cell. Others schemes involve higher-order interpolations functions which take a larger cell neighborhood into account. There is a trade off between the simplicity (and the computational efficiency of the interpolation step) and the smoothness of the image model. For example, the piecewise multi-linear scheme given in Eq. (2.2.1) is simple and fast to compute but the lack of differentiability can be an issue for some applications. Cubic B-spline scheme are commonly used to retrieve this issue being able to model the images as continuous and differentiable functions [Hou and Andrews 1978] [Keys 1981] [Kadosh et al. 2003] [Lekien and Marsden 2005].

Additionally to the definition of the Cartesian grid by the arrays (X_i) , (Y_j) and (Z_k) , a spatial transformation h can be specified allowing to place the image content in a real world coordinate system. Then a grid point y_{ijk} is spatially located at $h \star (X_i, Y_j, Z_k)^T$. For an image G defined in a Cartesian grid (that is, aligned with the coordinate system) and a point y in \mathbb{R}^d , the value $G_{(y)}$ can be computed by the interpolation scheme explained before where the neighboring set can be efficiently computed by independent *binary searches*. When the grid nodes are transformed the cell search procedure becomes a more complex and costly task [Knabner et al. 2003] [Zhang 2005a,b]. However, by using the Eulerian formulation $(h \star G)_{(y)} = G_{(h^{-1} \star y)}$ and this evaluation can be efficiently handled if $h^{-1} \star y$ is known.

We have implemented a very efficient data structure for an image container as a structured Cartesian grid with nodal values at each grid point, together with an interpolation function, an extrapolation rule and equipped with an affine spatial transformation (see later Section 2.5.2) to place the image in real world coordinates. The inverse of the affine transformation is stored in the data structure as a 4×4 matrix. This image container allows to easily manage the kind of data usually managed in the medical image area which are a collection of voxels, typically with non-isotropic voxel sizes, with an intensity value associated to each voxel, and usually defined for a patient attitude and position. Using this image container, a classical $M \times N$ two dimensional scalar image can be represented by setting $(Z_k) = (0)$, $(X_i) = (1, \dots, N)$, $(Y_j) = (1, \dots, M)$ and the spatial transformation equals to the identity. Also, the data structure allows to store dynamic or temporal data like videos or time dependent medical images like, for example, dynamic perfusion studies. Moreover, multidimensional data can be stored in each voxel in order to be able to represent RGB images, vectors, or more complex medical data such as *diffusion tensor images*. In these cases, vectorial data is treated as independent scalar images where each component is interpolated independently.

Image gradients

The computation of gradients of an image is a fundamental tool in the image registration process. The gradient of a scalar function G is a vector-valued function $\nabla G: \mathbb{A} \rightarrow \mathbb{R}^d$. Typically, the

gradient at a point within a grid cell is computed as the multi-linear interpolation of the gradients computed at grid locations by central finite differences scheme. However, to improve convergence of registration algorithms, it is better to define the gradient as the spatial derivatives of the interpolated image.

The gradient of an image G at the point y is

$$(\nabla G)_{(y)}^T = \mathcal{D}_z G_{(z)} \Big|_{(z=y)},$$

where the last term is the *Fréchet derivative* of G at the point y , which is the linear operator satisfying $(\mathcal{D}G)_{(y)}\bar{w} = \lim_{\epsilon \rightarrow 0}(G_{(y+\epsilon\bar{w})} - G_{(y)})/\epsilon$, for all $\bar{w} \in \mathbb{R}^d$. Alternatively, the Fréchet derivative can be defined by the operator providing the linear approximation around a point $G_{(y+\epsilon\bar{w})} = G_{(y)} + \epsilon(\mathcal{D}G)_{(y)}\bar{w} + O(\epsilon)$.

The gradient can be computed almost everywhere if a multi-linear interpolation is used (Eq. (2.2.1)). At those points where the interpolated function is not differentiable, such as the boundaries of the grid cells, a central difference scheme can be used as approximation of a subdifferential of the image [Bonnans et al. 2006].

For a deformed image, by using the chain rule, the gradient results

$$\begin{aligned} (\nabla(h \star G))_{(y)}^T &= \mathcal{D}_z(h \star G)_{(z)} \Big|_{(z=y)} \\ &= \mathcal{D}_z G_{(h^{-1} \star z)} \Big|_{(z=y)} \\ &= (\nabla G)_{(h^{-1} \star y)}^T (\mathcal{D}h^{-1})_{(y)}, \end{aligned}$$

where in this case $(\mathcal{D}h^{-1})$ is called the Jacobian of h^{-1} given by $(\mathcal{D}q)_{(y)} = \mathcal{D}_z(q \star z) \Big|_{(z=y)}$.

2.2.3 Examples of actions which modify attribute values

Up to now, the action of spatial transformations does not modify the attribute value the location-attribute pairs defining geometrical objects. Although it will not be used in the thesis, let us exemplify two types of geometrical objects which require to changing the attribute values in the Lagrangian formulation of the action.

In the first example, let us consider an edge detection process $Z(\cdot)$ such that for a given scalar function G , the edge detection result $Z(G): \mathbb{A} \rightarrow \mathbb{R}: y \mapsto \|(\nabla G)_{(y)}\|^2$, i.e. the Euclidean norm of the gradient of the image as a strength measure of the edges. Although, $Z(G)$ can be viewed as a scalar function, it is more interesting to consider that the action of a spatial transformation h on $Z(G)$ as the action on the underlying image, $h \star Z(G) = Z(h \star G)$. Then, the resulting value at a point $y \in \mathbb{A}$ results

$$\begin{aligned} (h \star Z(G))_{(y)} &= (Z(h \star G))_{(y)} \\ &= \|(\nabla(h \star G))_{(y)}\|^2 \\ &= (\nabla G)_{(h^{-1} \star y)}^T (\mathcal{D}h^{-1})_{(y)} (\mathcal{D}h^{-1})_{(y)}^T (\nabla G)_{(h^{-1} \star y)} \\ &\neq (Z(G))_{(h^{-1} \star y)}. \end{aligned}$$

Then, the action on a location-attribute pair is more complex than the simple Eulerian formulation of the action on a scalar function.

The second example can be found in [Younes et al. 2009]. Consider a geometrical object representing a spatially distributed density function $D: \mathbb{A} \rightarrow \mathbb{R}^+$. As D is a density function,

it is expected that integrations on every open neighborhood are preserved under spatial transformations. Then, for every open $\Omega \subseteq \mathbb{A}$ and any spatial transformation h

$$\int_{y \in \Omega} D(y) = \int_{z \in h \star \Omega} (h \star D)(z).$$

It can be shown that the previous holds if $(h \star D)(y) = \det((\mathcal{D}h^{-1})(y)) D_{(h^{-1} \star y)}$. In terms of location-attribute pairs, $D = \{(y, d_y) \mid y \in \mathbb{A}\}$, and the Lagrangian formulation of the action is $h \star (y, d_y) = [h \star y, \det((\mathcal{D}h^{-1})(y)) d_y]$.

2.2.4 Combined objects

There is an emerging interest in the registration of *combined information*, where on the same spatial domain different and complementary types of contents are described. In [Johnson and Christensen 2002] the MRI intensity and landmark information are simultaneously registered. [Postelnicu et al. 2009] proposes a registration guided by volumetric and cortical surface information. Similarly, [Du et al. 2011] uses at the same time the geometrical information from intensity values, cortex surfaces, sulcal lines and manually placed anatomical landmarks. Moreover, [Gori et al. 2013] includes also the neural tracts obtained from a tractography on diffusion tensor images. In those cases, where different geometrical information is defined in the same ambient space, we need apply a spatial transformation to all of them simultaneously.

Given different types of geometrical objects, for example points sets P , landmarks L , curves C_1 and C_2 , intensity values F , we can describe a combined object B simply by fusing all of them

$$B = (P, L, C_1, C_2, F, \dots)$$

and the action of a spatial transformation h is simply given by

$$h \star B = (h \star P, h \star L, h \star C_1, h \star C_2, h \star F, \dots).$$

Regarding its location-attribute representation, let \mathbb{F}_P be the space of attributes of the object P , the combined object $B = (P, L, C_1, C_2, F, \dots)$ can be described as the collection of location-attribute pairs of the form (y, b_y) where b_y belongs to $\mathbb{F}_P \times \mathbb{F}_L \times \mathbb{F}_{C_1} \times \dots \times \mathbb{F}_F \times \dots$. Explicitly the attribute value at the location y results in $b_y = (\varpi_y(P), \varpi_y(L), \varpi_y(C_1), \dots, \varpi_y(F), \dots)$.

2.2.5 Spatial transformations as geometrical objects

The identity spatial transformation e can be described by a geometrical object by $E = \{(y, y) \mid y \in \mathbb{A}\}$, where the attribute assigned to each pair is the original location of the point. Following the Lagrangian formulation, the action of the spatial transformation h on E results in $h \star E = H = \{(h \star y, y) \mid y \in \mathbb{A}\}$, and in terms of the Eulerian formulation, $H = \{(y, h^{-1} \star y) \mid y \in \mathbb{A}\}$.

Having a sequence of spatial transformations h_1, h_2, \dots, h_k , they can successively act on E obtaining

$$\underbrace{h_k \star (\dots \star (h_2 \star (h_1 \star E)) \dots)}_{(h_k \bullet \dots \bullet h_2 \bullet h_1) \star E = q \star E} = \{ (y, \underbrace{(h_1^{-1} \bullet h_2^{-1} \bullet \dots \bullet h_k^{-1}) \star y}_{q^{-1}}) \mid y \in \mathbb{A} \} = Q, \quad (2.2.2)$$

where Q is a single geometrical object encoding the successive composition of the given spatial transformations.

It is a common problem that an image degrades every time that it is transformed and stored in a structured grid. However, the spatial transformation h can be stored separately of the original

image data in a geometrical object H and store successive transformations as in Eq. (2.2.2). At the end of the process, a single interpolation of the image at the deformed coordinates is performed and in this way avoids to degrade the image information.

2.3 Parametrization of the action of spatial transformations

The set of all spatial transformations is the set of all homeomorphisms $\mathcal{H}_{\mathbb{A}}$ of the ambient space. With the representation of spatial transformations as a geometrical object given in Section 2.2.5 it is possible to explicitly describe any spatial transformation. Thus, any $h \in \mathcal{H}_{\mathbb{A}}$ can be univocally identified with a geometrical object H and this identification will be denoted by h^H , such that $H(y) = \varpi_y H = (h^H)^{-1} * y$ for any $y \in \mathbb{A}$. Another explicit identification is given by a displacement field, *i.e.* a function $\mathbf{u}: \mathbb{A} \rightarrow \mathbb{R}^d$, such that $h^{\mathbf{u}} * y = y + \mathbf{u}(y)$ (which is possible thanks to the equivalence $\mathbb{A} \equiv \mathbb{R}^d$). However, with these identifications an infinite number of parameters (or a very large number depending of the desired spatial resolution) are required to completely specify h . Moreover, besides the constructive process of the mapping, it is hard to check if a mapping is a valid homeomorphism.

For practical purposes and taking into account the constraints imposed by the applications, it is usual to deal with smaller subsets $\mathcal{S} \subset \mathcal{H}_{\mathbb{A}}$ of spatial transformations and more compact parameterizations can be used to identify the elements of \mathcal{S} . For example, if \mathcal{S} has a countable number of elements, then its elements can be indexed by a natural number, $\mathcal{S} = \{h^i \mid i \in \mathbb{Z}\}$. However, these are not the most interesting cases for practical purposes.

A set \mathcal{S} will be called a *continuously parameterizable spatial transformation set* if the following conditions hold:

- There exist an open, not necessarily connected, subset $\Upsilon \subseteq \mathbb{R}^m$ (for some m) and a surjective function $\Theta: \Upsilon \rightarrow \mathcal{S}$.
- For a continuous path of parameters, $\bar{p}_{(t)}: (0, 1) \rightarrow \mathbb{R}^m$, the path described in \mathbb{A} by the actions $\Theta(\bar{p}_{(t)}) * y$ is continuous for every point $y \in \mathbb{A}$.
- Conversely, if a subset $\mathcal{L} \subseteq \mathcal{S}$ fulfils that for every $y \in \mathbb{A}$ the set $\mathcal{L}_y = \{\ell * y \mid \ell \in \mathcal{L}\}$ is a continuous path in \mathbb{A} (that is, there exist an open interval $I \subseteq \mathbb{R}$ and a continuous function $\gamma_y: I \rightarrow \mathbb{A}$ with $\text{imag}(\gamma_y) = \mathcal{L}_y$), then, there exist a piecewise continuous function $\eta: I \rightarrow \Upsilon$ such that $\text{imag}(\Theta \circ \eta) = \mathcal{L}$.

The smallest integer number m that holds the previous conditions will be called the *degrees of freedom of the spatial transformation set*. It is remarkable that the whole set $\mathcal{H}_{\mathbb{A}}$ cannot be associated with a continuous parameterizable transformation set with a finite number of degrees of freedom. The function $\Theta(\cdot)$ will be called *spatial transformation model* and each spatial transformation $h \in \mathcal{S} = \text{imag}(\Theta)$ can be identified by an m -dimensional parameter $\bar{p} \in \Upsilon$ such that $h^{\bar{p}} = \Theta(\bar{p})$ although different values of \bar{p} could results in the same spatial transformation. In summary, a spatial transformation model gives a rule to transform a point y in terms of some finite dimensional parameter \bar{p} , such that $h^{\bar{p}} * y = \Theta(\bar{p}) * y$.

Additionally, a model Θ will be called a *differentiable spatial transformation model* if for any differentiable curve $\gamma_{(t)} \in \Upsilon$, *i.e.* a curve in the parameter space of the model, $\Theta(\gamma_{(t)}) * y$ is a differentiable curve in the ambient space for all points $y \in \mathbb{A}$. Note that this condition does not mean that a differentiable curve in \mathbb{A} should remains differentiable in \mathbb{A} after apply the transformation.

The importance of differentiable spatial transformation models lies in the area of optimization of functions defined on the set \mathcal{S} . Firstly, the function Θ^{-1} allows to characterize the continuity

and smoothness of the objective function by mapping functions from \mathcal{S} to Υ . Secondly, standard optimization techniques on vector spaces can be used on Υ because it is, by definition, an open subset of \mathbb{R}^m . It is usual that the objective function be determined by the state of a geometrical object and in this framework, the continuity and differentiability of the transformation model makes compatible a perturbation in the state of a geometrical object with a perturbation in Υ .

Along literature there are different taxonomies to categorize sets of spatial transformations [Holden 2008] [Sotiras et al. 2013]. Spatial transformation models can be organized in the following categories:

- **Parametric transformations:** this category includes those spatial transformations that are defined in terms of a few parameters and the action on a point of \mathbb{A} is usually expressed as an algebraic expression that includes the parameters and the coordinates of the point. Examples of these transformations are rigid body transformations, similarity transformations, affine, projective and Möbius transformations, among others. Transformations expressed by polynomials (usually of degree no larger than 5) in the coordinates of the points also can be considered as parametric, where the parameters are, for example, the coefficients of the polynomials. These polynomials may require up to hundreds of parameters to identify the transformation (see for example [Woods et al. 1998] where fifth degree polynomials were used in 3 dimensions requiring up to 168 parameters).
- **Object-dependent transformations:** transformations from this category are defined in terms of some parameters but also in terms of the object or image to be transformed. For example, in the case of the registration of a few points or landmarks, it is usual to describe the transformation as the displacements to be applied to each point. However, the spatial transformation must be able to be applied to the whole ambient space and therefore it is necessary to extend or interpolate these displacements to the entire domain. Thus, the resulting spatial transformation will also depend on the number and location of the original points.
- **Non-parametric transformations:** transformations in this category can be only described by specifying the mapping or action on every point of \mathbb{A} . However, it is impossible to store a computational description of a non-parametric function for a continuous space like \mathbb{R}^2 or \mathbb{R}^3 . Typically, as the representation of vector valued images, the transformation is described in a structured grid with high spatial resolution, where the mapping is specified explicitly for each node together with an interpolation scheme and an extrapolation rule. These descriptions typically require a very large number of parameters and descriptions with $\sim 10^6$ or $\sim 10^7$ degrees of freedom are not uncommon. A transformation is considered non-parametric when the number of nodes is so large that the function can be considered dense in the space environment. For example, for the transformation of an image, the displacement or mapping is specified for every pixel or voxel of the image. In the registration process it is considered that a spatial transformation is non-parametric if under an increase of the spatial resolution of their description (with the consequent increasing of the degrees of freedom) differences in the registration process are negligible. In this sense, a non-parametric spatial transformation model does not bias the results of a registration in more than the spatial resolution of the transformation model.

Spatial transformations from parametric and non-parametric categories are considered *global mappings* in the sense that they transform the whole ambient space independently of the geometrical objects defined on it. On the other hand, for the object-dependent case, both the number of parameters as the transformation itself will depend on the objects to be transformed.

2.4 Spatial transformation groups

As was already stated in Section 2.1, the set of all homeomorphisms of the ambient space $\mathcal{H}_{\mathbb{A}}$ has a group structure under the function composition $(\cdot \circ \cdot)$ operation and under its restriction $(\cdot \bullet \cdot) = (\cdot \circ \cdot)|_{\mathcal{H}_{\mathbb{A}}}$. The *group identity*, e , is associated with the function $\text{id}_{\mathbb{A}}$ and the *group inverse* operation, $\text{inv}(\cdot)$, by the inverse function $(\cdot)^{-1}$. Immediate consequences of the group structure are:

- $h \bullet e = e \bullet h = h$
- $h \bullet \text{inv}(h) = \text{inv}(h) \bullet h = e$
- $\text{inv}(\text{inv}(h)) = h$
- $\text{inv}(h \bullet q) = \text{inv}(q) \bullet \text{inv}(h)$
- $h \bullet (q \bullet s) = (h \bullet q) \bullet s$

for any spatial transformations h , q and $s \in \mathcal{H}_{\mathbb{A}}$. It is standard notation to omit the \bullet symbol and the juxtaposition of group elements is used as a shorthand for the composition operation $hq \equiv h \bullet q$.

This group structure is compatible with the action $(\cdot \star \cdot)$ of a spatial transformation on any geometrical object of any type, *i.e.* for any geometrical object B and any spatial transformations $h, q \in \mathcal{H}_{\mathbb{A}}$

$$\begin{aligned} e \star B &= B \\ h \star (q \star B) &= (hq) \star B \\ h \star B = C &\Leftrightarrow B = \text{inv}(h) \star C. \end{aligned}$$

It is said that spatial transformations act *at the left* on the geometrical objects by the action³ $(\cdot \star \cdot)$.

A nonempty set of spatial transformation \mathcal{G} is called a (*transformation*)⁴ *group* if along with every two transformations h and q belonging to the set \mathcal{G} , their composition hq and the inverse transformation h^{-1} belong to the same set \mathcal{G} . An obvious consequence of these requirements is that every transformation group contain the identity transformation e . In addition, group actions usually preserves interesting geometrical magnitudes or relations of the objects.

2.4.1 Matrix groups and matrix representations

Hitherto the relations between spatial transformations have been mainly described abstractly, *i.e.* only with the minimal algebraic structure of the homeomorphism group operations. This description gives sufficient generality to formalize geometric concepts involved in the registration of images. On the other hand, Section 2.3 introduces spatial transformation models that parameterize a set of transformations allowing, from a computational point of view, the calculation of actions on points and objects. However, these parametrization or identifications do not include

³ Alternatively, if there exist a map $\diamond: \mathcal{H}_{\mathbb{A}} \times \mathbb{A} \rightarrow \mathbb{A}$ such that $h \diamond (q \diamond B) = (qh) \diamond B$ for any object B , then it is said that spatial transformations act *at the right* on the objects by the action $(\cdot \diamond \cdot)$. A right-action \diamond can be obtained from a left-action \star by $(\cdot \diamond \cdot) = (\text{inv}(\cdot) \star \cdot)$. Without loss of generality, left actions will be always consider throughout this thesis.

⁴ The group structure of a set can be defined in an axiomatic way by defining the operations composition and inverse abstractly. Or can be defined by the use of an action on a set. In the later case the set is called a transformation group. See for example [Arnold 1992] for a practical justification to define the group structure in terms of the action on a space of objects.

in their definition the relationships between different spatial transformations. A classical example is presented when model the rotations in a three-dimensional space with Euler angles. Given two rotations parameterized with 3 Euler angles $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, where $\bar{\theta} = (\theta^x, \theta^y, \theta^z)^T$ corresponds to rotations around x , y and z coordinate axis (or parameterized in any other preferable convention of Euler angles [Vince 2011b, Appendix A]), and the same for $\bar{\phi}$, and a point $\mathbf{y} = (y^1, y^2, y^3)^T$, it is rather simple to algebraically obtain in terms of trigonometric expressions the coordinates of the transformed point $\boldsymbol{\theta} \star (\boldsymbol{\phi} \star \mathbf{y})$. However, it is more daunting to know which are the Euler angles $\bar{\omega}$ which parameterize $\boldsymbol{\theta} \star \boldsymbol{\phi} = \boldsymbol{\theta} \bar{\omega} \boldsymbol{\phi}$ and also those Euler angles corresponding to $(\boldsymbol{\theta})^{-1}$ (see for example [Schneider and Eberly 2002]). In another example, let us consider spatial transformations which their action is expressed as a second order polynomial on the coordinates of the points. Each of these transformations can be naturally identified with the coefficients of the polynomial providing simple algebraic expressions to numerically compute the action on any coordinate point. However, given two transformations from this model $\boldsymbol{\theta}^p$ and $\boldsymbol{\theta}^q$, the composition of both does not necessarily belong to the same model. Moreover, there is no guaranty to be able to describe the inverse transformation within the same model. To summarize, although the spatial transformation model allows to parameterize any transformations of a set usually giving convenient mathematical expressions or algorithms to perform the action on a coordinate point, the basic relations between spatial transformations are, in general, hardly expressed in terms of the parameters of the model.

In the following the concept of matrix groups will be introduced that will be a fundamental and practical tool for the analysis of the structure of most of the groups of spatial transformations used throughout this thesis.

The set of all square matrices of size $n \times n$ with real entries will be denoted by \mathbb{M}_n . Operations between matrices can be conducted *element by element*, wherewith matrices immediately inherit algebraic structures from the real numbers. Also, there is a widely known relation between matrices and linear combinations over vector spaces. This relation induces a *matrix multiplication* rule which results isomorphic to the composition of linear combinations [Hoffman and Kunze 1977]. A *matrix group* \mathbf{G} of degree n is a set of matrices from \mathbb{M}_n which has also a group structure under the matrix multiplication operation [Hall 2003] [Tapp 2005]. The identity element of the group is the identity matrix I_n , the inverse operation of the group is given by the matrix inversion operation and therefore any matrix group \mathbf{G} have to be composed by non-singular matrices.

The interest in the study of matrix groups is largely due to the fact that many of the spatial transformations groups used in imaging applications can be *represent with matrices*. A *representation* is an identification of every element of some abstract mathematical structure with a more concrete numerical and algebraic entity providing that the existing properties and relationships among the abstract elements are in accordance with the relations among their representations.

The concept of representation have been previously used to trivially identify each point of the ambient space, an element of an abstract set, with a d -dimensional array of \mathbb{R}^d , a numerical structure with a well defined and know algebraic structure⁵. Formally: given an algebraic structure $\mathcal{A} = (\mathcal{A}, \mathbf{op}_1, \mathbf{op}_2, \mathbf{op}_3, \dots)$, where \mathcal{A} is a set $\{a_\alpha\}$ and \mathbf{op}_i are operations (unary, binary or in general finitary operations) with arguments in \mathcal{A} and given another algebraic structure $\mathcal{B} = (\mathcal{B}, \hat{\mathbf{op}}_1, \hat{\mathbf{op}}_2, \hat{\mathbf{op}}_3, \dots)$, then a representation $\mathbf{rep}_{\mathcal{A} \rightarrow \mathcal{B}}$ is a bijective identification⁶ of every element from \mathcal{A} with an element from \mathcal{B} such that every algebraic operation \mathbf{op}_m can be written

⁵ Some books, such as [Gilmore 2005], prefer the term *realization* for this mapping and reserve the term *representation* for the special case of matrix representations.

⁶ Many authors, in special from the group theory area, define a representation as a surjective function and refer as *faithful representation* when it is a also one-to-one map.

down explicitly and on which calculations on \mathbf{A} can be carried out by

$$\mathbf{op}_m(a_i, a_j, \dots) = \mathbf{rep}_{\mathbf{A} \rightarrow \mathbf{B}}^{-1}\left(\tilde{\mathbf{op}}_m(\mathbf{rep}_{\mathbf{A} \rightarrow \mathbf{B}}(a_i), \mathbf{rep}_{\mathbf{A} \rightarrow \mathbf{B}}(a_j), \dots)\right),$$

where $\tilde{\mathbf{op}}_m$ is a finite composition of operations $\hat{\mathbf{op}}_k$ from \mathbf{B} .

Unexpectedly, almost all spatial transformation groups, with a finite number of degrees of freedom, encountered in applications from the image analysis scope are matrix groups. This effects an enormous simplification in the study transformation groups since almost all of what we would like to learn about groups can be determined by studying matrix groups.

Specifically, given a group of spatial transformations \mathcal{G} with elements h_α , identity e and the operations of composition $(\cdot \bullet \cdot)$ and inverse $\mathbf{inv}(\cdot)$, it is called *representable* by matrices from \mathbb{M}_n , for some n , or simply *linear group*, when the following identifications can be made:

- Every element h_α can be identified with a matrix $H_\alpha \in \mathbb{M}_n$.
- In particular, e is identified with the identity matrix I_n .
- The composition $h_\alpha \bullet h_\beta$ is identified with the matrix multiplication $H_\alpha H_\beta$.
- Group inverse operation $\mathbf{inv}(h_\alpha)$ is identified with the matrix inversion H_α^{-1} .
- In the case that the set \mathcal{G} is endowed with a topological structure, then its global topology is the same than the global topology of the set $\{H_\alpha\}$ as a subset of $\mathbb{R}^{n^2} \equiv \mathbb{M}_n$. Else, \mathcal{G} can be endowed with the topology of $\{H_\alpha\}$ via the previous identifications.

The study of the algebraic structure of a linear group is provided with the well known language of matrix algebra. In this language the properties of the group become elegant, compact and easy to handle algebraic expressions on matrices. Moreover, matrices are a basic tool in practical mathematics and in the solutions and implementations of real world applications being basic tools in many numerical software and libraries.

2.4.2 The action of matrix groups on spatial points

Matrices from \mathbb{M}_n also represent *linear combinations* of elements in an n -dimensional vector space \mathbb{V} . This suggests that, in addition to operations between spatial transformations, matrices can be useful to express the action of the transformation on a point of the ambient space. Typically, the points of the ambient space $\mathbb{A} \equiv \mathbb{R}^d$ can be mapped via a function $\psi: \mathbb{A} \rightarrow \mathbb{R}^n$ for some $n \geq d$, such that, for a spatial transformation h which is represented by $H \in \mathbb{M}_n$ and for any point $y \in \mathbb{A}$, results $\psi(h \star y) = H\psi(y)$. A common example is given by what is called *homogeneous coordinates* which allows to express affine transformations (transformations composed by a linear part followed by a translational part) as a matrix-vector formula. In a two dimensional application it is performed by:

- Let $y = (y^1, y^2)^T$ be a point in \mathbb{A} , the function ψ consists in raise the point to \mathbb{R}^3 by adding a 1 in the last coordinate: $\psi(y) = (y^1, y^2, 1)^T$.
- The affine transformation h is represented by a matrix H of the form

$$H = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix}$$

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a linear transform and t_x, t_y correspond to horizontal and vertical translations respectively.

- Perform the matrix-vector product

$$H\psi(\mathbf{y}) = \begin{pmatrix} ax^1 + by^2 + t_x \\ cx^1 + dy^2 + t_y \\ 1 \end{pmatrix}$$

- Finally, apply the function ψ^{-1} which remove the last coordinate and recover the two dimensional point structure.

Further examples will be shown in Section 5.1 for more complex groups such as projective and Möbius transformation groups.

This procedure could induce to think that the entries of the matrices can be used as a m -dimensional spatial transformation model $\Theta: \Upsilon \subseteq \mathbb{R}^m \rightarrow \mathcal{G}$. However a spatial transformation model demands that Υ be an open set of \mathbb{R}^m but, in general, the matrix entries of a matrix group of degree n shape a surface embedded in \mathbb{R}^{n^2} which is usually not open.

2.5 Spatial transformation models

In the following we will enumerate several spatial transformations models typically used in computer vision and image analysis areas. Their uses for medical image analysis are fundamental tools in image registration and computational anatomy. Although the list is far to be complete it is wide enough to gather the most common models used in applications.

2.5.1 Discrete spatial transformations models

By discrete spatial transformations models we refer to those sets where all its elements can be indexed by a natural number and therefore there is no way to connected them by continuous paths. Classical examples of these are **Inversion groups** and **Dihedral groups**.

2.5.2 Finite dimensional models

Here we will include spatial transformation model with a few dozens of degrees of freedom. Most of them are also spatial transformation groups which can be represented by matrix groups and in these cases we will specify how to express the corresponding matrix forms in terms of the parameters. For all the models, the action on a spatial point will be explicitly given in term of the coordinates of the point and of the parameters of the model.

Translations, $\mathcal{T}(d)$

This spatial transformation model consists in all the possible translations of an object defined in a d -dimensional ambient space. Several magnitudes from the geometrical objects such as distances between points, angles between lines, attitudes, are preserved under the action of this model.

Each spatial transformation can be parameterized by specifying the amount of translation along each coordinate, such that $\hbar^{\bar{t}}$, with $\bar{t} = (t^1, t^2, \dots, t^d)^T \in \mathbb{R}^d$. The action of a transformation on a point $\mathbf{y} = (y^1, y^2, \dots, y^d)^T \in \mathbb{A}$ is given by

$$\hbar^{\bar{t}} \star \mathbf{y} = (y^1 + t^1, y^2 + t^2, \dots, y^d + t^d)^T.$$

It is very easy to ascertain that this model is a group where the composition is given by $\hbar^{\bar{t}} \hbar^{\bar{d}} = \hbar^{\bar{t}+\bar{d}}$ and the inverse by $(\hbar^{\bar{t}})^{-1} = \hbar^{-\bar{t}}$.

Finally, it is a representable group by a matrix group, such that every element $\hbar^{\bar{t}} \in \mathcal{T}(d)$ can be represented by a $(d+1) \times (d+1)$ matrix of the form

$$H^{\bar{t}} = \left(\begin{array}{c|c} I_d & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right).$$

Isotropic scales, $\mathcal{S}^+(1)$

The transformations of this set generate resized versions of the objects around the origin of the coordinate system defined on \mathbb{A} . Several magnitudes from the geometrical objects are preserved under the action of these spatial transformations, such as angles, attitudes, and ratios of lengths. Also, the point at the origin remains at the same location.

Spatial transformations from this model can be parameterized with a positive scalar number $s \in \mathbb{R}^+$. The action on a point is $\hbar^s \star (y^1, y^2, \dots, y^d)^T = (sy^1, sy^2, \dots, sy^d)^T$. It is also a representable group and a matrix representation of \hbar^s can be given by $H^s = s$.

Non-isotropic scales, $\mathcal{S}^+(d)$

Using these transformations the objects can be resized differently along each spatial direction. For example, objects can be stretched along the horizontal direction but shrunk along the vertical one. These spatial transformations preserve angles, attitudes and the point at the origin.

Scale factors along each spatial direction, (s^1, s^2, \dots, s^d) , give a parametrization of the transformations and in order to avoid orientation inversions (changes in the handedness of a frame) only positive numbers are considered. The action on a point of the ambient space is given by $\hbar^{(s^1, s^2, \dots, s^d)} \star (y^1, y^2, \dots, y^d)^T = (s^1 y^1, s^2 y^2, \dots, s^d y^d)^T$.

The transformations set results in a group with matrix representation given by $d \times d$ matrices of the form

$$H^{(s^1, s^2, \dots, s^d)} = \left(\begin{array}{cccc} s^1 & 0 & \dots & 0 \\ 0 & s^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & s^d \end{array} \right).$$

Isotropic scale + Translations, $\mathcal{ST}(d)$

This spatial transformation model is the composition of uniform scales and translations. Under these spatial transformations the objects are isotropically resized around the origin and next translated. Although this type of transformation does not have many practical applications in medical imaging and computer vision applications, they will be used several times in Chapters 4 and 5 to exemplify many properties of about Lie groups.

A parametrization of the transformations is $\hbar^{(s, \bar{t})}$, where $s \in \mathbb{R}^+$ and $\bar{t} \in \mathbb{R}^d$. The action on a point is $\hbar^{(s, \bar{t})} \star y = sy + \bar{t}$.

The set has a group structure and a matrix representation is given by $(d+1) \times (d+1)$ matrices of the form

$$H^{(s, \bar{t})} = \left(\begin{array}{c|c} sI_d & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right).$$

Rotations, $\mathcal{SO}(d)$

Rotations are the transformations which preserve distances between points, angles and a keep fixed the point at the origin. With these requirements it can be seen that rotations have to act

linearly on the coordinates of a point and therefore can be represented by a matrix from \mathbb{M}_d . Furthermore, in order to preserve distances and angles these matrices have to satisfy $RR^T = I_d$. This constraint specifies $d(d+1)/2$ independent quadratic equations leaving $d(d-1)/2$ degrees of freedom. Moreover, in order to preserve the handedness rotations are also requested to fulfil $\det(R) = 1$.

Rotations in 2-dimensions (when $\mathbb{A} \equiv \mathbb{R}^2$) can be parameterized by a single number θ while three parameters are required for the 3-dimensional case.

There are several parametrizations proposed to describe rotations, specially in \mathbb{R}^3 (quaternions [Shoemake 1985] [Hanson 2006] [Vince 2011a], Euler angles conventions [Craig 2004] [Vince 2011b], axis-angle representation [Murray et al. 1994] [Grassia 1998]), but also for higher dimensions [Gallier and Xu 2002].

Once the parametrization is chosen, the action of this spatial transformation model is given by $h^{\bar{\theta}} \star y = R_\theta y$, and R_θ is a matrix representation of the rotation. The set of rotations has a group structure.

Rigid body transformations, $\mathcal{SE}(d)$

These spatial transformations are composed of rotations and translations. It is the most general spatial transformation model which preserves the distances between all pair of points and the handedness of the objects.

A parametrization of the transformations is $h^{(\bar{\theta}, \bar{t})}$, where $\bar{\theta}$ parameterize a rotation and $\bar{t} \in \mathbb{R}^d$ is a translation. The action is given by $h^{(\bar{\theta}, \bar{t})} \star y = R_\theta y + \bar{t}$. Rigid body transformations have a group structure which is usually called *special Euclidean group*. The transformations can be represented by matrices in \mathbb{M}_{d+1} with the form

$$H^{(\bar{\theta}, \bar{t})} = \left(\begin{array}{c|c} R_\theta & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right).$$

Similarity transformations, $\mathcal{SIM}(d)$

It is a combination of rotations, translations and isotropic scales⁷.

The elements are parameterized by $\langle^{(s, \bar{\theta}, \bar{t})}$, with $s > 0$, and its action is $h^{(s, \bar{\theta}, \bar{t})} \star y = sR_\theta y + \bar{t}$.

It is Lie group and a matrix representation is given by

$$H^{(s, \bar{\theta}, \bar{t})} = \left(\begin{array}{c|c} sR_\theta & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right).$$

General linear transformation, $\mathcal{GL}(d)$

These spatial transformations transform linearly the ambient space. The coordinate system origin remains at the origin and the straight lines remain straight.

This set has a group structure and the transformations can be represented by $d \times d$ non-singular matrices. A natural parametrization of the spatial transformations is carried on by the matrix elements itself, such that the matrix $M \in \mathbb{M}_d$ with non-null determinant represents to the spatial transformation $h^M \in \mathcal{GL}(d)$. Obviously, the degrees of freedom are d^2 .

⁷ In other works this transformations are called conformal [Dubrovin et al. 1984] [Woods 2003] or similitudes [Guggenheim 1967]. Sometimes it is defined allowing also inversions, but here it will be defined only for positive scales.

Orientation preserving general linear group, $\mathcal{GL}^+(d)$

This set is given by selecting from the previous set of spatial transformations those ones which preserve the handedness. The model has a group structure and its elements are represented by $d \times d$ matrices having strictly positive determinant.

Special linear, $\mathcal{SL}(d)$

This model is given by particularizing to the spatial transformation from $\mathcal{GL}^+(d)$ which remains fixed the area or volume of the objects. The model has a group structure and its elements are represented by $d \times d$ matrices determinant equals to 1. Regarding the parametrization, due to the constraint on the determinant, one degree of freedom was removed from the model. Therefore, a parametrization can be performed by specifying $d^2 - 1$ elements of the matrices.

Orientation preserving general affine, $\mathcal{GA}^+(d)$

It is composed of transformations in $\mathcal{GL}^+(d)$ group and translations. It is the most general transformation model preserving collineation, parallelism and handedness.

The action is $h^{(M, \bar{t})} \star y = h^M \star y + \bar{t}$, where M is a parameterization of the orientation preserving general linear group, *i.e.* a $d \times d$ with positive determinant.

This model is a group with matrix representation given by

$$H^{(M, \bar{t})} = \begin{pmatrix} M & \bar{t} \\ 0_d^T & 1 \end{pmatrix} .$$

Orientation preserving special affine, $\mathcal{SA}^+(d)$

Is the subset of transformations from \mathcal{GA}^+ that also preserve the area or volume of the objects. It is also a group and its matrix representation is

$$H^{(M, \bar{t})} = \begin{pmatrix} M & \bar{t} \\ 0_d^T & 1 \end{pmatrix} ,$$

where $\det M = 1$.

In the previous spatial transformations model, some of them include translations. In these models the spatial transformations have a *semidirect product* structure, *i.e.* each linear group \mathcal{G} is extended to the group $\mathcal{G} \ltimes \mathcal{T}$ with \mathcal{T} being the group of translations in d -dimensions. Each spatial transformation in $\mathcal{G} \ltimes \mathcal{T}$ can be identified by a pair (h, t) with $h \in \mathcal{G}$ and $t \in \mathcal{T}$. The action on an ambient space point y is performed by, first apply \mathcal{G} (which is typically an action around the origin of the coordinate system) and later apply the translation: $(h, t) \star y = t \star (h \star y)$. By the identification of $\mathbb{R}^d \equiv \mathbb{A}$, the operation composition results in $(h_1, t_1) \bullet (h_2, t_2) = (h_1 h_2, h_1 \star t_1 + t_2)$.

However, in some cases, it is possible to construct a *direct product* structure such that $(h_1, t_1) \bullet (h_2, t_2) = (h_1 h_2, t_1 + t_2)$. But, in order to make this composition law being compatible with the successive actions on an object, we have to modify the action operation.

Centered transformation ($\mathcal{G}(d) \times \mathcal{T}(d)$)

In the ambient space the origin of the coordinate system has been chosen arbitrarily. The linear transformations act on the objects remaining fixed to this arbitrary point, for example the objects are rotated or scaled around the origin, and then translate the object. An alternative formulation

is obtained by rotating the space around a specified center and after that translate the object together with the center point.

To do that we define an *object with a center* (B, c) and the actions defined on them. Actually each point of the ambient space $y \in A$ is replaced by the pair (y, c) , where c is point which is arbitrarily defined as the center of the object.

Now, given a spatial transformation of the form (h, t) , its action on a pair point-center is given by

$$(h, t) \star (y, c) = (t \star (h \star (y - c) + c), t \star c)$$

When the spatial transformation model $\mathcal{G}(d)$ has a group structure, then the model $(\mathcal{G}(d) \times \mathcal{T}(d))$ will also have a group structure. Moreover, if the former is representable by matrices from M_n then the later can be represented by matrices of the form

$$H^{(\bar{M}, \bar{t})} = \left(\begin{array}{c|c|c} M & 0_{n,d} & 0_n \\ \hline 0_{d \times d} & I_d & \bar{t} \\ \hline 0_d^T & 0_d^T & 1 \end{array} \right).$$

Weak affine transformation model

It is a transformation model composed of a non-isotropic scale following by a rotation and by a translation [Bankman 2008, Chapter 32]. Its action on a point $y \in A$ is given by $h^{(\bar{s}, \bar{\theta}, \bar{t})} \star y = (t^{\bar{t}} r^{\bar{\theta}} s^{\bar{s}}) \star y$, where $s^{\bar{s}}$ is the non-isotropic scale with factors $\bar{s} = (s^1, s^2, \dots, s^d)^T$, $r^{\bar{\theta}}$ is the rotation parameterized with $\bar{\theta}$ and $t^{\bar{t}}$ is the translation given by $\bar{t} = (t^1, t^2, \dots, t^d)^T$.

Its action can be described by a matrix times the point represented by its homogeneous coordinates. The matrix takes the following form

$$H^{(\bar{s}, \bar{\theta}, \bar{t})} = \left(\begin{array}{c|c} R_{\theta} \text{diag}(\bar{s}) & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right).$$

The Weak affine transformation model is a continuously parameterized model with $2d + d(d-1)/2 = d(d+3)/2$ degrees of freedom. However, it has no group structure since the composition and the inverse transformations do not belong to the model [Flusser and Zitová 2013].

Polynomial transformations model

The components of \mathbb{R}^d form a field, therefore, a polynomial function can be defined on \mathbb{R}^d .

The action of a general polynomial transformation h of degree D on a point $y = (y^1, y^2, \dots, y^d)^T$ of a d -dimensional ambient space is given by

$$(h \star y)^j = \sum_{n_1=0}^D \sum_{n_2=0}^D \dots \sum_{n_d=0}^D c_{n_1 n_2 \dots n_d}^j (y^1)^{n_1} (y^2)^{n_2} \dots (y^d)^{n_d}, \quad (2.5.1)$$

where $(h \star y)^j$ refers to the j -th coordinate of the transformed point.

Each polynomial transformation can be identified by the $d(D+1)^d$ coefficients $h^{(c_{n_1 n_2 \dots n_d}^j)}$. There must take into account that not all $d(D+1)^d$ -dimensional parameters $c_{n_1 n_2 \dots n_d}^j$ defines a spatial transformation, because for some values the algebraic expression Eq. (2.5.1) produce folds in the ambient space and therefore are no homeomorphisms.

The polynomial transformation model does not form a group (except for the trivial case of $d = D = 1$). The composition operation increase the order of the polynomial and the inverse transformation cannot be described by a polynomial.

2.5.3 Object-dependent transformation models

In some cases a parameterization is selected for the specification of spatial transformations but this parameterization takes sense only for a given object.

- It is common, for example, to parameterize a deformation of an object by specifying the deformation in some point of the object and interpolating and extending the deformation to the whole ambient space. This case will be consider in Section 3.5 where a minimum regularity energy interpolation is considered.
- A very common parameterization of deformation fields is by b-splines functions defined on a structured grid. Although the parameterization depends on the selected grid, the choice grid itself depends on the object that will be deformed. The extension and the resolution of the grid have to be properly selected taken into account the extension and the frequencies presented in the object.

The deformation is specified in the nodes of the grid and a interpolation scheme is often used to extend the deformation to the whole space. Often the parametrization on the nodes is by specifying the displacements of nodes and interpolate the displacement in-between the nodes, for example with the multi-linear interpolation described in Section 2.2.2. In other approaches the parametrization of the nodes is by specifying the coefficients of a basis spline associated to each node. This technique is commonly known as *free form deformation* [Hsu et al. 1992] [Rueckert et al. 1999] [Rueckert et al. 2006]. Note that in this type of transformations, the number of parameters can be quite large since a three-dimensional grid with $10 \times 10 \times 10$ nodes requires the specification of 3000 parameters to characterize the transformation.

2.5.4 Infinite dimensional spatial transformation models

The most general description of a spatial transformation is given by describing the transformed location for every point in the ambient space. A similar approach was given previously in Section 2.2.5 where spatial transformations were described as geometrical objects.

There are two paradigms commonly used for the infinite dimensional models. The first one is called the *small deformation paradigm* and it is helpful to the description of transformation with relative high smoothness. In this paradigm, it is relaxed the conditions that spatial transformation are homeomorphisms, and transformation are only considered as continuous mappings from the ambient space to itself. The second one is called *large deformation paradigm* and in those parametrization, only homeomorphisms of the ambient space are inherently allowed.

Displacement maps

This model parameterizes the spatial transformation explicitly by the vector field $\mathbf{u}: \mathbb{A} \rightarrow \mathbb{R}^d: y \mapsto \mathbf{u}(y) = \left(u_{(y)}^1, u_{(y)}^2, \dots, u_{(y)}^d \right)^T$ which give the displacement to apply to each point of the ambient space. Then, the action of spatial transformations parameterized with this model is $h^\mathbf{u} * y = y + \mathbf{u}(y)$. In order to be able to manage and store displacement maps in a computer system, a finite dimensional description must be used. For example, it is common to specify the displacement maps as d scalar functions using the approach given in Section 2.2.2.

An issue about this parameterization is that it is hard to impose constraints to the displacement fields to guarantee that the resulting spatial transformation is a homeomorphism. Under some mild conditions, such as continuity of the displacement field, it can be proved that there is a $\rho > 0$ such that the mapping $id_{\mathbb{A}} + \epsilon \mathbf{u}: y \mapsto y + \epsilon \mathbf{u}(y)$ is a homeomorphism if

$|\epsilon| < \rho$ [Younes 2010, Chapter 8]. This description of the spatial transformations is included in the *small deformation paradigm*.

Flow of diffeomorphisms

To build spatial transformations which allow large displacements it can be used the small deformations together with the property that the composition homeomorphisms results in a homeomorphism. Therefore, given two homeomorphisms $h^{\mathbf{u}_1}$ and $h^{\mathbf{u}_2}$ parameterized by their displacement fields, *i.e.* belonging to the small deformation paradigm, the composition

$$(h^{\mathbf{u}_2} \bullet h^{\mathbf{u}_1}) * \mathbf{y} = \mathbf{y} + \mathbf{u}_2(\mathbf{y} + \mathbf{u}_1(\mathbf{y}))$$

will be a homeomorphism. This process can be performed indefinitely given a spatial transformation of the form

$$h^{(\dots, \mathbf{u}_{m+1}, \mathbf{u}_m, \mathbf{u}_{m-1}, \dots, \mathbf{u}_2, \mathbf{u}_1)} = h^{\mathbf{u}_{m+1}} \bullet h^{\mathbf{u}_m} h^{\mathbf{u}_{m-1}} \bullet \dots \bullet h^{\mathbf{u}_2} h^{\mathbf{u}_1}$$

that will be a homeomorphism if every displacement fields \mathbf{u}_m are small enough.

This compositional scheme to obtain spatial transformations suggests that a large transformation can be described by a spatially continuous velocity field, $\mathbf{w}: \mathbb{A} \times [0, 1] \rightarrow \mathbb{R}^d$: $(\mathbf{y}, \tau) \mapsto \mathbf{w}(\mathbf{y}, \tau) = \begin{pmatrix} w_{(y,\tau)}^1, \dots, w_{(y,\tau)}^d \end{pmatrix}^T$ such that the action $h^{\mathbf{w}} * \mathbf{y}$ is given by the solution $\mathbf{x}_{(t=1)}$ of the following initial value problem (IVP):

$$\begin{cases} \dot{\mathbf{x}}_{(t)} = \mathbf{w}(\mathbf{x}, t) \\ \mathbf{x}_{(0)} = \mathbf{y}. \end{cases} \quad (2.5.2)$$

Conditions about the existence of the solution of the IVP can be found in [Hirsch et al. 2004].

Since Eq. (2.5.2) is valid for all initial conditions $\mathbf{y} \in \mathbb{A}$, the velocity field induces a *flow of deformation mappings* given by

$$\begin{cases} \dot{\varphi}_{(t)} = \mathbf{w}(\varphi_{(t)}, t) \\ \varphi_{(0)} = \text{id}_{\mathbb{A}}. \end{cases} \quad (2.5.3)$$

The mappings resulting from Eq. (2.5.3) are invertible and differentiable. Furthermore, if \mathbf{w} is \mathcal{C}^k -differentiable with $k > 1$, then $\varphi_{(t)}$ is \mathcal{C}^{k+1} -differentiable. A \mathcal{C}^k -differentiable invertible mapping is usually called a \mathcal{C}^k *diffeomorphism*.

The associated spatial transformation is $h^{\mathbf{w}} \equiv \varphi_{(t=1)}$. Regarding its inverse, it can be obtained by the solution of

$$\begin{cases} \dot{\psi}_{(t)} = -\mathbf{w}(\psi_{(t)}, 1-t) \\ \psi_{(0)} = \text{id}_{\mathbb{A}}, \end{cases}$$

resulting $(h^{\mathbf{w}})^{-1} = \psi_{(t=1)}$.

Finally, it can be remarked that for any diffeomorphism φ there exist a velocity field \mathbf{w} that describes it [Trouvé 1995] [Dupuis et al. 1998] [Beg et al. 2005].

Stationary Velocity Fields (SVF)

With the previous description it is possible to describe any smooth spatial transformation, however the memory requirements are too large as a time-dependent velocity field must be stored. A simplification can be performed by removing the time dependence. Then, for a given *stationary*

velocity field (SVF) $\mathbf{v}: \mathbb{A} \rightarrow \mathbb{R}^d: y \mapsto \mathbf{v}(y) = \left(v_{(y)}^1, v_{(y)}^2, \dots, v_{(y)}^d \right)^T$ the spatial transformation $\hbar^\mathbf{v}$ has action given by $\hbar^\mathbf{v} \star y = \mathbf{x}_{(t=1)}$, where $\mathbf{x}_{(t)}$ is the solution of the IVP

$$\begin{cases} \dot{\mathbf{x}}_{(t)} = \mathbf{v}(\mathbf{x}) \\ \mathbf{x}_{(0)} = y. \end{cases} \quad (2.5.4)$$

In general, as a deformation mapping, $\hbar^\mathbf{v}$ is given by the solution at $t = 1$ of

$$\begin{cases} \dot{\varphi}_{(t)} = \mathbf{v} \circ \varphi_{(t)} \\ \varphi_{(0)} = \text{id}_{\mathbb{A}}. \end{cases} \quad (2.5.5)$$

Due to the resemblance of the evolution equations for linear ODEs, the deformation mapping $\varphi_{(t=1)}$ is usually denoted by $\exp(\mathbf{v})$. For any \mathbf{v} (with appropriate conditions such that $\varphi_{(t)}$ exists for all t), the mappings $\varphi_{(t)}$ belongs to a one-parameter subgroup, $\varphi_{(r+s)} = \varphi_{(r)} \circ \varphi_{(s)} = \varphi_{(s)} \circ \varphi_{(r)}$. Moreover, the inverse of the mapping is obtained by the integration of the negated velocity, i.e. $(\hbar^\mathbf{v})^{-1} = \hbar^{-\mathbf{v}}$.

Unlike the description with time-dependent velocity fields $\mathbf{w}(y, t)$, not all diffeomorphic mappings can be described by a SVF. Although, for every SVF spatial transformation, its inverse is also a SVF transformation, the composition of two SVF transformation, in general, does not a SVF transformation. Therefore, SVF transformations does not have a group structure.

Practical aspects of SVF transformations

The description of the SVF is usually given by d scalar functions in a structured grid, together with interpolation and extrapolation rules as was performed in Section 2.2.2. Several practical approaches to compute the solution of ODEs (2.5.4) and (2.5.5) have been proposed [W. 1997] [Kindermann and Georgiev 2002] [Orchard 2005] [Arsigny et al. 2006] [Bossa et al. 2008].

For example, for a large enough N , $\varphi_{(1/N)}$ can be approximated by $\text{id}_{\mathbb{A}} + \frac{1}{N}\mathbf{v}$. Using the one-parameter subgroup property $\varphi_{(k/N+1/N)} = \varphi_{(1/N)} \circ \varphi_{(k/N)}$ the solution solution of (2.5.5) is given by

$$\hbar^\mathbf{v} = \exp(\mathbf{v}) \approx \underbrace{\varphi_{(1/N)} \circ (\varphi_{(1/N)} \circ (\dots \circ \varphi_{(1/N)}))}_{N \text{ times}}.$$

If we particularize the mapping to a point y , we can obtain the action

$$\hbar^\mathbf{v} \star y \approx \underbrace{\varphi_{(1/N)}(\varphi_{(1/N)}(\dots(\varphi_{(1/N)}(y))))}_{N \text{ times}},$$

which for the case of $\varphi_{(1/N)} = \text{id}_{\mathbb{A}} + \frac{1}{N}\mathbf{v}$ resembles the Euler method for integrate the ODE (2.5.4)

$$\mathbf{x}_{(k/N)} \approx \mathbf{x}_{(k/N)} + \frac{1}{N}\mathbf{v}(\mathbf{x}_{(k/N)}),$$

with $\mathbf{x}_{(0)} = y$. The more the value of N , the more accurate the solution $\hbar^\mathbf{v} \star y$ (within the floating point precision).

The use of the one-parameter property can be performed in the converse order, $\varphi_{((k+1)/N)} = \varphi_{(k/N)} \circ \varphi_{(1/N)}$, or in general $\varphi_{((k+1)/N)} = \varphi_{((k-r)/N)} \circ \varphi_{((r+1)/N)}$ for any r . This property induces

the use of a *scaling and squaring scheme* [Higham 2005] [Arsigny 2006]

$$\begin{aligned}
 \varphi_{(\frac{1}{2^M})} &\approx \text{id}_{\mathbb{A}} + \frac{1}{2^M} \mathbf{v} \\
 \varphi_{(\frac{1}{2^{M-1}})} &\approx \varphi_{(\frac{1}{2^M})} \circ \varphi_{(\frac{1}{2^M})} \\
 \varphi_{(\frac{1}{2^{M-2}})} &\approx \varphi_{(\frac{1}{2^{M-1}})} \circ \varphi_{(\frac{1}{2^{M-1}})} \\
 &\vdots \\
 \varphi_{(\frac{1}{2})} &\approx \varphi_{(\frac{1}{4})} \circ \varphi_{(\frac{1}{4})} \\
 \exp(\mathbf{v}) \approx \varphi_{(1)} &\approx \varphi_{(\frac{1}{2})} \circ \varphi_{(\frac{1}{2})},
 \end{aligned}$$

where the first approximation is accurate for small M . For example, $M = 8$ is equivalent to an Euler process with $N = 256$ but only 8 evaluations of the mappings are required. However, these evaluations are usually performed by multi-linear interpolations on a structured grid. If we assume that the grid is defined in terms of the spatial frequencies of \mathbf{v} , for example using the Nyquist–Shannon sampling Theorem, then the first estimation $\varphi_{(1/2^M)}$ will be sufficiently accurate. In [Bergner et al. 2006], the frequency behaviour of the function composition was analysed. There, it was argued that the essentially band-limiting frequency of a composition of the form $h(x) = f(g(x))$ if given by $\nu_h = \nu_g \max_x |f'(x)|$, where ν_h and ν_g denote the maximum essential frequencies of h and g , respectively. Therefore, as the scaling and squaring process proceed, the grid describing the mappings should increase the resolution in order to remain the Nyquist–Shannon assumptions and guarantee accurate interpolations. As a final result, from the Euler process it can be seen that the limiting frequency for $\varphi_{(1)}$

$$\begin{aligned}
 \nu_{\varphi_{(1)}} &\approx \nu_{\mathbf{v}} \left(\max_x |(\mathcal{D}\varphi_{1/N})(x)| \right)^{N-1} \\
 &\approx \nu_{\mathbf{v}} \left(\max_x \left| \mathcal{D} \text{id}_{\mathbb{A}} + \frac{1}{N} \mathcal{D}\mathbf{v} \right| \right)^{N-1} \\
 &\approx \nu_{\mathbf{v}} \left(1 + \frac{1}{N} \max | \mathcal{D}\mathbf{v} | \right)^{N-1} \\
 \text{if } N \rightarrow +\infty, \quad \nu_{\varphi_{(1)}} &\approx \nu_{\mathbf{v}} \exp \left(\max_x | \mathcal{D}\mathbf{v} | \right).
 \end{aligned}$$

Therefore, an appropriate grid to describe $\varphi_{(1)}$ should have more spatial resolution than the one that describes \mathbf{v} . Moreover, the larger the amplitude of \mathbf{v} , the larger the value of $\max | \mathcal{D}\mathbf{v} |$ and the larger the frequencies of the final $\varphi_{(1)}$.

In practice, to avoid controlling those frequencies and manage different grid containers, we prefer explicit evolution processes such that the Euler scheme or a Runge–Kutta scheme with adaptive step-size to integrate Eq. (2.5.4) for each point.

Figure 2.5.1 shows the deformation mapping obtained by the integration of a SVF. The velocity field is illustrated by small arrows. Illustrative time evolution path for some points are shown as red curves.

Jacobian of deformation mappings

Another important aspect of the deformation mappings are the evaluation of their Jacobian. The common approach is to compute the Jacobian by a finite differences scheme. For example in 3D,

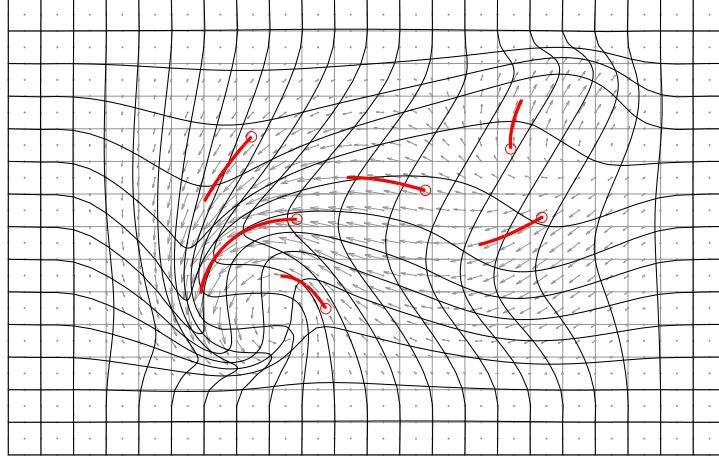


Figure 2.5.1. Integration of a stationary velocity field (SVF) in 2D. Illustrative evolution from $t = 0$ to $t = 1$ for some points are shown as red curves. Final deformation mapping $\exp(\mathbf{v}) = \varphi_{(t=1)}$ is shown as black grid.

given $\varphi: \mathbb{A} \equiv \mathbb{R}^3 \rightarrow \mathbb{A}$ the Jacobian is given

$$(\mathcal{D}\varphi)_{(y)} = \mathcal{D}_z\varphi(z)|_{(z=y)} = \begin{pmatrix} \partial_{y^1}(\varphi(y))^1 & \partial_{y^2}(\varphi(y))^1 & \partial_{y^3}(\varphi(y))^1 \\ \partial_{y^1}(\varphi(y))^2 & \partial_{y^2}(\varphi(y))^2 & \partial_{y^3}(\varphi(y))^2 \\ \partial_{y^1}(\varphi(y))^3 & \partial_{y^2}(\varphi(y))^3 & \partial_{y^3}(\varphi(y))^3 \end{pmatrix},$$

where

$$\partial_{y^i}(\varphi(y))^j \approx \frac{(\varphi(y + \epsilon e_j))^i - (\varphi(y - \epsilon e_j))^i}{2\epsilon},$$

with ϵ small enough, $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$ and $e_3 = (0, 0, 1)^T$ the canonical basis of \mathbb{R}^d .

A more accurate procedure to compute the Jacobian of a velocity generated spacial velocity (for both cases, time-dependent $\mathbf{w}(y, t)$ and SVF $\mathbf{v}(y)$) is by making use of Eqs. (2.5.3) or (2.5.5). For simplicity, let us consider the SVF case given in Eq. (2.5.5). Applying \mathcal{D} to both sides, it is obtained

$$\begin{aligned} \mathcal{D}\dot{\varphi}_{(t)} &= \mathcal{D}(\mathbf{v} \circ \varphi_{(t)}) \\ &= ((\mathcal{D}\mathbf{v}) \circ \varphi_{(t)}) (\mathcal{D}\varphi_{(t)}), \end{aligned}$$

where the last equality is obtained by applying the chain rule. Calling to the $d \times d$ Jacobian matrix $J = \mathcal{D}\varphi$ and exchanging the order in the derivatives $\mathcal{D}\dot{\varphi} = (\mathcal{D}\varphi)^\cdot = \dot{J}$, the following ODE is obtained

$$\dot{J}_{(x_{(t)})} = (\mathcal{D}\mathbf{v})_{(x_{(t)})} J_{(x_{(t)})},$$

where the corresponding evaluation point x follows a trajectory given by Eq. (2.5.4). Finally, to the Jacobian of the spatial transformation $\hbar^{\mathbf{v}} = \varphi_{(1)}$ at a spatial point $y \in \mathbb{A}$, $(\mathcal{D}\hbar^{\mathbf{v}})_{(y)}$ is given by the solution $J_{(y, 1)}$ of the following ODE

$$\begin{aligned} \dot{J}_{(y, t)} &= (\mathcal{D}\mathbf{v})_{(x_{(t)})} J_{(y, t)} \\ \dot{x}_{(t)} &= \mathbf{v}(x_{(t)}) \\ \text{with initial conditions} \quad J_{(y, 0)} &= I_d \\ \text{and} \quad x_{(0)} &= y. \end{aligned} \tag{2.5.6}$$

In a practical realization of the evolution, both the point \mathbf{x} and the associated Jacobian J are evolved simultaneously up to $t = 1$ obtaining the action $\mathcal{H}^Y \star \mathbf{y}$ and the Jacobian matrix $(\mathcal{D}\mathcal{H}^Y)_{(\mathbf{y})}$, respectively. In a numerical procedure to solve (2.5.6), with a step size Δ , at each iteration k the following update are performed:

- the time evolution t_k is updated to $t_k + \Delta$.
- the position of the point is updated by $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{v}(\mathbf{x}_k)$
- the Jacobian matrix can be update by $J_k = J_k + \Delta (\mathcal{D}\mathbf{v})_{(\mathbf{x}_k)} J_k$. However, this update does not preserve the underlying group structure of the Jacobian matrices (these concerns will be treated later in Chapters 4 and 5). A more accurate update can be performed by considering $(\mathcal{D}\mathbf{v})$ remains fixed along the evolution. With this assumption, the Jacobian matrix is update $J_{k+1} = e^{\Delta(\mathcal{D}\mathbf{v})_{(\mathbf{x}_k)}} J_k$, where $e^M = \sum_k M^k/k!$ is the matrix exponential function.

CHAPTER

3

Image registration

3.1	Registration of geometrical objects	50
3.2	Similarity and dissimilarity measures between geometrical objects	51
3.2.1	Distances between landmarks	51
3.2.2	Distance between unlabeled points	52
3.2.3	Distance between points and curves	52
3.2.4	Measures of the overlapping between regions	52
3.2.5	Dissimilarity measures between scalar images	53
3.3	Regularity measures	54
3.4	Registration using parametric spatial transformation models	55
3.4.1	Registration of scalar functions	58
3.5	Registration using object-dependent spatial transformation models	63
3.5.1	Exact matching of landmarks	64
3.5.2	Inexact matching formulation	66
3.5.3	Inexact matching for registering scalar functions	68
3.6	Registration using non-parametric spatial transformation models	69

3.1 Registration of geometrical objects

Let F (referring to *fixed*) and M (referring to *moving*) be two images, or two geometrical objects of any type. When both objects represent the same or similar content, it makes sense to identify features which exists in both objects. The relevant features to match could be of different nature: geometrical, such as corners or points at a high curvature location of a border; visual, such as points with a distinctive appearance, texture or color compared to its neighbours; structural or points with a high-level anatomical description, such as the *anterior- and posterior-commissures* of the brain (AC-PC). The goal of the registration process is to find a spatial transformation which achieves, as much as possible, the matching of those relevant features. Roughly, if the features of F are located at $\{f_i\}$ and their *corresponding* features in the object M are in locations $\{m_i\}$, then the registration consists in obtaining the spatial transformation h^* which best satisfies $f_i = h^* \star m_i$ for all i . Once h^* is found, the objects F and $h^* \star M$ have their significant features spatially aligned and ideally both object should share the same visual aspect. Nonetheless, h^* can act on any point of the ambient space, resulting in a dense correspondence, aligning not only the relevant features but also any point of the objects, and in general any point of the ambient space. For example, consider the geometrical objects shown in Figure 3.1.1, the significant features are, both at geometrical and anatomical level, the fingertips and the interdigital folds. It is expected that the registration puts in correspondence not only to those points, but also the whole contours. Moreover, a *flat point*, *i.e.* a point which is geometrically indistinguishable from its neighbours because their geometrical attributes are all the same, located nearby to the medial line of a finger is expected to be in correspondence with a point nearby to the medial line of the corresponding finger. Therefore, the registration of M towards F can be defined as the process of finding the spatial transformation h^* which best aligns (spatially) the significant features of F and $h^* \star M$ and which “best interpolates or extends” the transformation in between the significant features. This process will be denoted by $\mathcal{R}(M, F) = h^*$ such that every point of the fixed object F is spatially aligned with its corresponding point of the transformed object $\mathcal{R}(M, F) \star M$.

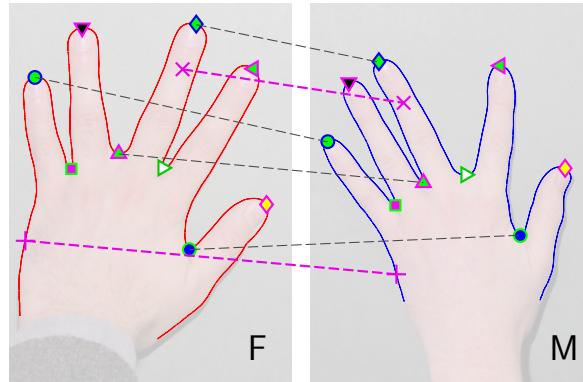


Figure 3.1.1. Correspondences between geometrical objects. The correspondences between the fingertips and the interdigital folds are well defined. The registration process is expected to be able to also find the correspondences for points '+' and 'x'.

However, there are missing formalisms in the previous definition of registration: how to quantify how much close are to geometrical objects or they significant features?; and, how to quantify how well the spatial transformation interpolates or extends the correspondences between the significant features?

It should be noted that exact correspondences are not known, in fact one of the objectives the main objective of the registration is to establish the spatial correspondences, although in the best case only a few sparse correspondences along the ambient space are known. In practice,

registration is formulated as to find the spatial transformation that minimizes a discrepancy or dissimilarity between objects at the same time that getting a spatial transformation with the maximum regularity as possible. This is induced by the fact that under some reasonable conditions about the process of the image acquisition, the less dissimilar be the images the more spatially matched will be their corresponding points. Therefore a by-product of the alignment of the corresponding points is that the geometrical objects look more and more similar [Rohlfing and Avants 2012]. Regarding the regularity of the transformation, it is introduced to allow to find the correspondences between flat points. When there is no available information in the geometrical attributes that makes preferable one alignment to another one, because both alignments results in the same discrepancy, then the more regular alignment should be preferred. This is a similar scenario of underdetermined linear systems which have more variables or degrees of freedom than equalities to satisfy. As in the case of ill-posed linear problems, the registration problem should be modified in order to give preferences to some suitable regular solutions. For example, if a point y lies *in-between* two points $a, b \in \mathbb{A}$, the transformed point $h * y$ should lie *in-between* the transformed points $h * a$ and $h * b$.

Then, the registration problem is usually formulated as the minimization of a functional of the form

$$\mathcal{R}_S(M, F) = \underset{h \in S}{\operatorname{argmin}} \text{reg}(h) + \lambda \text{dis}(F, h * M), \quad (3.1.1)$$

where $S \subseteq \mathcal{H}_{\mathbb{A}}$ is a set of spatial transformations and the notation \mathcal{R}_S makes explicit that the registration process is performed over spatial transformations from S . Also, **dis** quantifies the matching between both objects such that they look visually similar, **reg** is a measure of how much the spatial transformation deforms or wrinkles the ambient space, and λ is a tradeoff factor between both terms.

For interested readers, general reviews of the field may be found in several papers [Brown 1992] [Hawkes 1998] [Maintz and Viergever 1998] [Lester and Arridge 1999] [Hill et al. 2001] [Crum et al. 2003] [Zitova and Flusser 2003] [Crum et al. 2004] [Szeliski 2006] [Fleet and Weiss 2006] [Fischer and Modersitzki 2008] [Sotiras et al. 2013] [Oliveira and Tavares 2014] and books [Modersitzki 2004] [Goshtasby 2005] [Modersitzki 2009] [Hajnal and Hill 2010] [Goshtasby 2012].

3.2 Similarity and dissimilarity measures between geometrical objects

A *similarity measure* or *matching measure* is a tool to compute how much spatially aligned are the points of a geometrical object with respect to their corresponding points in another geometrical object [Modersitzki 2004]. Conversely, a *dissimilarity measure* quantifies how far are the corresponding points [Goshtasby 2012]. It is immediate to see that given a function **sim()**, the function **-sim()** can be used as dissimilarity.

In this section several commonly used dissimilarity measures for different kind of geometrical objects will be overviewed. For a wide and recent review of the subject refer to [Goshtasby 2012] [Sotiras et al. 2013].

3.2.1 Distances between landmarks

As was described in Section 2.2.1, landmarks are labeled spatial points which can be found consistently and with well defined correspondences along the studied objects. Let us start by considering that both geometrical objects, the fixed F and the moving M , can be described as a

ordered set of N landmarks such that $F = \{f_1, f_2, \dots, f_N\}$ and $M = \{m_1, m_2, \dots, m_N\}$, where each f_i is in correspondence with m_i .

A suitable and consistent dissimilarity measure between landmarks have to get its minimum value when all landmarks coincide in space. Using the fact that the ambient space (which is considered in this thesis to be the Euclidean space \mathbb{R}^2 or \mathbb{R}^3) is a metric space with distance function $\text{distance}_{\mathbb{A}}(\cdot, \cdot)$, there are several possible elections for the dissimilarity measure between landmarks such as $\max_l \text{distance}_{\mathbb{A}}(f_l, m_l)$ or $\sum_l \omega_l \text{distance}_{\mathbb{A}}(f_l, m_l)^\alpha$, where $\omega_l > 0$ is a weight for each landmark and $\alpha > 0$. May be the most algebraically convenient is

$$\text{dis}(F, M) = \sum_{l=1}^N \|f_l - m_l\|^2, \quad (3.2.1)$$

where $\|\cdot\|$ is the Euclidean norm. This is the most commonly used dissimilarity measure between geometrical objects defined by landmarks in a Euclidean space.

3.2.2 Distance between unlabeled points

Let F and M be two geometrical objects made up of collections of unlabeled points, $F = \{f_\alpha\}$ and $M = \{m_\alpha\}$. A dissimilarity measure between them can be defined by the *Hausdorff distance* defined as

$$\text{dis}(F, M) = \max \left\{ \sup_{f \in F} \inf_{m \in M} \text{distance}_{\mathbb{A}}(f, m), \sup_{m \in M} \inf_{f \in F} \text{distance}_{\mathbb{A}}(f, m) \right\}. \quad (3.2.2)$$

Several modified and more robust versions of this distance were proposed in [Dubuisson and Jain 1994], for example for discrete sets of points

$$\text{dis}(F, M) = \sum_{f \in F} \min_{m \in M} \text{distance}_{\mathbb{A}}(f, m) + \sum_{m \in M} \min_{f \in F} \text{distance}_{\mathbb{A}}(f, m).$$

3.2.3 Distance between points and curves

A dissimilarity measure between curves and points can be obtained inspired by Eq. (3.2.2).

Given F , representing a curve in \mathbb{A} , $F = \{\gamma(t) \mid t \in [0, L] \subset \mathbb{R}\}$, and M an object defined by N unlabeled points $M = \{m_i\}_{i=1}^N$, a dissimilarity measure between these two objects can be defined

$$\text{dis}(F, M) = \sum_{i=1}^N \min_{t \in [0, L]} \text{distance}_{\mathbb{A}}(m_i, \gamma(t)). \quad (3.2.3)$$

Unlike previous dissimilarity measures, note that dissimilarity given by Eq. (3.2.3) is not *symmetric*.

3.2.4 Measures of the overlapping between regions

Several similarity and dissimilarity measures between geometrical regions (interior of closed curves or surfaces) have been proposed in the literature. For an exhaustive survey see [Choi et al. 2010]. *Dice coefficient* and *Jaccard index* are the most used in the area of medical imaging. Both are defined in terms of the areas or volumes of regions. Given two geometrical regions F and M , Dice coefficient is $DC = 2 \frac{\text{area}(F \cap M)}{\text{area}(F) + \text{area}(M)}$ while Jaccard index is $JI = \frac{\text{area}(F \cap M)}{\text{area}(F \cup M)}$. Both measures provide values between 0 (not overlapping at all) and 1 (perfect overlap). Noticing that $\text{area}(F \cup M) = \text{area}(F) + \text{area}(M) - \text{area}(F \cap M)$ it is easy to show that both measures are related by $JI = DC/(2 - DC)$.

These overlapping measures are typically used to evaluate the performance of segmentation algorithms [Warfield et al. 2006] [Crum et al. 2006] or registration procedures [Gaens et al. 1998] [Klein et al. 2009] [Rohlfing 2012], but up to our knowledge, they are not commonly used as similarity measures to guide the registration process.

Another interesting measure is *Hamming distance* (also called *symmetric difference*) $HD = \text{area}(F) + \text{area}(M) - 2\text{area}(F \cap M)$. HD was used as a matching term to register shapes and to computed median shapes in [Berkels et al. 2010].

3.2.5 Dissimilarity measures between scalar images

An ideal dissimilarity criterion for scalar intensity images should result in low values when similar tissues classes are spatially aligned. There are two fundamental cases: when fixed and moving image come from the same acquisition process; and when they are from different modalities.

In the first case, called *mono-modal registration*, it is reasonable to consider that the appearance (intensity value and texture) of corresponding anatomical structures are the same. Different matching criteria have been devised to deal under the assumption of equal intensities on homologous structures. The most commonly used is the *sum of squared differences* (SSD). This dissimilarity measure is defined by

$$\mathbf{dis}(F, M) = \int_{y \in \mathbb{A}} (F_{(y)} - M_{(y)})^2 \quad (3.2.4)$$

and has optimal properties under an independent and identically distributed (i.i.d.) Gaussian noise model in the image acquisition process.

The assumption that the appearance is invariant along different subjects and acquisitions is rarely achieved. For example, a common situation in MR images is that inhomogeneities of the magnetic field produce a smoothly spatial disturbance of the intensity values [Hou 2006], although these inhomogeneities are typically removed by a preprocess step [Sled et al. 1998] [Manjón et al. 2007]. Therefore, other popular choices of matching term have been devised: *sum of absolute deviations* [Giachetti 2000]; *cross-correlation* criteria [Avants et al. 2008]; *correlation coefficient* [Kim and Fessler 2004]; and *correlation ratio* [Hermosillo et al. 2002]. Each of these matching criteria is based on different assumptions between the signal intensities of both images and on the noise model in the acquisition process [Roche et al. 2000].

The use of intensity information can lead to ambiguous correspondences for flat zones, *i.e.* an region with the same intensity vale for all points. For example, although the gradient of the image does not vanish at any place, a deformation tangential to the level sets of the intensity function shall result in the same deformed image. To cope with these shortcomings, several works propose to increase the dimensionality of the feature space by introducing attributes representing the geometric structure of the underlying anatomy. For example in [Shen and Davatzikos 2002], a vector with the local intensity moments are used as features to be matched. Other approaches include multi-scale Gabor filters as local characteristic [Ou et al. 2011], *local correlation coefficient* [Cachier et al. 2003] [Lorenzi et al. 2013], spherically distributed gray-level information [Liao and Chung 2012], or residual complexity [Myronenko and Song 2010].

For images from different modalities, information theoretic approaches, such as *mutual information* [Collignon et al. 1995] [Viola and Wells III 1997] [Pluim et al. 2003] or *Kullback–Leibler divergence* [Chung et al. 2002] [Guetter et al. 2005], have been proposed with well accepted success. Alternatively, several works proceed by reducing the multi-modality registration to a mono-modal one by simulating or generating one modality from the information of the other [Roche et al. 2001] [Wein et al. 2008] [Michel and Paragios 2010]. Finally, [Van den Elsen et al. 1995], [Droske and Rumpf 2004] and [Haber and Modersitzki 2006]

assumed that the anatomical and geometrical information is encoded in the borders of the structures and proposed the matching of the gradients or ridges of the image.

It is also possible to define matching terms between images or scalar functions and other types of objects. For example, a geometrical object represented by a curve can be matched with the edges presented in an image. This framework can be used as a connection between the registration process and the segmentation of structures within the images [Lombaert 2012] as it is performed by *snakes* techniques [Kass et al. 1988], by the geodesic active contours [Caselles et al. 1997], or by the *edge-free active contours* [Chan and Vese 2001].

A larger list of several matching measures used on different areas of image analysis can be found in [Goshtasby 2012, Chapter 2] while the most relevant measures used in medical imaging are reviewed in [Hajnal and Hill 2010] [Sotiras et al. 2013].

The choice of a dissimilarity measure or matching term is certainly application dependent, but it must take into account that it must satisfy that the better or more exact the obtained correspondences, the smaller the dissimilarity measure [Rohlfing 2012] [Rohlfing and Avants 2012].

3.3 Regularity measures

There are two main scenarios where a regularization term $\text{reg}(\cdot)$ must be included in the registration process. The first one is to favoring some transformations than others and avoid multiple solutions in the minimization problem. The second one refers to the case where the regularization, together with a probabilistic model, serves to introduce an *a priori* knowledge and reg is used to penalizing those spatial transformations with low probability.

For the case of parametric models of spatial transformations, due to their low degrees of freedom, they can be considered inherently smooth and regularization terms can only intervene if a probabilistic model is considered. The most important use of regularization terms is to avoid excessive wrinkling of deformation mappings when spatial transformation models are parametric models with a large number of parameters, object-dependent, or non-parametric models.

Three different frameworks are commonly used. The most popular one is inspired from *infinitesimal strain theory of elasticity* and quantify the regularization as a functional norm on the displacement field generated by spatial transformations. Given a spatial transformation $\boldsymbol{\kappa}^{\mathbf{u}}$ described by the displacement field \mathbf{u} , such that $\boldsymbol{\kappa}^{\mathbf{u}} \star \mathbf{y} = \mathbf{y} + \mathbf{u}_{(\mathbf{y})} = \mathbf{y} + \left(u_{(\mathbf{y})}^1, u_{(\mathbf{y})}^2, \dots, u_{(\mathbf{y})}^d \right)^T$. Regularization terms are typically *quadratic differential forms* $\text{reg}(\boldsymbol{\kappa}^{\mathbf{u}}) = \int_{\mathbf{y} \in \mathbb{A}} \mathbf{u}_{(\mathbf{y})}^T (\mathbf{L}\mathbf{u})_{(\mathbf{y})}$ where \mathbf{L} is a linear differential operator. Sometimes, the \mathbf{L} operator is a scalar operator which acts independently on each component of the displacement [Cachier and Ayache 2004], $\mathbf{L}\mathbf{u} = (Lu^1, \dots, Lu^d)^T$, and in these cases $\text{reg}(\boldsymbol{\kappa}^{\mathbf{u}}) = \sum_{c=1}^d \int_{\mathbb{A}} u^c (Lu^c)$. Most classic examples are: *linearized elastic potential* with $(\mathbf{L}\mathbf{u})^c = \mu \sum_i \partial_i^2 u^c + (\lambda + \mu) \partial_c (\sum_i \partial_i u^i)$, where μ and λ are the Lamé constants; the *diffusion smoothness* where $Lu^c = \sum_i \partial_i^2 u^c$; the *curvature energy* or *thin-plate splines* where $Lu^c = \sum_i \partial_i^4 u^c$. While the previous regularization functions are isotropic and rotationally invariant, a *locally oriented smoothness* induced by gradient intensities of the embedded images was proposed in [Nagel and Enkelmann 1986] and evaluated in optical flow scenarios.

The second framework considers non-linear operators on the deformation. The ambient space was modeled as hyper-elastic materials in [Rabbitt et al. 1995] [Droske and Rumpf 2004] [Rouchdy et al. 2007] and non-linear energies are associated to the deformation of the ambient space. [Le Guyader and Vese 2009] proposed a Saint Venant–Kirchhoff elasticity energy for a combined registration and segmentation approach and further results about the existence of

optimal minimizer can be found in [Derfoul and Le Guyader 2013]. A log-Euclidean treatment of the Cauchy–Green strain tensor was proposed in [Pennec et al. 2005] [Pennec 2006b] as an invariant measure of the regularity of the deformation. In [Yanovsky et al. 2008] the regularization term was derived from information theory leading to a deformation with minimal bias from compression-expansion maps. The total-variation on the gradient of the displacement field were used as regularization measure in [Hömke et al. 2007] [Zach et al. 2007] [Chumchob and Chen 2010]. An extension of the linear curvature functional where proposed in [Chumchob et al. 2011].

In the third framework the regularity or smoothness is measured on the velocity field which generates the spatial transformation. Every differentiable homeomorphism can be obtained by integrating a time-dependent velocity field, and a regularity measure can be obtained by endowing to the functional space of velocities with a functional norm [Dupuis et al. 1998] [Trouvé 1998] [Miller et al. 2002] [Younes 2010]. That is, for a spatial transformation $\hbar^{\mathbf{v}}$ obtained by the flow integration of $\mathbf{v}_{(t)}$, a metric on the \mathbf{v} space can be given by $\|\mathbf{v}\|_L^2 = \int_{t \in [0,1]} \int_{y \in \mathbb{A}} \mathbf{v}(t)_{(y)}^T (L\mathbf{v}(t))_{(y)}$. For a given spatial transformation \hbar there are several velocity field \mathbf{v}_i whose flows generate $\hbar^{\mathbf{v}_i} = \hbar$, that one with a minimal norm $\|\mathbf{v}\|_L$, which under some conditions of the operator L , infers a metric structure on the space of differentiable homeomorphisms. This metric can be used as a regularity measure on the spatial transformations: $\text{reg}(\hbar) = \inf_{\mathbf{v}} \{ \|\mathbf{v}\|_L \text{ such that } \hbar^{\mathbf{v}} = \hbar \}$. This approach was used as regularization for landmark matching [Joshi and Miller 2000] [Vaillant et al. 2004] [Marsland and Twining 2004], points sub-sets [Glaunès et al. 2004a], curves and surfaces [Vaillant and Glaunès 2005] [Cotter 2008], volumetric intensity images [Beg et al. 2005] [Avants et al. 2008], vector fields [Cao et al. 2005] and with combined objects like surfaces and intensities [Avants et al. 2006] [Auzias et al. 2011]. Alternatively, the regularization over the velocities can be using in the *stationary velocity field* setting [Ashburner 2007] [Vercauteren et al. 2008] [Hernandez et al. 2009] [Vercauteren et al. 2009] [Bossa et al. 2010b] [Lorenzi and Pennec 2013b]. Although under the SVF setting, velocity norms do not infer valid metrics on the spatial transformations, $\|\mathbf{v}\|_L^2 = \int_{\mathbb{A}} \mathbf{v}^T (L\mathbf{v})$ can still be used as a prior to regularize deformations in the registration process.

3.4 Registration using parametric spatial transformation models

Let us start by considering an example of registering two geometrical objects composed of N landmark points $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N)$ and $\mathbf{M} = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_N)$ in the two dimensional space $\mathbb{A} \equiv \mathbb{R}^2$ under the spatial transformation model $\mathcal{SIM}(2)$ given in Section 2.5.2. Every point can be represented by its coordinates in a fixed reference system, $\mathbf{f}_l = (f_l^1, f_l^2)^T$ and $\mathbf{m}_l = (m_l^1, m_l^2)^T$.

Every spatial transformation of $\mathcal{SIM}(2)$ model can be identified by $\hbar^{(\sigma, \theta, \bar{t})}$ where $\sigma \in \mathbb{R}^+$ is a scale factor around the origin of the coordinate system, θ is the angle of rotation around the origin, and $\bar{t} = (t^1, t^2)^T$ is the translation along both coordinates. The action on a point results $\hbar^{(\sigma, \theta, \bar{t})} \star \mathbf{m} = \sigma R_\theta \mathbf{m} + \bar{t}$, where R_θ is the rotation matrix $(\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix})$ with $s = \sin(\theta)$ and $c = \cos(\theta)$ and satisfies $R_\theta R_\theta^T = R_\theta^T R_\theta = I$ and $\det(R_\theta) = 1$.

Using the dissimilarity function given in Eq. (3.2.1) the registration problem to obtain the parameters of $\mathcal{R}_{\mathcal{SIM}(2)}(\mathbf{M}, \mathbf{F})$ results in

$$\begin{aligned} (\sigma^*, \theta^*, \bar{t}^*) &= \underset{\sigma, \theta, \bar{t}}{\operatorname{argmin}} \mathbf{dis}(\mathbf{F}, \hbar^{(\sigma, \theta, \bar{t})} \star \mathbf{M}) \\ &= \underset{\sigma, \theta, \bar{t}}{\operatorname{argmin}} \sum_{l=1}^N \|\mathbf{f}_l - (\sigma R_\theta \mathbf{m}_l + \bar{t})\|^2, \end{aligned}$$

which is equivalent to

$$\begin{aligned} & \min_{\sigma, R_\theta, \bar{t}} \sum_{l=1}^N \|f_l - (\sigma R_\theta m_l + \bar{t})\|^2 \\ & \text{subject to } R_\theta^T R_\theta = I \\ & \quad \det(R_\theta) = 1 \\ & \quad \sigma > 0. \end{aligned} \tag{3.4.1}$$

Problem (3.4.1) is known as the *Procrustes problem*¹ and can be solved algebraically using the *singular value decomposition* [Umeyama 1991].

The previous minimization can be decoupled noticing that

$$\min_{\sigma, R_\theta, \bar{t}} \sum_{l=1}^N \|f_l - \sigma R_\theta m_l - \bar{t}\|^2 = \min_{\sigma, R_\theta} \min_{\bar{t}} \sum_{l=1}^N \|f_l - \sigma R_\theta m_l - \bar{t}\|^2$$

and the optimal translation \bar{t}^* is obtained by

$$\begin{aligned} \bar{t}^* &= \operatorname{argmin}_{\bar{t}} \sum_{l=1}^N \|f_l - \sigma R_\theta m_l - \bar{t}\|^2 \\ &= \frac{1}{N} \sum_l \sigma R_\theta m_l - \frac{1}{N} \sum_l f_l, \end{aligned}$$

i.e. the optimal translation is the one which aligns the *center of mass* of both landmark sets. Thus, without loss of generality, it can be assumed that both landmark sets have their center of masses at the origin of coordinates.

To compute the optimal rotation, assuming that both centers are at the origin, note that the functional $\sum_l \|f_l - \sigma R_\theta m_l\|^2$ can be written as $\|F - \sigma R_\theta M\|^2$, where F and M are the $2 \times N$ matrices composed by all the landmarks coordinates of the objects, $F = (f_1, f_2, \dots, f_N) = \begin{pmatrix} f_1^1 & f_2^1 & \dots & f_N^1 \\ f_1^2 & f_2^2 & \dots & f_N^2 \end{pmatrix}$. The optimal rotation problem can be written as

$$\begin{aligned} \min_{R_\theta} \sum_{l=1}^N \|f_l - \sigma R_\theta m_l\|^2 &= \min_{R_\theta} \|F - \sigma R_\theta M\|^2 \\ &= \min_{R_\theta} \operatorname{trace}((F - \sigma R_\theta M)^T (F - \sigma R_\theta M)) \\ &= \operatorname{trace}(F^T F) + \sigma^2 \operatorname{trace}(M^T M) - \min_{R_\theta} 2\sigma \operatorname{trace}(F^T R_\theta M). \end{aligned}$$

To minimize the last term, let $UDV^T = FM^T$ be the singular value decomposition of FM^T , then $R_\theta^* = U \operatorname{diag}(1, \dots, \det(UV^T)) V^T$ is the optimal rotation matrix [Umeyama 1991] (see also [Horn 1987] [Horn et al. 1988]). Finally, the optimal scaling factor is obtained by $\sigma^* = \operatorname{trace}(R_\theta^* M F^T) / \operatorname{trace}(M^T M)$.

The resulting spatial transformation is illustrated in Figure 3.4.1. In order to highlight that the transformation is applied to the whole ambient space, a grid representing the coordinate system of \mathbb{A} is also shown.

The previous algebraic procedure is valid for any dimension d of the ambient space. However, it is not possible to obtain algebraic solutions for all parametric transformation models, even for simple models such as linear transformations. Then, an alternative formulation to obtain $\mathcal{R}(M, F)$ is by using an iterative descend minimization process.

¹ In some references the Procrustes problem does not include translations.

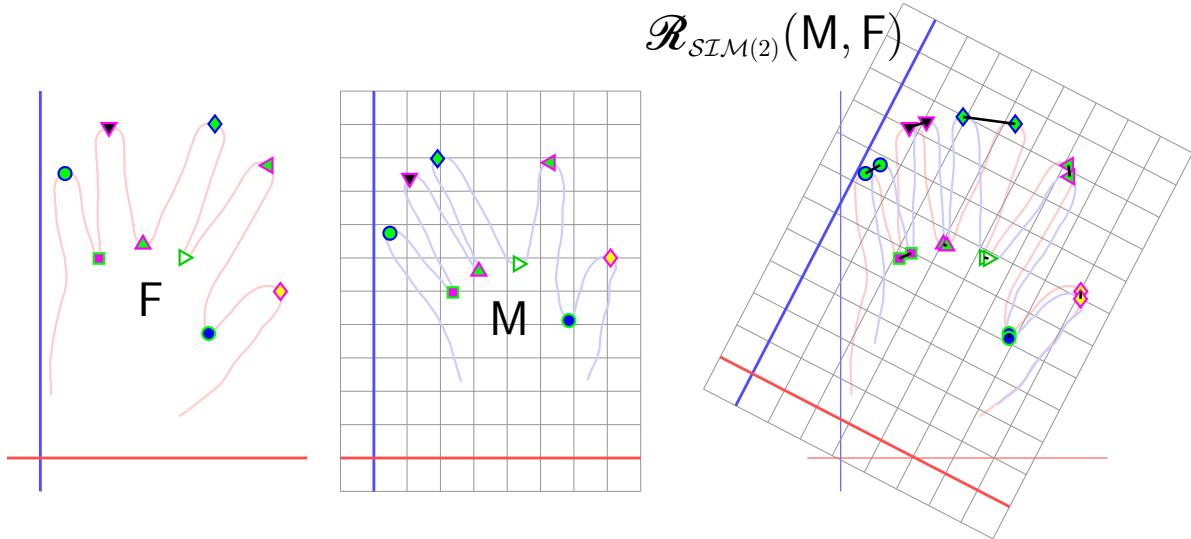


Figure 3.4.1. Registration of objects made up by landmarks under the action of $\text{SIM}(2)$.

The registration process returns the spatial transformation which acting on the moving object M get as close as possible their landmarks to the landmarks of the fixed object F .

For example, considering the special linear model $\mathcal{SL}(d)$ given in Section 2.5.2, each spatial transformation of this model can be identified by h^M , where $M \in \mathbb{M}_d$ fulfils $\det(M) = 1$. The action on a point is simply $h^M \star y = My$ and the registration process is performed by minimizing

$$\min_{M \in \mathbb{M}_d} \sum_{l=1}^N \|f_l - Mm_l\|^2$$

subject to $\det(M) = 1$,

which is a quadratic functional subject to a polynomial constraint and no algebraic solution is known.

This problem can be considered in a more general formulation. Almost all dissimilarity measures are the summation or the integral of individual matching terms, for example in the previous example $\text{dis}(F, M) = \sum_l \text{lwdis}_l(m_l)$ where the *location-wise dissimilarity* $\text{lwdis}_l(m_l) = \|f_l - m_l\|^2$. Let us assume that the space of spatial transformation considered in the registration is a *differentiable spatial transformation model* (see Section 2.3) described by $\Theta: \bar{p} \mapsto \Theta(\bar{p}) = h^{\bar{p}}$ where \bar{p} belongs to some open subset $\Omega \subset \mathbb{R}^m$, with m the degrees of freedom of the model. Then the registration process results

$$\min_{\bar{p} \in \Omega} \text{dis}(F, \Theta(\bar{p}) \star M) = \min_{\bar{p} \in \Omega} \sum_{l=1}^N \text{lwdis}_l(\Theta(\bar{p}) \star m_l),$$

which can be solved by classical optimization procedures such as a gradient descend method or any other like conjugate gradients, (quasi-)Newton optimizations or the Levenberg–Marquardt technique [Nocedal and Wright 2006] [Chong and Zak 2013]. In those methods the optimal parameter \bar{p}^* of $\text{dis}(\bar{p}) = \text{dis}(F, \Theta(\bar{p}) \star M)$ is obtained by successively approximate it by the sequence $\bar{p}_j, \bar{p}_{j+1}, \dots$, where $\bar{p}_{j+1} = \bar{p}_j + \alpha \bar{d}$, \bar{d} is a direction satisfying² $(\mathcal{D}_{\bar{p}} \text{dis}(\bar{p})) \bar{d} < 0$ and α is a step-size sufficiently small (although not so small!) such that $\text{dis}(\bar{p}_{j+1}) < \text{dis}(\bar{p}_j)$. Alternatively,

² The symbol \mathcal{D} refers to the *Fréchet derivative*. That is, given a function $f: \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}$ its Fréchet derivative at \bar{p} is the linear operator $(\mathcal{D}_{\bar{p}} f(\bar{p}))$ satisfying $(\mathcal{D}_{\bar{p}} f(\bar{p})) \bar{w} = \lim_{\epsilon \rightarrow 0} (f(\bar{p} + \epsilon \bar{w}) - f(\bar{p})) / \epsilon$, for all $\bar{w} \in \mathbb{R}^m$.

\bar{p}^* can be computed as the asymptotic solution in $t \rightarrow \infty$ of the following *initial value problem* (IVP)

$$\begin{cases} \dot{\bar{p}}(t) = \bar{d}_{(\bar{p}(t))} \\ \bar{p}(0) = \bar{p}_0 \end{cases},$$

with $\bar{d}_{(\bar{p}(t))}$ equals to, for example, $-\mathcal{D}_z \mathbf{dis}(z)|_{(z=\bar{p}(t))}$.

One of the key ingredients of those methods is the derivative of the function to minimize

$$\mathcal{D}_{\bar{p}} \mathbf{dis}(\mathsf{F}, \Theta(\bar{p}) \star \mathsf{M}) = \sum_l \mathcal{D}_z \mathbf{lwdis}_l(z)|_{(z=\Theta(\bar{p}) \star \mathsf{m}_l)} \mathcal{D}_{\bar{p}}(\Theta(\bar{p}) \star \mathsf{m}_l), \quad (3.4.2)$$

where terms under the summation come from the chain rule. In each of these terms, the first factor is commonly called *spatial force* or *image force* [Yanovsky 2008] and can be considered as a vector field from the ambient space to \mathbb{R}^d which describes for each point of \mathbb{A} in which direction is preferable to move it in order to increase the value of its location-wise dissimilarity. Flat points of a geometrical object are those having spatial force equals to zero, in this example, the points of the ambient space where no landmarks are defined. Regarding second factors from the chain rule in Eq. (3.4.2), they can be considered as the operator which projects the spatial forces over the parameters of the spatial transformation model such that a small perturbation in \bar{p} along that direction produces the displacements in \mathbb{A} which “best aligns” with the spatial forces. Notice that this factor is well defined since Θ was assumed to be a differentiable spatial transformation model.

Using the previous formulation, let us compute the registration under the action of $\mathcal{SL}(2)$ model. First note that all matrices $M \in \mathbb{M}_d$ with $\det(M) = 1$ can be described by its *singular value decomposition* (SVD) $M = UDV^T$ where U and V are $d \times d$ rotation matrices where each one can be described with $d(d - 1)/2$ parameters (in general, in SVD they are orthogonal matrices but a different convention can be chosen to force positive determinants of U and V by allowing negative singular values). Matrix D is a diagonal matrix with determinant equals to 1, $D = \text{diag}(\sigma^1, \sigma^2, \dots, \sigma^{d-1}, \prod_i 1/\sigma^i)$ with each $\sigma_j \neq 0$. Therefore, the set of spatial transformations $\mathcal{SL}(d)$ can be described by a differentiable spatial transformation model of $d^2 - 1$ degrees of freedoms. In particular, spatial transformations from $\mathcal{SL}(2)$ (or equivalently the matrix group $\mathbf{SL}(2)$) can be parameterized by

$$\Theta: \mathbb{R} \times \mathbb{R}_{\neq 0} \times \mathbb{R} \rightarrow \mathbf{SL}(2): (\theta_1, \sigma, \theta_2) \mapsto \begin{pmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ 0 & 1/\sigma \end{pmatrix} \begin{pmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{pmatrix},$$

with $s_i = \sin(\theta_i)$ and $c_i = \cos(\theta_i)$.

For the particular case of registering landmarks with the dissimilarity measure given in Eq. (3.2.1), $\mathbf{lwdis}_l(z) = \|\mathbf{f}_l - z\|^2$ and therefore spatial forces result $\mathcal{D}_z \mathbf{lwdis}_l(z) = 2(z - \mathbf{f}_l)$. That is, each landmark of the moving object prefers to move towards its corresponding landmark in the fixed object to decrease its local-wise dissimilarity (the negated spatial forces direction).

In Figure 3.4.2 the negated spatial forces are shown. In addition, the paths of the landmarks after the projection on the spatial transformation model are shown as green curves. The result of the registration $\mathcal{R}_{\mathcal{SL}(2)}(\mathsf{M}, \mathsf{F})$ is shown in Figure 3.4.3.

3.4.1 Registration of scalar functions

The previous concept of spatial force is also useful for registration of images. Remember that images are considered as continuous scalar functions which interpolate the values of the pixels or voxels. Although it is desirable that the resulting function be differentiable, in practice it is enough with an interpolation resulting in an almost everywhere differentiable function.

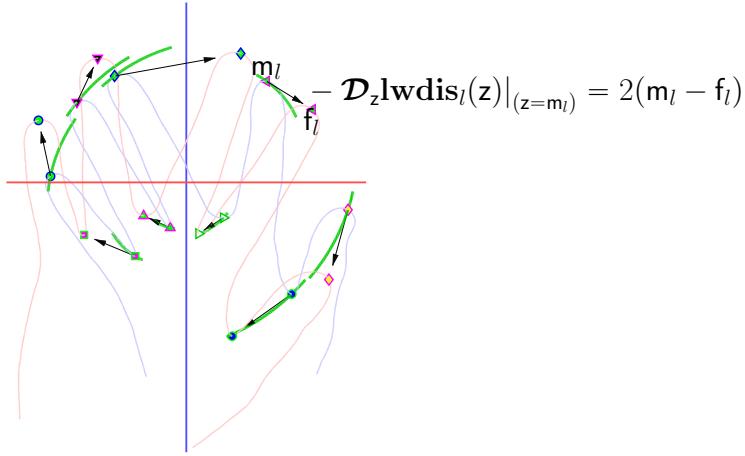


Figure 3.4.2. Illustration of the negated spatial forces. Each landmark of the moving object (the blue one) prefers to move towards its corresponding landmark in the fixed object (the red one) to decreased its local-wise dissimilarity. However, when these forces are projected on the spatial transformation model, each landmark move along the green curves to decrease the dissimilarity between both objects.

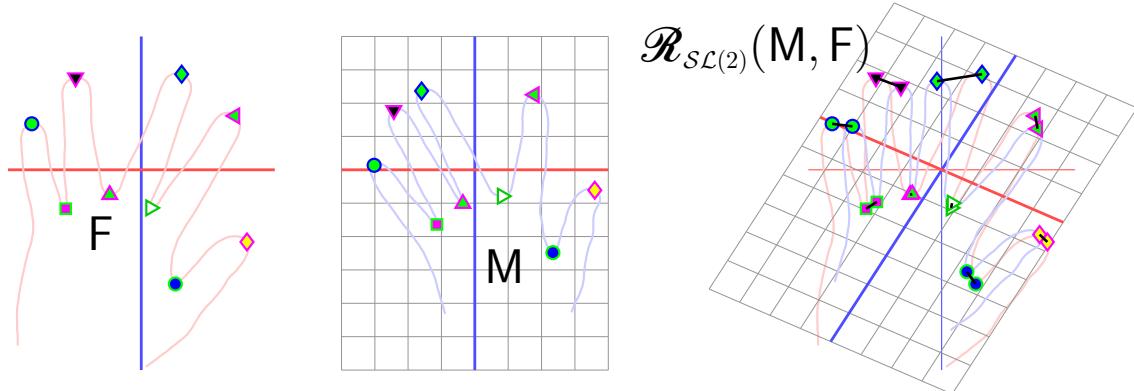


Figure 3.4.3. Registration of landmark objects under the $\mathcal{SL}(2)$ spatial transformation model.

As was already stated, the most extended dissimilarity measure between two images F and M is the *sum of squared differences* (SSD)

$$\text{dis}(F, M) = \int_{y \in \mathbb{A}} (F_{(y)} - M_{(y)})^2 .$$

As in the previous section, given a differentiable spatial transformation model Θ , the registration problem becomes

$$\min_{\bar{p} \in \Omega \subset \mathbb{R}^m} \text{dis}(F, \Theta(\bar{p}) \star M) = \min_{\bar{p} \in \Omega \subset \mathbb{R}^m} \int_{y \in \mathbb{A}} (F_{(y)} - (\Theta(\bar{p}) \star M)_{(y)})^2 . \quad (3.4.3)$$

In Section 2.2.2 it was already explained that $(\Theta(\bar{p}) \star M)_{(y)} = M_{(\Theta(\bar{p})^{-1} \star y)}$. In addition, the dissimilarity function can also be written in terms of location-wise dissimilarities $\text{lwdis}_y: \mathbb{A} \rightarrow \mathbb{R}: z \mapsto (F_{(y)} - M_{(z)})^2$ such that

$$\text{dis}(F, \Theta(\bar{p}) \star M) = \int_{y \in \mathbb{A}} \text{lwdis}_y(\Theta(\bar{p})^{-1} \star y) .$$

Taking derivatives with respect to \bar{p} it is obtained

$$\mathcal{D}_{\bar{p}} \text{dis}(\mathbf{F}, \Theta(\bar{p}) \star \mathbf{M}) = \int_{y \in \mathbb{A}} \mathcal{D}_z \text{lwdis}_y(z) |_{(z=\Theta(\bar{p})^{-1} \star y)} \mathcal{D}_{\bar{p}}(\Theta(\bar{p})^{-1} \star y), \quad (3.4.4)$$

where $\mathcal{D}_z \text{lwdis}_y(z) = 2(M_{(z)} - F_{(y)}) (\nabla M)_{(z)}$ which means that if $M_{(z)}$ is smaller than $F_{(y)}$ then is preferable to move the point z towards a higher value of M .

Within this formulation, the craft of the image registration process is to choose a convenient transformation model, compute Eq. (3.4.4) and perform a descend procedure in the space of the transformation model parameters $\Omega \subset \mathbb{R}^m$. There are also some techniques to improve the convergence of the optimization steps and also to make robust the process to local minimal [Klein 2008] [Sun et al. 2013] such as perform successive optimizations at different space-scales.

Let us take as example the registration of binary images under similarity transformations model $SIM(2)$. The action of these spatial transformations is given by

$$\Theta(\sigma, \theta, \bar{t}) \star y = \sigma \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \end{pmatrix} + \begin{pmatrix} t^1 \\ t^2 \end{pmatrix},$$

where the domain of Θ is $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$ and with $c = \cos(\theta)$ and $s = \sin(\theta)$.

After some algebraic manipulation, it can seen that

$$\begin{aligned} \mathcal{D}_{\bar{p}}(\Theta(\bar{p})^{-1} \star y) &= (\partial_\sigma, \partial_\theta, \partial_{t^1}, \partial_{t^2})(\Theta(\bar{p})^{-1} \star y) \\ &= \frac{1}{\sigma} \begin{pmatrix} \frac{-c(y^1-t^1)-s(y^2-t^2)}{\sigma} & -s(y^1-t^1)+c(y^2-t^2) & -c & -s \\ \frac{s(y^1-t^1)-c(y^2-t^2)}{\sigma} & -c(y^1-t^1)-s(y^2-t^2) & s & -c \end{pmatrix} (3.4.5) \end{aligned}$$

Given fixed and moving images, \mathbf{F} and \mathbf{M} respectively, as is shown in Figure 3.4.4. The registration $\mathcal{R}_{SIM(2)}(\mathbf{M}, \mathbf{F})$ was computed with the previous procedure and with some strategies such as applying smoothing filters to allow the derivability of \mathbf{M} . The result is shown in the right-most panel of Figure 3.4.4.

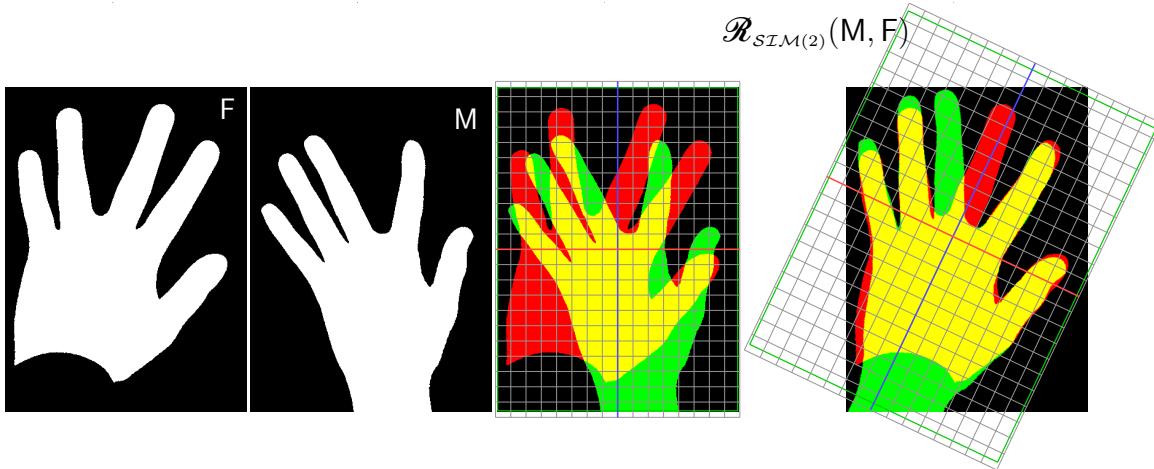


Figure 3.4.4. Registration of scalar functions under the action of $SIM(2)$ spatial transformation model.

In the previous form, the registration of images is making use of the Eulerian action of the transformations, unlike the landmarks matching case where the spatial transformation acts

in Lagrangian form over the landmarks. In the landmarks case, the spatial forces “push” the landmarks towards a better configuration while in the Eulerian form the spatial forces “pull” the moving image in order to decrease the dissimilarity. However, an alternative Lagrangian framework can be formulated for images. A new transformation model $\tilde{\Theta}$ can be proposed such that $\tilde{\Theta}(\bar{q}) = \text{inv}(\Theta(\bar{p}))$ for $\bar{q} \in \tilde{\Omega} \subseteq \mathbb{R}^m$. In fact, if the original transformation model Θ is able to describe the direct and the inverses transformations (which is the case of, for example, spatial transformation groups), then $\tilde{\Theta}$ can be selected to be the same as Θ . In terms of $\tilde{\Theta}$ the registration process results in the following optimization problem

$$\min_{\bar{q} \in \tilde{\Omega} \subseteq \mathbb{R}^m} \int_{y \in \mathbb{A}} \left(F(y) - M_{(\tilde{\Theta}(\bar{q}) * y)} \right)^2.$$

Taking derivatives with respect to \bar{q} it is obtained

$$\mathcal{D}_{\bar{q}} \int_{y \in \mathbb{A}} \left(F(y) - M_{(\tilde{\Theta}(\bar{q}) * y)} \right)^2 = \int_{y \in \mathbb{A}} \mathcal{D}_z \text{lwdis}_y(z)|_{(z=\tilde{\Theta}(\bar{q}) * y)} \mathcal{D}_{\bar{q}}(\tilde{\Theta}(\bar{q}) * y). \quad (3.4.6)$$

While Eq. (3.4.3) is understand as transform the moving image M and interpolate it at fixed coordinates y , Eq. (3.4.6) can be understand as: first, assign to each point y of the ambient space with the intensity value $F(y)$; and second, transform those points and interpolate on they the intensity values of M (remaining fixed at its original position); finally, perform the optimization in order to minimize the differences between the assigned and the interpolated values. In this Lagrangian formulation, transformations “push” the query points towards their best position in an immovable M instead to “pull” M and sample it on fixed points. Within this formulation the registration of images is understood in the same fashion than the registration of other types of geometrical objects, where the transformations acts “actively” on the constituent points of the objects. Moreover, by choosing an algebraically simple form of $\tilde{\Theta}$ simplifies the computation of the projection $\mathcal{D}_{\bar{q}}(\tilde{\Theta}(\bar{q}) * y)$ (compare with Eq. (3.4.5)) leaving the inversion of $\tilde{\Theta}(\bar{q}^*)$ to the end of the process and only to report the final result of $\mathcal{R}(M, F)$.

In Figure 3.4.5 it is shown an intermediate step of the optimization considering the Lagrangian formulation. The points of the ambient space are initially assigned with the corresponding intensity value of the fixed image (points on high level intensity are shown in red). Image forces act on each point to move it towards a zone with an intensity more similar to its original assignment. The moving image M remains fixed and the query points, together with their assigned intensity values, are transformed actively by $\tilde{\Theta}(\bar{q})$.

It is interesting to remark that this construction is, in general, not equivalent to reformulating the registration problem as find the transformation which minimize $\text{dis}(\Theta(p)^{-1} * F, M)$. Let us consider the optimization given in Eq. (3.4.3)

$$\int_{y \in \mathbb{A}} \left(F(y) - M_{(\Theta(\bar{p})^{-1} * y)} \right)^2$$

and perform the variable substitution $x = \Theta(\bar{p})^{-1} * y \Rightarrow y = \Theta(\bar{p}) * x$, then the integrand results

$$\begin{aligned} (F(y) - M_{(\Theta(\bar{p})^{-1} * y)})^2 dy &= (F_{(\Theta(\bar{p}) * x)} - M_{(x)})^2 \det(\mathcal{D}\Theta(\bar{p}))|_x dx \\ &= ((\Theta(\bar{p})^{-1} * F)(x) - M_{(x)})^2 \det(\mathcal{D}\Theta(\bar{p}))|_x dx, \end{aligned}$$

which have the same optimum than $\text{dis}(\Theta(p)^{-1} * F, M)$ only when the spatial transformation model gives spatially constant $\det(\mathcal{D}\Theta(\bar{p}))$. Moreover, if a regularization term is added to the registration, the tradeoff between matching and smoothness must be appropriately adjusted [Sabuncu et al. 2009].

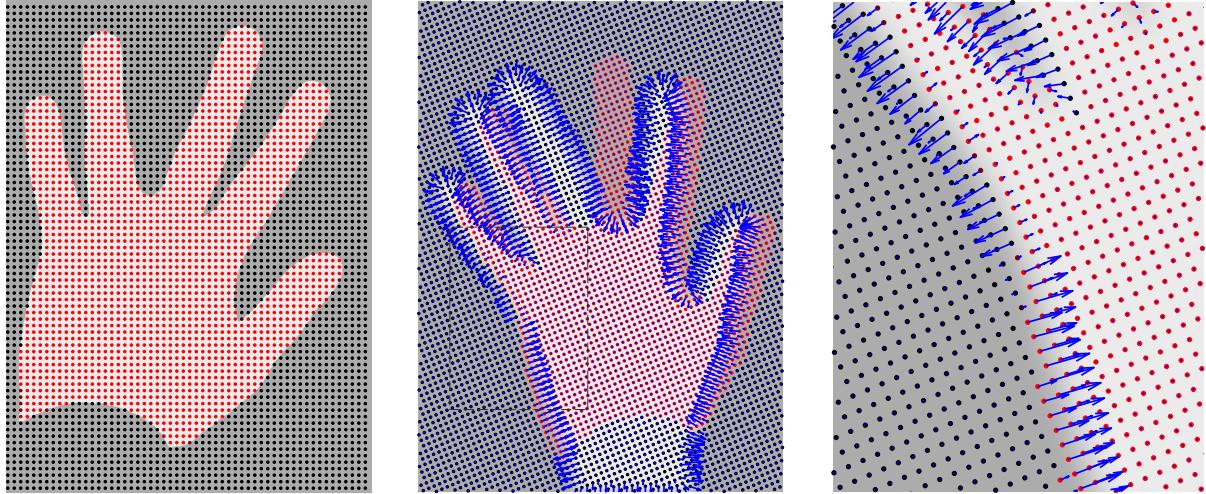


Figure 3.4.5. Image forces acting on the query points. The image forces “push” the points towards a zone in a immovable M with an intensity value more similar to their assignments.

Figure 3.4.6 shows registration results of sagittal slices from brain MRIs under $\mathcal{SE}(2)$, $\mathcal{SIM}(2)$ and $\mathcal{GA}^+(2)$ spatial transformation models. It can be some differences in the matching of scalps, ventricles and skull zones surrounding cerebellum. Roughly, spatial transformation models with more degrees of freedom get better alignments.

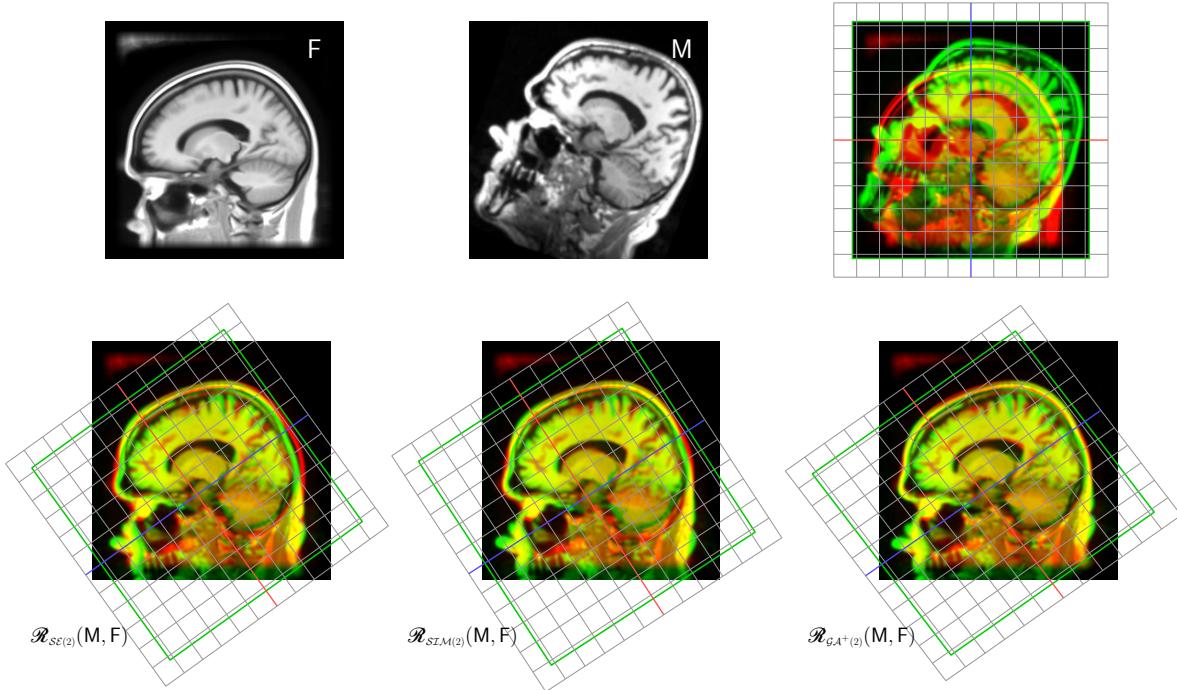


Figure 3.4.6. Sagittal slices from brain MRIs registered under different spatial transformation models.

The same procedures can be applied to volumetric 3D images. It is common in the area of medical imaging that the volumetric images are provided together with their specific location in the coordinate system given by the acquisition device. Figure 3.4.7 show the registration of two

brain MRIs (T1-weighted) under affine transformations in 3D ($\mathcal{GA}^+(3)$ model).

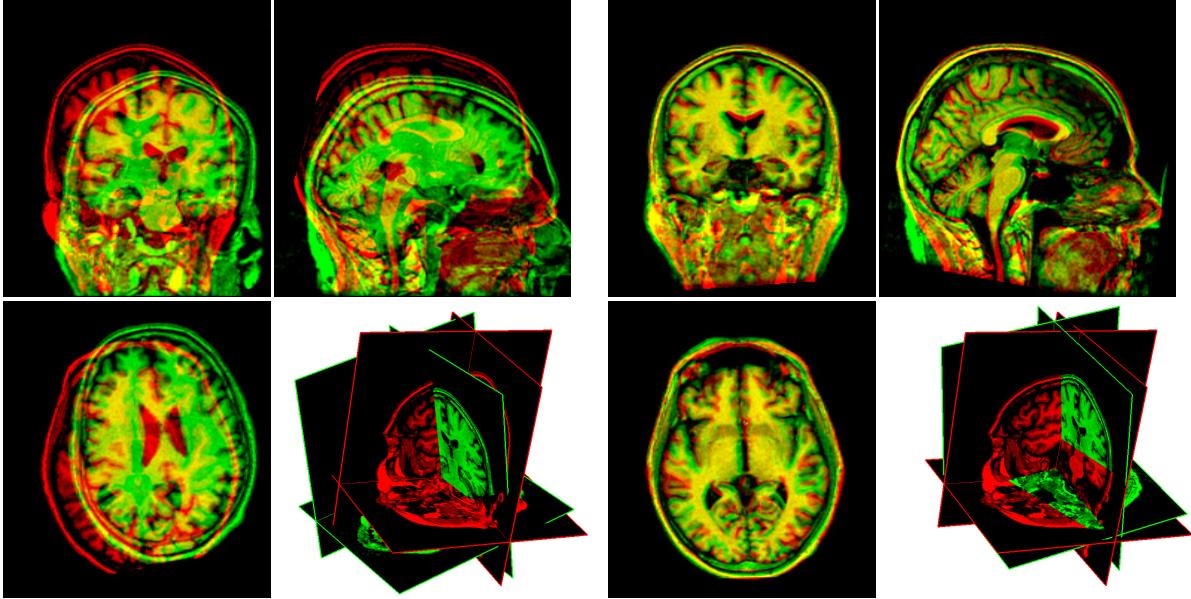


Figure 3.4.7. Registration of two volumetric brain MRIs under $\mathcal{GA}^+(3)$ spatial transformation model. Left: images at original position (in the coordinate system of the acquisition device). Right: images after registration, moving image M (shown in green) is registered towards the fixed image F (shown in red).

3.5 Registration using object-dependent spatial transformation models

As can be seen clearly in the previous examples shown in Figure 3.4.1 and Figure 3.4.3 parametric transformation models usually do not have enough flexibility to perform accurate registrations. They are not even able to achieve alignment between the landmarks and they find much less correspondences of other non-landmarked locations, which is the major objective of the registration. Therefore, it is necessary to increase the degrees of freedom of the spatial transformation models. An initial solution is to use object-dependent models in the small deformation paradigm.

An example of object-dependent models is where the ambient space is deformed guided by a finite set of points. These models are parameterized by the points together with their displacements and the deformation on the whole ambient space must be interpolated from those displacements. Let us start working in the small-deformation paradigm where transformations $h^{\mathbf{u}}$ are described by the displacement field $\mathbf{u}: \mathbb{A} \rightarrow \mathbb{R}^d$ such that the action is $h^{\mathbf{u}} \star y = y + \mathbf{u}(y)$. In this small-deformation paradigm, the only imposed constraint on the displacement field is that it must be a continuous function of \mathbb{A} and if the displacements are small enough, the whole transformation will be a valid homeomorphism.

There are two common approaches to work with the sparse displacements and their interpolation: the exact matching formulation; and the inexact matching formulation.

3.5.1 Exact matching of landmarks

In the case of registering geometrical objects made up by landmarks, $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N)$ and $\mathbf{M} = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_N)$, the only important displacements in the matching term given by Eq. (3.2.1) are $\mathbf{u}(\mathbf{m}_l)$. The displacements at other points \mathbf{x} of \mathbb{A} do not intervene in the dissimilarity function and therefore any displacement $\mathbf{u}(\mathbf{x})$ is allowed, *i.e.* the points of \mathbb{A} that are not landmarks are flat points. In order to prevent the arbitrariness of the solution of $\mathcal{R}(\mathbf{M}, \mathbf{F})$ a regularization term is considered on the displacement field. In the exact matching approach, the registration is formulated as

$$\begin{aligned} & \underset{\mathbf{h}\mathbf{u}}{\text{minimize}} \quad \mathbf{reg}(\mathbf{h}\mathbf{u}) \\ & \text{subject to} \quad \mathbf{h}\mathbf{u} \star \mathbf{m}_l = \mathbf{f}_l \quad \text{for } l = 1, \dots, N, \end{aligned} \quad (3.5.1)$$

for an appropriate regularization function $\mathbf{reg}(\cdot)$ (see Section 3.3).

The most typical form of regularization term is by consider energies of the form

$$\mathbf{reg}(\mathbf{h}\mathbf{u}) = \sum_{c=1}^d \int_{\mathbb{A}} u^c(Lu^c),$$

where L is a scalar, linear and differential operator together with some boundary conditions and u^c are the components of \mathbf{u} such that $\mathbf{u}(\mathbf{y}) = \left(u_{(y)}^1, u_{(y)}^2, \dots, u_{(y)}^d \right)^T$.

Lagrange multipliers on the constraints can be used to solve Eq. (3.5.1), resulting in the augmented functional

$$\mathbf{reg}'(\mathbf{h}\mathbf{u}) = \sum_{c=1}^d \int_{\mathbb{A}} u^c(Lu^c) + \sum_{l=1}^N \bar{\lambda}_l^T (\mathbf{m}_l + \mathbf{u}(\mathbf{m}_l) - \mathbf{f}_l),$$

where $\bar{\lambda}_l = (\lambda_l^1, \lambda_l^2, \dots, \lambda_l^d)^T$ is a vector-valued Lagrange multipliers for each landmark. Extremal conditions of Eq. (3.5.1) are (see [Marsland and Twining 2004])

$$\partial_{\bar{\lambda}_l} \mathbf{reg}' = 0 \Rightarrow \mathbf{u}(\mathbf{m}_l) = \mathbf{f}_l - \mathbf{m}_l \quad (3.5.2)$$

$$\partial_{u^c} \mathbf{reg}' = 0 \Rightarrow (Lu^c)(x) = -\frac{1}{2} \sum_{l=1}^N \lambda_l^c \delta(x^c - f_l^c), \quad (3.5.3)$$

where $\delta(\cdot)$ is the Dirac delta distribution. Due to the linearity of the the operator L , solutions of problem (3.5.1) take the form

$$u^c(x) = \sum_k^K \alpha_k^c \mathbf{g}_k(x) + \sum_{j=1}^N \beta_j^c \mathbf{G}_{\mathbf{m}_j}(x), \quad (3.5.4)$$

where \mathbf{g}_k are K linear independent functions in the null space of the operator L , *i.e.* $(L\mathbf{g}_k)(x) = 0$ for all x , and $\mathbf{G}_y(x)$ are the *Green's functions* of the operator L , *i.e.* $(L\mathbf{G}_y)(x) = \delta(x - y)$.

Replacing Eq. (3.5.4) in Eq. (3.5.1) the following problem is obtained for each component $c = 1, \dots, d$:

$$\begin{aligned} & \underset{u^c}{\text{minimize}} \quad \left(\int_{\mathbb{A}} u^c(Lu^c) = \sum_{i=1}^N \sum_{j=1}^N \beta_i^c \beta_j^c \mathbf{G}_{\mathbf{m}_i}(\mathbf{m}_j) \right) \\ & \text{subject to} \quad \sum_k^K \alpha_k^c \mathbf{g}_k(\mathbf{m}_l) + \sum_{j=1}^N \beta_j^c \mathbf{G}_{\mathbf{m}_j}(\mathbf{m}_l) = \mathbf{f}_l^c - \mathbf{m}_l^c \quad \text{for } l = 1, \dots, N. \end{aligned} \quad (3.5.5)$$

The optimal coefficients α_k^c and β_j^c to solve problem (3.5.5) can be computed by the *quadratic programming problem with linear constraints*

$$\begin{aligned} & \underset{\bar{\alpha}^c, \bar{\beta}^c}{\text{minimize}} \quad \begin{pmatrix} \bar{\beta}^c \\ \bar{\alpha}^c \end{pmatrix}^T \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{\beta}^c \\ \bar{\alpha}^c \end{pmatrix} \\ & \text{subject to} \quad (G \quad g) \begin{pmatrix} \bar{\beta}^c \\ \bar{\alpha}^c \end{pmatrix} = \bar{d}^c, \end{aligned}$$

with $N \times N$ matrix $(G)^{i,j} = \mathbf{G}_{\mathbf{m}_i}(\mathbf{m}_j)$, together with the $N \times K$ matrix $(g)^{l,k} = \mathbf{g}_k(\mathbf{m}_l)$ and the N -dimensional constraint vector $(d^c)^l = \mathbf{f}_l^c - \mathbf{m}_l^c$. After some algebraic work, solutions of this particular quadratic programming problem are

$$\begin{aligned} \bar{\alpha}^c &= (g^T G^{-1} g)^{-1} g^T G^{-1} \bar{d}^c \quad (3.5.6) \\ \bar{\beta}^c &= G^{-1} (\bar{d}^c - g \bar{\alpha}^c) \end{aligned}$$

$$= G^{-1} \left(I_N - g (g^T G^{-1} g)^{-1} g^T G^{-1} \right) \bar{d}^c. \quad (3.5.7)$$

In general, spatial transformations \mathfrak{h} can be described by specifying the displacements (\bar{d}_n) of a set of N sparse *knotpoints* (\mathbf{p}_n) , such that $\mathfrak{h}^{(\mathbf{p}_n)}(\bar{d}_n) \star \mathbf{p}_i = \mathbf{p}_i + \bar{d}_i$ for all $i = 1, \dots, N$. To compute the action on any point $y \in \mathbb{A}$, first the displacement field \mathbf{u} is computed by solving problem (3.5.1) with Eqs. (3.5.6) and (3.5.7), and the action $\mathbf{h} \star y = y + \mathbf{u}(y)$. Then, given N fixed knotpoints (\mathbf{p}_n) and for a particular operator L , the spatial transformation model is defined by specifying the N displacements $\Theta((\bar{d}_n)) = \mathfrak{h}^{(\mathbf{p}_n)}(\bar{d}_n)$. Its action is

$$(\Theta((\bar{d}_n)) \star y)^c = \sum_{k=1}^K \alpha_k^c \mathbf{g}_k(y) + \sum_{n=1}^N \beta_n^c \mathbf{G}_{\mathbf{p}_n}(y).$$

Replacing the α s and β s coefficients it is obtained the important fact that for each point $y \in \mathbb{A}$ there exist a linear functional D_y such that

$$(\Theta((\bar{d}_n)) \star y)^c = y^c + D_y \bar{d}^c, \quad (3.5.8)$$

where D_y is a $1 \times N$ row matrix given by

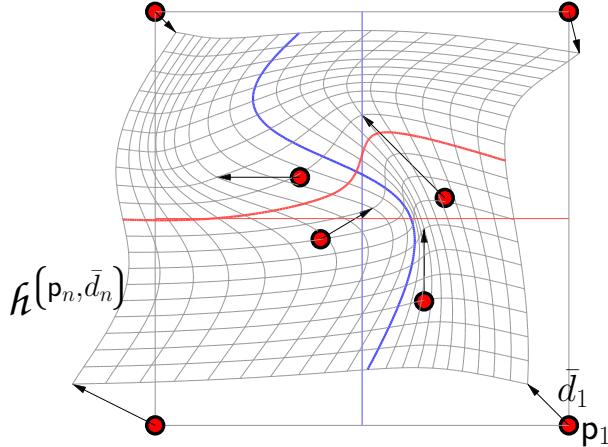
$$D_y = (\mathbf{G}_{\mathbf{p}_1}(y), \dots, \mathbf{G}_{\mathbf{p}_N}(y), \mathbf{g}_1(y), \dots, \mathbf{g}_K(y)) \begin{pmatrix} G^{-1} \left(I_N - g (g^T G^{-1} g)^{-1} g^T G^{-1} \right) \\ (g^T G^{-1} g)^{-1} g^T G^{-1} \end{pmatrix}$$

Moreover, the regularization term results in

$$\begin{aligned} \mathbf{reg}(\Theta((\bar{d}_n))) &= \sum_{c=1}^d (\bar{\beta}^c)^T G (\bar{\beta}^c) \\ &= \sum_{c=1}^d (\bar{d}^c)^T R(\bar{d}^c), \end{aligned} \quad (3.5.9)$$

with $R = \left(I - g (g^T G^{-1} g)^{-1} g^T G^{-1} \right)^T G^{-1} \left(I - g (g^T G^{-1} g)^{-1} g^T G^{-1} \right)$.

Finally, for a given linear operator L , if its Green's functions $\mathbf{G}_y(x)$ and its kernel $\text{span}(\{\mathbf{g}_1, \dots, \mathbf{g}_K\})$ are known, the spatial transformation $\mathfrak{h}^{(\mathbf{p}_n)}(\bar{d}_n)$ is completely defined by



An example of Green's function based displacement field interpolation.

the algebraic operations summarized in this section. Some commonly used L operators and their Green's functions are listed in [Marsland and Twining 2004] [Twining and Marsland 2008]

When no known Green's functions in algebraic form exists for L , numerical procedures can be performed to compute solutions u^c of problem (3.5.5). These numerics can be based on spectral methods [Cachier and Ayache 2004] or in approximations of the operator L as a convolution operator by sparse matrices.

To illustrate the method, some results for different L operators are shown in Figure 3.5.1.

3.5.2 Inexact matching formulation

The second approach for the alignment of sparse geometrical objects is by considering a generative model of the instances. Again, let us consider geometrical objects made up by N landmarks and let us assume that an instance of a geometrical object is obtained by a random process of the form $\mathbf{X} = \mathbf{h} \star \mathbf{T} + \varepsilon$, where \mathbf{T} is a *template* object composed by the landmarks $(\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N)$. Regarding \mathbf{h} , it is a spatial transformation with a probability density distribution of the form $\Pr(\mathbf{h}) \propto e^{-\text{reg}(\mathbf{h})}$ such that the larger the regularity of \mathbf{h} , the smaller the chance to get \mathbf{h} . Regarding ε , it can be considered as a perturbation process or as measure error with a known probability density distribution. For this landmark case let us assume that each of the N landmarks is affected by independent and identically distributed Gaussian noise such that, for all $l = 1, \dots, N$, every component ε_l^c distributes as a Gaussian with zero mean and standard deviation $\sigma \mathcal{N}(0; \sigma)$.

The idea of the inexact matching approach is to find what is the most likely spatial transformation \mathbf{h} which have generated a measured object \mathbf{X} when the atlas \mathbf{T} and the distribution of ε are known. According to *Bayes' rule* $\Pr(\mathbf{h}|\mathbf{X}) \Pr(\mathbf{X}) = \Pr(\mathbf{h}) \Pr(\mathbf{X}|\mathbf{h})$, and taking into account the ε distribution

$$\Pr(\mathbf{X}|\mathbf{h}) = \prod_{l=1}^N \Pr(\mathbf{x}_l - \mathbf{h} \star \mathbf{t}_l) \propto \prod_{l=1}^N e^{-\frac{\|\mathbf{x}_l - \mathbf{h} \star \mathbf{t}_l\|^2}{\sigma^2}}.$$

The *maximum a posteriori* estimator of \mathbf{h} , knowing \mathbf{X} and \mathbf{T} , is the one that minimizes

$$\text{reg}(\mathbf{h}) + \frac{1}{\sigma^2} \sum_{l=1}^N \|\mathbf{x}_l - \mathbf{h} \star \mathbf{t}_l\|^2 = \text{reg}(\mathbf{h}) + \frac{1}{\sigma^2} \mathbf{dis}(\mathbf{X}, \mathbf{h} \star \mathbf{T}),$$

where $\mathbf{dis}(\cdot, \cdot)$ is the dissimilarity matching function given in Section 3.2.1.

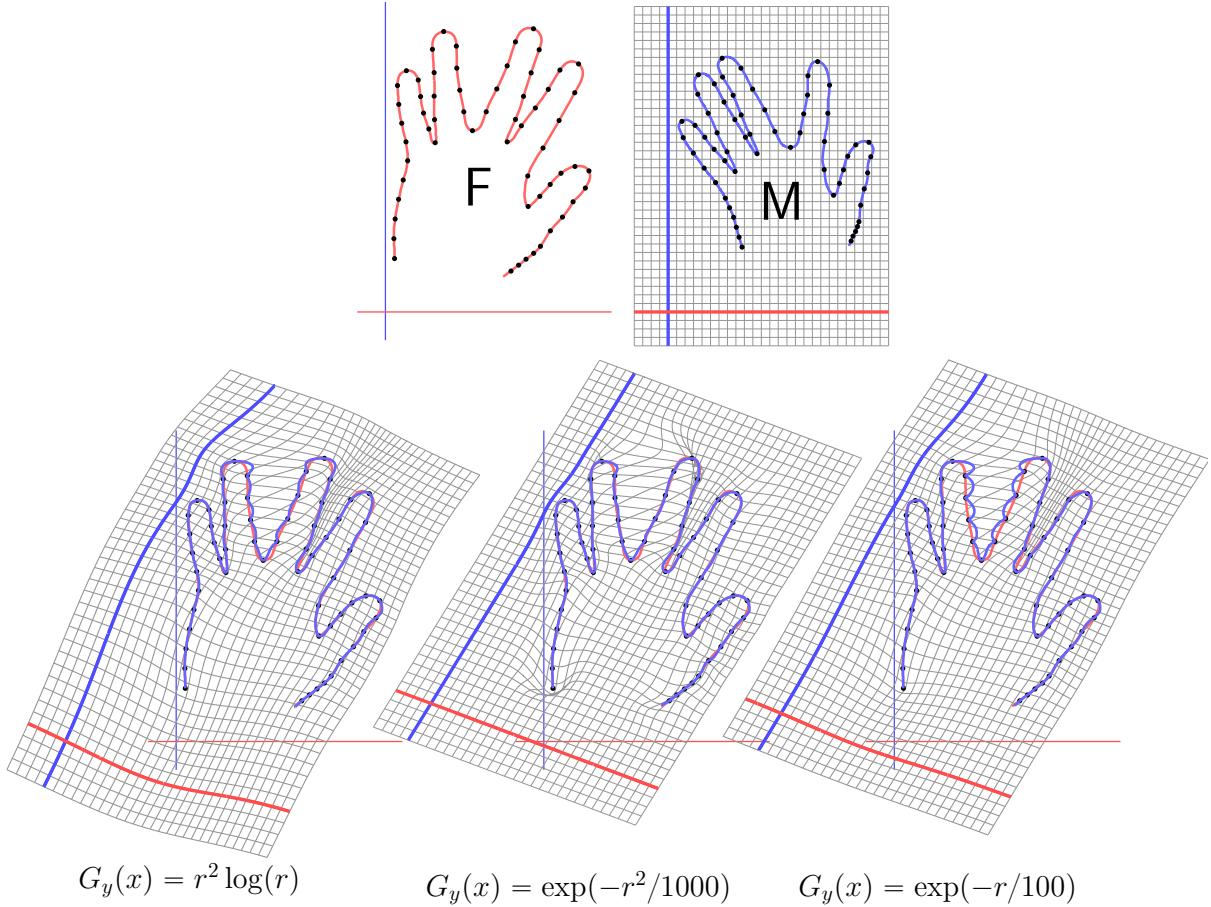


Figure 3.5.1. Matching landmarks under three different L operators. The Green's coordinates are explicitly given for each example in term of $r = \|x - y\|$.

Comparing with the exact matching approach described in the previous section, in the current inexact approach the dissimilarity term acts as *soft-constraint* with a tradeoff between the matching and the regularity terms given by σ^2 , *i.e.* the noise level considered in the generative model. The smaller the value of σ , the closer the solution to the exact matching case.

As an example, let us consider the registration of landmark objects T and X using as spatial transformation model the displacements (\bar{d}_n) on some fixed knotpoints (p_n) . To measure the regularity, an operator L with Green's function $\mathbf{G}_y(x)$ can be used. Following Eq. (3.5.9), the inexact matching problem can be rewritten as

$$\underset{(\bar{d}_n)}{\text{minimize}} \quad \mathbf{dis}(X, \Theta((\bar{d}_n)) \star T) + \sigma^2 \sum_{c=1}^d (\bar{d}^c)^T R \bar{d}^c.$$

Using that each component of the action, $(\Theta((\bar{d}_n)) \star t_l)^c = t_l^c + D_{t_l} \bar{d}^c$, the problem results in the

quadratic problem

$$\begin{aligned} \underset{\bar{d}^1, \bar{d}^2, \dots, \bar{d}^d}{\text{minimize}} \quad & \begin{pmatrix} \bar{d}^1 \\ \vdots \\ \bar{d}^d \end{pmatrix}^T \begin{pmatrix} D_{(t)} & & \\ & \ddots & \\ & & D_{(t)} \end{pmatrix}^T \begin{pmatrix} D_{(t)} & & \\ & \ddots & \\ & & D_{(t)} \end{pmatrix} \begin{pmatrix} \bar{d}^1 \\ \vdots \\ \bar{d}^d \end{pmatrix} \\ & + \begin{pmatrix} \bar{d}^1 \\ \vdots \\ \bar{d}^d \end{pmatrix}^T \begin{pmatrix} \sigma^2 R & & \\ & \ddots & \\ & & \sigma^2 R \end{pmatrix} \begin{pmatrix} \bar{d}^1 \\ \vdots \\ \bar{d}^d \end{pmatrix} + 2 \begin{pmatrix} t_1^1 - x_1^1 \\ \vdots \\ t_N^1 - x_N^1 \\ t_1^2 - x_1^2 \\ \vdots \\ t_N^d - x_N^d \end{pmatrix}^T \begin{pmatrix} D_{(t)} & & \\ & \ddots & \\ & & D_{(t)} \end{pmatrix} \begin{pmatrix} \bar{d}^1 \\ \vdots \\ \bar{d}^d \end{pmatrix}, \end{aligned}$$

which can be solved efficiently by most scientific numerical packages.

3.5.3 Inexact matching for registering scalar functions

It is possible to use the inexact matching formulation to solve the registration of scalar functions. The generative model takes the same form $\mathbf{X}_{(y)} = (\mathbf{h} \star \mathbf{T})_{(y)} + \varepsilon_y$, where \mathbf{T} is a template image and ε_y is a noise in the intensity measured at the point y . Again, if it is assumed that noise distributes identically and independently, along all $y \in \mathbb{A}$, as a Gaussian $\mathcal{N}(0; \sigma)$, the maximum a posteriori estimation of \mathbf{h} is obtained by solving

$$\underset{\mathbf{h}}{\text{minimize}} \quad \text{reg}(\mathbf{h}) + \frac{1}{\sigma^2} \int_{y \in \mathbb{A}} (\mathbf{X}_{(y)} - (\mathbf{h} \star \mathbf{T})_{(y)})^2,$$

where the SSD dissimilarity function between images is recovered [Roche et al. 2000].

If we consider a spatial transformation model given by the displacements (\bar{d}_n) of some fixed knotpoints and regularized with an operator L , the problem takes the form

$$\underset{(\bar{d}_n)}{\text{minimize}} \quad \int_{y \in \mathbb{A}} (\mathbf{X}_{(y)} - (\Theta((\bar{d}_n)) \star \mathbf{T})_{(y)})^2 + \sigma^2 \sum_{c=1}^d (\bar{d}^c)^T R \bar{d}^c.$$

However, at this time we cannot proceed as in Section 3.4.1 where $(\mathbf{h} \star \mathbf{T})_{(y)}$ is computed by $\mathbf{T}_{(\mathbf{h}^{-1} \star y)}$ because, \mathbf{h}^{-1} cannot be described by the same spatial transformations model. Even more, as we working in the small-deformation formulation, the displacement fields generated by $\Theta((\bar{d}_n))$ can be non-invertible mappings. However, as was ramarked in Section 3.4.1, the previous integral functional can be rewritten as

$$\int_{y \in \mathbb{A}} (\mathbf{X}_{(y)} - (\mathbf{h} \star \mathbf{T})_{(y)})^2 = \int_{y \in \mathbb{A}} (\mathbf{X}_{(y)} - \mathbf{T}_{(\mathbf{h}^{-1} \star y)})^2 = \int_{z \in (\mathbf{h}^{-1} \star \mathbb{A})} \det(\mathbf{D}\mathbf{h})|_z (\mathbf{X}_{(\mathbf{h} \star z)} - \mathbf{T}_{(z)})^2,$$

where the variable substitution $z = \mathbf{h}^{-1} \star y$ was performed and the Jacobian of the transformation $\det(\mathbf{D}\mathbf{h})$ must be included. The evaluation of this functional does not require the computation of the inverse spatial transformation $(\Theta((\bar{d}_n)))^{-1}$ but requires the computation of the Jacobian which follows from Eq. (3.5.4)

$$\mathbf{D}_y(\Theta((\bar{d}_n)) \star y)^c|_{(y=z)} = \sum_k \alpha_k^c (\mathbf{D}\mathbf{g}_k)|_z + \sum_{j=1}^N \beta_j^c (\mathbf{D}G_{p_j})|_z,$$

and, as it was done in Eq. (3.5.8), for each z the Jacobian matrix can be written as a linear operation of the knotpoints displacements

$$\mathbf{D}_y(\Theta((\bar{d}_n)) \star y)^c|_{(y=z)} = J_z \bar{d}^c.$$

Finally, the registration problem for scalar functions is formulated in the inexact matching framework as

$$\underset{(\bar{d}_n)}{\text{minimize}} \quad \int_{\mathbf{z} \in \mathbb{A}} \det(J_{\mathbf{z}} \bar{d}) \left(\mathbf{X}_{(D_{\mathbf{z}} \bar{d})} - \mathbf{T}_{(\mathbf{z})} \right)^2 + \sigma^2 \sum_{c=1}^d (\bar{d}^c)^T R \bar{d}^c.$$

For the case of registration of scalar functions, the knotpoints are commonly selected as a regular array of points covering the image extent. The number of knotpoints must be chosen according the desired level of deformation in the images which typically depends on the frequencies present in the images.

3.6 Registration using non-parametric spatial transformation models

In this section let us consider the registration of landmark objects with spatial transformations parameterized by stationary velocity fields (SVF). Inherently this spatial transformation model provides invertible and differentiable deformation mappings. These are parameterized by stationary velocity fields described in a grid or non-parametrically (in the sense that is described with very high spatial resolution as to be considered dense in the continuous space) by specifying a velocity for each voxel. Even though the description of the velocity is non-parametric, interpolation and extrapolation rules must be specified because the construction of the deformation mapping depends of the continuous evolution of the mapping by integrating the trajectory of the points.

Let us start considering landmark objects $\mathbf{F} = \{f_1 \dots f_N\}$ be the fixed object and \mathbf{M} be the moveable object. The registration problem is formulated as an inexact matching by

$$\underset{\mathbf{v}}{\text{argmin}} \sum_{l=1}^N \|\varphi_{(1)}(m_l) - f_l\|^2 + \sigma^2 \int_{\mathbb{A}} \mathbf{v}_{(y)} (L\mathbf{v})_{(y)}, \quad (3.6.1)$$

where $\mathbf{v}: \mathbb{A} \rightarrow \mathbb{R}^d$ and the deformation mapping $\varphi_{(1)} = \exp(\mathbf{v}): \mathbb{A} \rightarrow \mathbb{A}$ is the solution at $t = 1$ of the IVP

$$\begin{cases} \dot{\varphi}_{(t)} = \mathbf{v} \circ \varphi \\ \varphi_{(0)} = \text{id}_{\mathbb{A}}. \end{cases}$$

An operator L is chosen as regularization of the velocity field and it will be considered that the smoothness of the deformation mapping can be indirectly controlled by the regularity of the velocity field.

In order to perform the optimization (3.6.1) a descent direction of the functional is needed. From the first term of the functional, it is obtained

$$\mathcal{D}_{\mathbf{v}} \left(\sum_{l=1}^N \|\varphi_{(1)}(m_l) - f_l\|^2 \right) = 2 \sum_{l=1}^N \overline{(\varphi_{(1)}(m_l) - f_l)^T} (\mathcal{D}_{\mathbf{v}} \varphi_{(1)}(m_l)). \quad (3.6.2)$$

Equation (3.6.2) is a projection of the image force $2 \overline{(\varphi_{(1)}(m) - f)^T}$, localized at $\varphi_{(1)}(m_l)$, to the space of velocities \mathbf{v} .

Lets start by considering a small perturbation of the velocity field along the direction \mathbf{w} , $\mathbf{v} \rightarrow \mathbf{v} + \epsilon \mathbf{w}$ and computing the Gâteaux derivative along this perturbation

$$d_{\epsilon} \varphi_{(1)}^{\mathbf{v} + \epsilon \mathbf{w}} \Big|_{(\epsilon=0)} = \lim_{\epsilon \rightarrow 0} \frac{\varphi_{(1)}^{\mathbf{v} + \epsilon \mathbf{w}} - \varphi_{(1)}^{\mathbf{v}}}{\epsilon}, \quad (3.6.3)$$

where the velocity guiding the deformation is remarked in the notation $\varphi_{(t)}^{\mathbf{v}}$ or $\varphi_{(t)}^{\mathbf{v}+\epsilon\mathbf{w}}$.

Following from the evolution equation for $\varphi_{(t)}$, at any t holds

$$d_t \varphi^{\mathbf{v}+\epsilon\mathbf{w}} = (\mathbf{v} + \epsilon\mathbf{w}) \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}} = \mathbf{v} \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}} + \epsilon\mathbf{w} \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}} + O(\epsilon)$$

and taking the derivative with respect to ϵ

$$\begin{aligned} d_\epsilon d_t \varphi^{\mathbf{v}+\epsilon\mathbf{w}} &= d_t d_\epsilon \varphi^{\mathbf{v}+\epsilon\mathbf{w}} \\ &= d_\epsilon (\mathbf{v} \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}}) + d_\epsilon (\epsilon\mathbf{w} \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}}) \\ &= ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}}) d_\epsilon (\varphi^{\mathbf{v}+\epsilon\mathbf{w}}) + \mathbf{w} \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}} + \epsilon ((\mathcal{D}\mathbf{w}) \circ \varphi^{\mathbf{v}+\epsilon\mathbf{w}}) d_\epsilon (\varphi^{\mathbf{v}+\epsilon\mathbf{w}}). \end{aligned}$$

Particularizing for $\epsilon = 0$,

$$d_t (d_\epsilon \varphi^{\mathbf{v}}) = ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) (d_\epsilon \varphi^{\mathbf{v}}) + \mathbf{w} \circ \varphi^{\mathbf{v}},$$

and by substituting $\psi = d_\epsilon \varphi^{\mathbf{v}}$, the last expression takes the form of a non-homogeneous differential equation

$$\dot{\psi} = ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) \psi + \mathbf{w} \circ \varphi^{\mathbf{v}}. \quad (3.6.4)$$

In order to obtain the solution of the non-homogeneous Eq. (3.6.4), first note that by computing the Jacobian of the evolution equation for $\varphi_{(t)}$, it is obtained

$$\mathcal{D}\dot{\varphi} = (\mathcal{D}\varphi)^\cdot = ((\mathcal{D}\mathbf{v}) \circ \varphi) \mathcal{D}\dot{\varphi}$$

which denoting $J = \mathcal{D}\varphi$ takes the form $\dot{J} = ((\mathcal{D}\mathbf{v}) \circ \varphi) J$ and can be identified with the homogeneous part of Eq. (3.6.4) (see also Section 2.5.4).

To obtain a particular solution of Eq. (3.6.4) we can use the technique of *variation of parameters* which express $\psi_{(t)}$ as a linear combination of the form $\varsigma_{(t)}^i c_{(t)}$, where $\{\varsigma_{(t)}^i\}_i$ is a basis of solutions of the homogeneous part, and the problem reduces to obtain $c_{(t)}$. It can be noted that the deformation mapping $\varphi_{(t)}$ is invertible and differentiable for all t . Thus, the Jacobian matrix $J_{(t)} = \mathcal{D}\varphi|_{(t)}$ will be non-singular and therefore, their columns provide a basis of solutions of the homogeneous part of Eq. (3.6.4). Therefore

$$\begin{aligned} \dot{\psi} &= ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) \dot{\psi} + \mathbf{w} \circ \varphi^{\mathbf{v}} \\ (Jc)^\cdot &= \dot{J}c + J\dot{c} = ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) Jc + \mathbf{w} \circ \varphi^{\mathbf{v}} \\ ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) Jc + J\dot{c} &= ((\mathcal{D}\mathbf{v}) \circ \varphi^{\mathbf{v}}) Jc + \mathbf{w} \circ \varphi^{\mathbf{v}} \\ J\dot{c} &= \mathbf{w} \circ \varphi^{\mathbf{v}} \Rightarrow \dot{c} = J^{-1}(\mathbf{w} \circ \varphi^{\mathbf{v}}). \end{aligned}$$

The last expression has solution

$$c_{(t)} = c_{(0)} + \int_{s=0}^t J_{(s)}^{-1} \mathbf{w}(\varphi_{(s)}^{\mathbf{v}}) ds.$$

Putting it all together and particularizing Eq. (3.6.3) for a point y , it is obtained

$$d_\epsilon \varphi_{(1)}^{\mathbf{v}+\epsilon\mathbf{w}}(y) \Big|_{(\epsilon=0)} = J_{(1)}(y) \int_0^1 J_{(s)}^{-1}(y) \mathbf{w}(\varphi_{(s)}^{\mathbf{v}})(y) ds.$$

Practical aspects

Let us summarize how $\xi_{(y)}^{\mathbf{w}} = d_\epsilon \varphi_{(1)}^{\mathbf{v} + \epsilon \mathbf{w}}(y) \Big|_{(\epsilon=0)}$ can be efficiently computed. Given \mathbf{v} and \mathbf{w} as velocity fields from the whole ambient space, $\mathbf{v}, \mathbf{w}: \mathbb{A} \rightarrow \mathbb{R}^d$, and a point y where to evaluate $\xi^{\mathbf{w}}$.

1. Choose a small time-step Δ , and initialize the variables: $t = 0$, $\mathbf{x} = \mathbf{y}$, $\xi_{(y)}^{\mathbf{w}} = \mathbf{0}_d$, $J = I_d$
2. Evaluate the velocity fields at \mathbf{x} : $V = \mathbf{v}(\mathbf{x})$, $W = \mathbf{w}(\mathbf{x})$
3. Evaluate the $d \times d$ Jacobian matrix of the velocity field at \mathbf{x} : $D = (\mathcal{D}\mathbf{v})_{(\mathbf{x})}$.
4. Update: $t = t + \Delta$
5. Update: $\mathbf{x} = \mathbf{x} + \Delta V$
6. Update: $\xi_{(y)}^{\mathbf{w}} = \xi_{(y)}^{\mathbf{w}} + \Delta J^{-1} W$
7. Perform multiplicative update of the Jacobian matrix: $J = e^{\Delta D} J$ (see Section 2.5.4).
8. If $t < 1$, go to step 2. Else return $J \xi_{(y)}^{\mathbf{w}}$

Projecting spatial forces on velocities

The most used procedure to infer changes in the velocity field from changes in the deformation mappings is by the use of the *Baker–Campbell–Hausdorff* (BCH) formula [Bossa et al. 2007] [Vercauteren et al. 2008] [Dru et al. 2010]. Given two velocity fields \mathbf{v} and \mathbf{w} , BCH formula proposes an expression for the velocity \mathbf{z} such that

$$\exp(\mathbf{z}) = \exp(\mathbf{v}) \bullet \exp(\mathbf{w}).$$

Actually, the formula is well posed for linear operators and matrix exponential function, with well known convergence properties and efficient algorithms to compute it and relates [Casas and Murua 2009] [Bonfiglioli and Fulci 2011] [Casas et al. 2012].

In the of matrices, the formula make use of the *matrix commutation* operation $[A, B] = AB - BA$. For the case of velocity fields, it must be replaced by the *Lie bracket* operation

$$[\mathbf{v}, \mathbf{w}]_{(y)} = (\mathcal{D}\mathbf{v})_{(y)} \mathbf{w}_{(y)} - (\mathcal{D}\mathbf{w})_{(y)} \mathbf{v}_{(y)},$$

where velocity fields must be defined in the whole domain and accurate high order derivatives are required to obtain reliable results. Typically, low-pass filters are applied to the velocities in order to compute finite difference schemes for the derivatives. Moreover, this filtering could help to define extend sparse defined velocities to the whole domain (such as in the case of localized image forces). However, a formal and rigourous definition for the BCH formula for the case of diffeomorphisms requires some hard conditions on the space of the velocities [Kriegl and Michor 1997] [Glöckner 2006] and the case of stationary velocity fields is not completely understood.

In this section an alternative strategy to the projection of the perturbations from the deformation to the velocities is proposed. The proposed strategy is restricted to the space of continuous linear velocities with finite dimensional basis. Lets return to the Eq. (3.6.2), where the derivatives $\mathcal{D}_{\mathbf{v}} \varphi_{(1)}(\mathbf{m})$ is needed to project the forces to the space of velocities. To compute

this projection, let us consider that \mathbf{v} is a linear combination of basis functions $\mathbf{b}_i: \mathbb{A} \rightarrow \mathbb{R}$, such that

$$\mathbf{v} = \sum_i \begin{pmatrix} a_i^1 \mathbf{b}_i \\ a_i^2 \mathbf{b}_i \\ \vdots \\ a_i^d \mathbf{b}_i \end{pmatrix} = \sum_i \sum_{c=1}^d a_i^c \mathbf{e}_c \mathbf{b}_i,$$

where a_i^c are the coefficients for each component of the velocity and \mathbf{e}_c are the canonical bases of \mathbb{R}^d (in 2D: $\mathbf{e}_1 = (1, 0)^T$ and $\mathbf{e}_2 = (0, 1)^T$). Then, the contribution of the force from the dissimilarity between $\varphi_{(1)}(\mathbf{m})$ and \mathbf{f} , to each basis \mathbf{b}_i at each coordinate c is given by

$$\begin{aligned} \mathcal{D}_{a_i^c} \|\varphi_{(1)}^{\mathbf{v}}(\mathbf{m}) - \mathbf{f}\|^2 &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left(\|\varphi_{(1)}^{\mathbf{v} + \epsilon \mathbf{e}_c \mathbf{b}_i}(\mathbf{m}) - \mathbf{f}\|^2 - \|\varphi_{(1)}^{\mathbf{v}}(\mathbf{m}) - \mathbf{f}\|^2 \right) \\ &= 2 \overline{(\varphi_{(1)}^{\mathbf{v}}(\mathbf{m}) - \mathbf{f})}^T \left(\mathbf{d}_\epsilon \varphi_{(1)}^{\mathbf{v} + \epsilon \mathbf{e}_c \mathbf{b}_i}(\mathbf{m}) \right) \\ &= 2 \overline{(\varphi_{(1)}^{\mathbf{v}}(\mathbf{m}) - \mathbf{f})}^T \xi_{(\mathbf{y})}^{\mathbf{e}_c \mathbf{b}_i}. \end{aligned} \quad (3.6.5)$$

If the functions \mathbf{b}_i have compact support, such as the case of multi-linear or b-splines basis, the image force $2 \overline{(\varphi_{(1)}^{\mathbf{v}}(\mathbf{m}) - \mathbf{f})}^T$ contributes to the coefficient a_i^c only if the trajectory $\varphi_{(t)}^{\mathbf{v}}(\mathbf{m})$ intersect with the support of \mathbf{b}_i .

Figure 3.6.1 shows a velocity field guiding three landmarks, the image forces and the projection of the forces to the space of velocities. For each landmark, the derivative with respect to a_i^c is computed. If more than one landmark trajectory influence in the same a_i^c , their contributions are added.

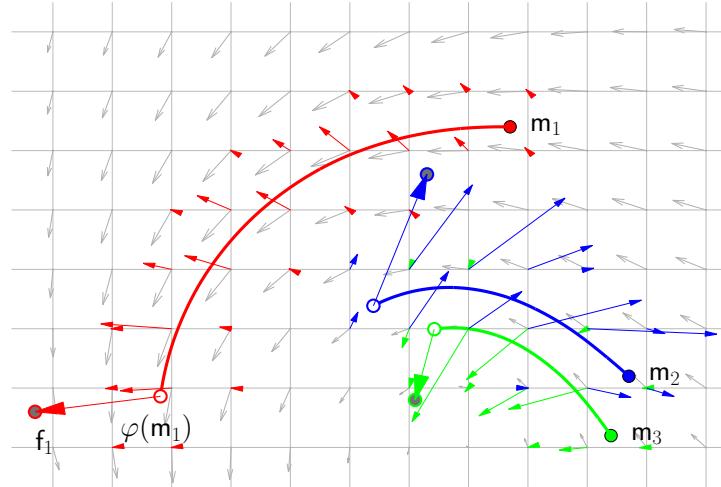


Figure 3.6.1. Evolution of landmarks in a stationary velocity field and the projection of image forces on the velocities. Each point evolves along each curve. The image force is computed between its final position and the desired position. Finally, the forces are projected on the space of velocities.

Example

An experiment was performed to show the results of the SVF registration process. Landmarks in 2D were registered using a regularization of the form

$$\begin{aligned}\text{reg}(\boldsymbol{\kappa}^v) &= \int_{\mathbb{A}} \mathbf{v}_{(y)}^T (L\mathbf{v})_{(y)} \\ &= \sum_{c=1}^d \int_{\mathbb{A}} v_{(y)}^c (Lv^c)_{(y)} \\ &= \sum_{c=1}^d \int_{\mathbb{A}} v_{(y)}^c (v^c - \alpha \nabla^2 v^c)_{(y)}.\end{aligned}$$

The objects to register are shown in the top row of Figure 3.5.1. Registration processes were perform by optimizing

$$\mathcal{R}_{SVF}(\mathbf{M}, \mathbf{F}) = \underset{\boldsymbol{\kappa}^v}{\operatorname{argmin}} \left(\sum_l \|f_l - \boldsymbol{\kappa}^v \star m_l\|^2 + \sigma^2 \int_{\mathbb{A}} \mathbf{v}_{(y)}^T (L\mathbf{v})_{(y)} \right), \quad (3.6.6)$$

where σ^2 is a parameter to control the trade-off between *dissimilarity* and *regularization*. The parameter α provides a measure of the spatial scales. The process to optimize Eq. (3.6.6) is by a descent procedure along the negated derivative or along the conjugate gradient direction. At each step, a backtracking line-search was performed. The velocities are parameterized in a 450×600 grid with cell with multi-linear interpolation and circular boundary condition. The gradient are computed by the projection of the forces on the 540 000 basis functions.

Several registration were performed at different values of σ^2 and α and are shown in Figure 3.6.2. It can be seen that the larger the value of σ^2 , the larger the weight of the regularization term, and the lower the matching of the points. On the other hand, the larger the larger of α , the larger the spatial scale.

The same procedure can be performed to register scalar images

$$\mathcal{R}_{SVF}(\mathbf{M}, \mathbf{F}) = \underset{\boldsymbol{\kappa}^v}{\operatorname{argmin}} \left(\int_{\mathbb{A}} (\mathbf{F}_{(y)} - (\boldsymbol{\kappa}^v \star \mathbf{M})_{(y)})^2 + \sigma^2 \int_{\mathbb{A}} \mathbf{v}_{(y)}^T (L\mathbf{v})_{(y)} \right). \quad (3.6.7)$$

In the case of SVF spatial transformation models, $(\boldsymbol{\kappa}^v)^{-1} = (\Theta(\mathbf{v}))^{-1} = \Theta(-\mathbf{v}) = \boldsymbol{\kappa}^{-v}$, and the process explained in Section 3.4.1 of defining $\bar{\Theta}(\bar{q}) = \text{inv}(\Theta(\bar{p}))$ can be used. Therefore, the problem (3.6.7) can be posed as

$$\mathcal{R}_{SVF}(\mathbf{M}, \mathbf{F}) = \left(\underset{\boldsymbol{\kappa}^{-v}}{\operatorname{argmin}} \left(\int_{\mathbb{A}} (\mathbf{F}_{(y)} - (\mathbf{M})_{(\boldsymbol{\kappa}^{-v} \star y)})^2 + \sigma^2 \int_{\mathbb{A}} \mathbf{v}_{(y)}^T (L\mathbf{v})_{(y)} \right) \right)^{-1}.$$

The descend direction is computed by the projection on the space velocity of the images forces

$$2((\mathbf{M})_{(\boldsymbol{\kappa}^{-v} \star y)} - \mathbf{F}_{(y)}) (\nabla \mathbf{M})_{(\boldsymbol{\kappa}^{-v} \star y)}$$

by computing $\xi_{(y)}^{e_i b_i}$, for every point y used as query point in the matching energy.

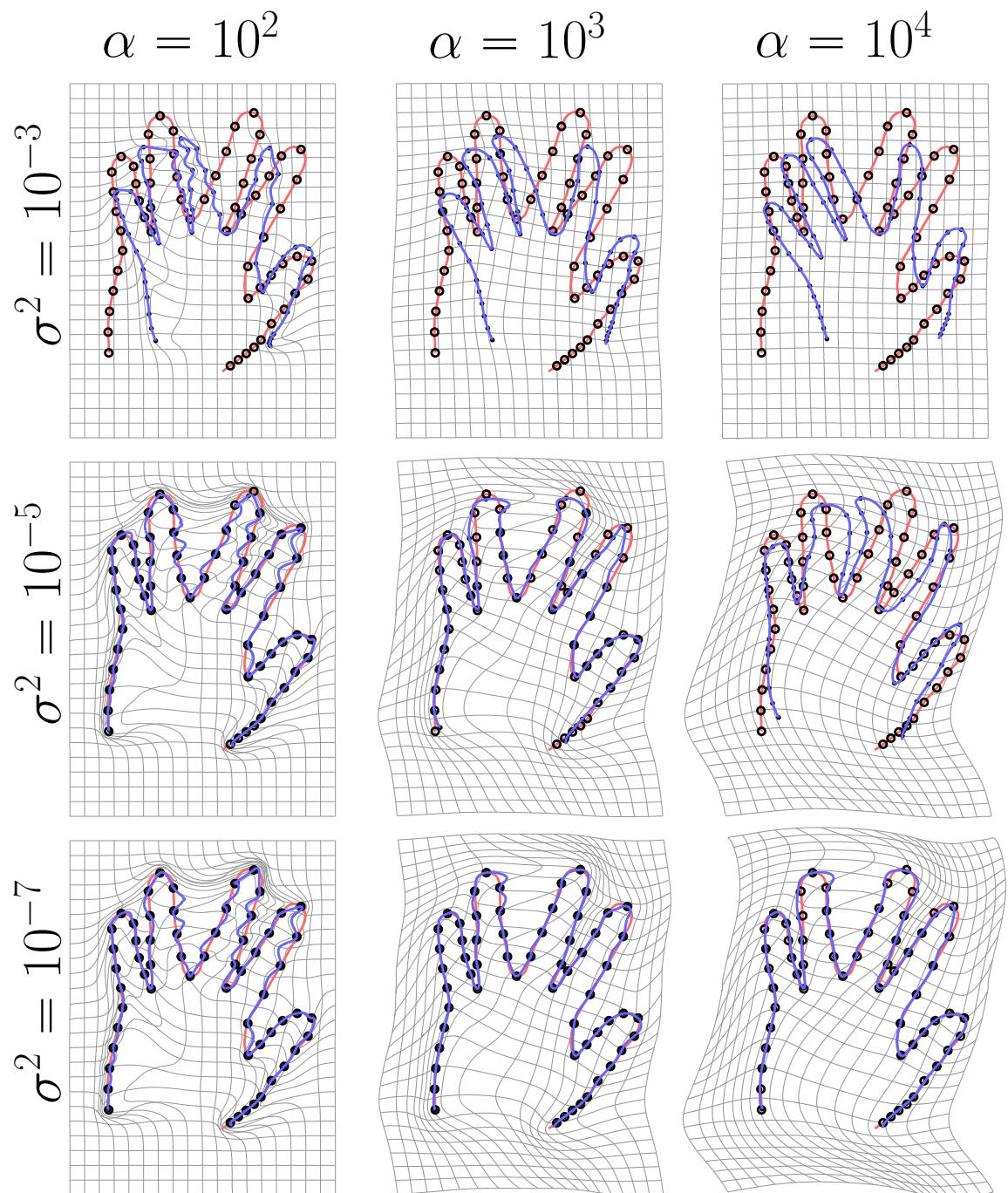


Figure 3.6.2. Landmark registration with SVF. Different values of the regularization kernel size α and the of the trade-off factor σ^2 .

CHAPTER

4

Left-invariant geodesics

4.1	Introduction	76
4.2	Background	78
4.2.1	Charts and manifolds, Lie groups, matrix groups	78
4.2.2	Tangent vectors, tangent space and vector fields	80
4.2.3	Translating elements, curves and velocities	80
4.2.4	Conjugation, Adjoint and adjoint	81
4.2.5	Group exponential	82
4.2.6	Riemannian metrics, lengths, geodesics and $\text{Exp}(\cdot)$	83
4.2.7	Isometries	85
4.2.8	Invariant Riemannian metrics	86
4.2.9	Totally geodesic subgroups	87
4.2.10	Bi-invariant metrics	87
4.2.11	Geodesics as “straightest” curves	88
4.3	Computing left-invariant Riemannian geodesics	92
4.3.1	Computing geodesic in coordinate charts	93
4.3.2	Solving geodesics in the group exponential chart	95
4.3.3	Evolution of geodesics described in the algebra	97
4.3.4	A coordinate-free solution based on (E–P)	100
4.3.5	Geodesics using Optimal Control	103
4.3.6	Symmetric representation of the geodesics evolution	105
4.4	Conserved quantities along geodesics	105
4.5	Reducing to a first-order equation	106
4.6	Sensitivity with respect to initial velocity	108
4.7	$\text{Exp}(\cdot)$ under different metrics	109
4.8	Expressions for the right-invariant metric case	110

In the computational anatomy framework, the study of the variability among anatomies is performed by the study of the variability of the spatial transformation relating the anatomical structures. The set of all spatial transformations is the homeomorphisms $\mathcal{H}_{\mathbb{A}}$ of the ambient space and it is common to select a subset of $\mathcal{H}_{\mathbb{A}}$ as the allowed transformations for the registration of the anatomies. When this subset remains invariant interesting geometrical properties of the geometrical objects, then it is very common that the set has a group structure under the composition and inversion operations inherited from $\mathcal{H}_{\mathbb{A}}$. In the following the selected set of spatial transformations will be denoted as \mathcal{G} . Unfortunately, it is not possible in general to endow \mathcal{G} with a vector space structure which allows the use of classical statistical tools for the analysis of spatial transformations. However, an intrinsic statistical framework can be used in the cases where \mathcal{G} has a metric space structure. Therefore, we have to be able measure distances between any two spatial transformation from \mathcal{G} in order to perform appropriate statistical analysis.

The present chapter covers the road to assign a metric structure on a group. In particular, on Lie groups which can be represented by matrix groups. Although, the representability will allow us to express all the operations in the well known language of matrix operations, some footnotes will be included to generalize the concepts to non-representable groups, such as some quotient groups or infinite dimensional groups. The mathematical nature of Lie groups is beyond the classical arithmetic, algebra or calculus. It relies on the areas of *group theory*, *differentiable manifolds*, *differential geometry*. The main goal of the chapter is to provide a bridge for researchers from computer vision and image analysis fields to fill the gap between differential geometry and technical disciplines. This bridge can also be helpful in the recent and growing interest in the study of other matrix manifolds used in image analysis applications without a group structure, such as symmetric and positive definite tensors, Grassmann manifolds, or Stiefel manifolds [Lui 2012]. Practical aspects and some illustrative applications are relegated to the next chapter.

4.1 Introduction

Groups of spatial transformations are key elements in the fields of computer graphics, computer vision, image analysis and medical imaging. Among the most interesting groups are those that can be continuously parameterized with a finite number of parameters. Most of them can be represented by matrix groups (see Section 2.5.2) and therefore are Lie groups [Tapp 2005], which means that are also differentiable manifolds. As it was done in several works along the last years, an appropriate framework to work is by using differential and Riemannian geometry concepts on those manifolds. Then, a suitable distance can be defined from a Riemannian metric assigned to the manifold that the group elements form. Additionally, it is common that computer vision and image analysis problems present a symmetry related to a set of spatial transformations. A wise selection of the metric should take into account these symmetries leading to left-invariant, right-invariant or bi-invariant metrics.

Many applications have been proposed where the choice of a metric is crucial. Within the field of statistics on manifold data, applications include mean and variance estimation [Pennec and Ayache 1998] [Gramkow 2001] [Moakher 2002] [Arsigny 2006] [Pennec 2006a] [Boisvert et al. 2008], regression and curve fitting [Machado et al. 2010] [Fletcher 2013] [Hinkle et al. 2013], dimensionality reduction [Said et al. 2007] [Sommer et al. 2013], hypothesis testing for group comparison [Zacur et al. 2013], classification [Gorczowski et al. 2010] [Bossa et al. 2011] and statistical sampling [Kuffner 2004] [Makadia et al. 2006]. Interpolation of spatial transformations [Shoemake 1985] [Park and Ravani 1997] [Kang and Park 1999] [Alexa 2002] [Begelfor and Werman 2005] [Li and Hao 2006] [Boumal 2013] and filtering of manifold data [Rahman et al. 2005] [Fillard et al. 2005] [Lenglet et al. 2006] [Fletcher and Joshi 2007] [Gur and Sochen 2009] [Rosman et al. 2012] [Jia and Evans 2013] also make use of a distance

function. Intrinsic differential calculus on groups, which is essential in geometric optimization, is defined for a given metric [Gabay 1982] [Smith 1994] [Absil et al. 2007] [Absil et al. 2009]. Moreover, the choice of a metric defines spatial transformation neighborhoods that are commonly used in global optimization techniques [Olsson et al. 2006] [Li and Hartley 2007] [Hartley and Kahl 2009], as well as in derivative free optimization methods [Gwak et al. 2003] [Lee et al. 2007] [Chang et al. 2011]. A distance between spatial transformations can also be used to evaluate the performance of image registration algorithms when a ground truth transformation is available [Penney et al. 1998] [Roche et al. 2001] [Skerl et al. 2006]. For free and transitive group actions, distance between transformations can be used to define a distance between objects. For example: in robotics, rigid transformations were used to define distances between configurations [Žefran et al. 1999]; in medical imaging, transformations between subcortical structures within the brain [Gorcowski et al. 2010] [Bossa et al. 2011] and spinal vertebrae [Adler et al. 2002] [Boisvert et al. 2008] were analyzed. Furthermore, a distance between deformations on the plane was used to measure distances between 2D curves [Younes 1998]. This distance between spatial deformations was used as a regularizer in ill-posed optimization problems [Beg et al. 2005]. Other applications are in the area of computer vision, such as object tracking [Tuzel et al. 2008] [Kwon and Park 2010], visual servoing of robot manipulators [Drummond and Cipolla 2000] [Bayro-Corrochano and Ortegón-Aguilar 2007], and dynamical analysis of visual flows [Lin et al. 2009]. The particular cases of rotations and rigid transformations have been deeply analyzed within the fields of robotics [Murray et al. 1994] [Park 1995] [Park et al. 1995] [Žefran et al. 1998] [Žefran and Kumar 1998] [Noakes and Popiel 2007] and geometric mechanics [Arnold 1989] [Park and Kim 2000] [Holm et al. 2009] where the use of left-invariant metrics on groups are well known because they guide the trajectories of a moving object without external forces. On the other hand, right-invariant metrics have been shown to guide the dynamics of inviscid flows [Arnold and Khesin 1998] [Constantin and Kolev 2002] [Khesin 2005] [Khesin 2008].

When a Riemannian metric is assigned to the manifold of spatial transformations the length of a curve segment can be measured. Those curves on the manifold which are locally length-minimizing are called Riemannian geodesics and play an important role in many applications of computer vision and image analysis. The concept of geodesics is often found in computer graphics and image analysis disciplines for other applications, such as implicit or parametric surfaces [Kimmel et al. 1995] [Dedieu and Nowicki 2005] [Seong et al. 2008], triangle meshes [Sifri et al. 2003] [Surazhsky et al. 2005] [Peyré and Cohen 2006] [Bommes and Kobbelt 2007] [Balasubramanian et al. 2009] images [Caselles et al. 1997] [Sethian 1999] [Boykov and Kolmogorov 2003] [Criminisi et al. 2008], but those will not be considered in this thesis.

In Riemannian framework, the length of the shortest curve connecting two transformations defines the Riemannian distance and endows a metric structure to the set of spatial transformations [Do Carmo 1992] [Gallot et al. 2004]. The shortest curve segment between two given spatial transformations belongs to the set of Riemannian geodesics. Geodesic curves or segments can be unambiguously defined in terms of a point of the manifold and a velocity in the tangent space of this point. The resulting geodesic emanates from the given point with the given initial velocity. This identification define a mapping from a tangent space of the manifold to the manifold itself which is called the Riemannian exponential function. On the other hand, the group exponential function also maps vectors from the tangent space at the identity to the group. It is well known that for matrix groups this mapping is given by the matrix exponential. Although in a few particular groups the Riemannian exponential function for some invariant metrics and translated group exponentials provide the same curve [Milnor 1976] [Do Carmo 1992] [Gallier 2013], the Riemannian exponential and the group exponential functions are totally different concepts. Unluckily, it should be noted that these concepts have not always been used with suffi-

cient accuracy in some publications. In particular group exponential curves were interpreted as shortest curves in [Tuzel et al. 2005] [Rahman et al. 2005] [Bossa and Olmos 2006] [Porikli et al. 2006] [Govindu 2006] [Bayro-Corrochano and Ortegón-Aguilar 2007] [Tuzel et al. 2008] [Lin et al. 2009] [Subbarao and Meer 2009] [Kwon and Park 2010] [Li et al. 2012] [Kwon et al. 2014]

Outline of the chapter

A summary of several concepts from Lie group and Riemannian geometry theories is presented in Section 4.2 with the aim of providing a self-contained compendium written for computer graphic and imaging researchers. At the same time this section introduces the notation used throughout the chapter.

Section 4.3 gathers several approaches to compute left-invariant Riemannian geodesics on matrix groups. The left-invariance case was arbitrarily chosen but it is straightforward to obtain right-invariant geodesics as will be indicated in Section 4.2.8. Whenever possible, the resolution methods will be written using matrix operations allowing straightforward numerical implementations. The methods include the integration of ordinary differential equations (ODEs) described in different ways: on specific chart coordinate Section 4.3.1, on a chart coordinate valid for any matrix group Section 4.3.2, solving the Euler–Poincaré equation Section 4.3.3, a coordinate-free formulation of Euler–Poincaré equation Section 4.3.4, a coupled state-costate equation system derived from Optimal Control Section 4.3.5 and a symmetric representation of geodesics Section 4.3.6.

Section 4.4 presents the derivation of two well-known conserved quantities along left-invariant geodesics, using results from Section 4.3.5. These quantities allow formulation in Section 4.5 of the geodesic evolution as an equivalent initial value problem (IVP) expressed only in terms of the transformation elements. Finally, the derivative of the Riemannian exponential function with respect to the initial velocity is given in Section 4.6. This is known as *Jacobi field* and it is used in practical applications to compute the inverse of Riemannian exponential function. Section 4.7 shows how to find geodesics in a broad family of Riemannian metrics using a known geodesic defined for a particular metric.

Finally, Section 4.8 gathers the right-invariant versions of the most important expressions considered along the chapter.

4.2 Background

This section gathers some concepts and properties involved in the definition of geodesics on transformation groups. In particular Riemannian metrics and the Riemannian exponential function are briefly explained. Also, the concept of group exponential is revisited in order to make clear the differences between both maps and avoid their misleading use. The knowledge of these concepts and properties is essential for dealing with the computation of geodesics and distances. Further details can be found in classical textbooks [Do Carmo 1992] [Hall 2003] [Gallot et al. 2004] [Tapp 2005] [Burns and Gidea 2005] [Absil et al. 2009].

4.2.1 Charts and manifolds, Lie groups, matrix groups

Almost all classical spatial transformation groups of finite degrees of freedom allow a matrix representation. Briefly, this means that any elements χ, γ from a group \mathcal{G} can be identified with matrices $X, Y \in \mathbf{G} \subset \mathbb{M}_n$ (for some n), such that the composition $\chi \bullet \gamma$ is identified with the matrix multiplication XY (see 2.4.1). For simplicity only matrix groups are considered in this chapter although some footnotes are added to generalize the concepts for abstract structures. As

it is shown in [Tapp 2005], all matrix groups are *Lie groups*, *i.e.* the elements form a *smooth manifold*¹.

The modern concept of a manifold relies on the concepts of charts and atlases [Scholz 1999]. Let \mathcal{G} be a non-empty set of spatial transformations, and let χ be an element of \mathcal{G} . A *local coordinate chart* (\mathcal{D}, φ) is a bijection that maps elements in $\mathcal{D} \subset \mathcal{G}$ to an open set $\Omega \subset \mathbb{R}^k$ for some k . The array $\varphi(\chi) = \bar{x} = [x^1, x^2, \dots, x^k]^T \in \Omega$ (hereinafter the overline means that the corresponding quantity is a column array while superscripts refer to its components) contains the coordinates of χ in the chart (\mathcal{D}, φ) .

An *atlas* for \mathcal{G} is a countable collection of charts $\{(\mathcal{D}_i, \varphi_i)\}_i$ such that the union of all \mathcal{D}_i covers the set \mathcal{G} . This construction allows definition of a *topology* on the set \mathcal{G} since open sets in \mathbb{R}^k (or, more precisely in each Ω_i) define *open sets in \mathcal{G}* via the functions φ_i^{-1} . If for any pair of charts $(\mathcal{D}_i, \varphi_i)$ and $(\mathcal{D}_j, \varphi_j)$ with non empty intersection $\mathcal{D}_i \cap \mathcal{D}_j$, the function $\varphi_i \circ \varphi_j^{-1}: \mathbb{R}^k \rightarrow \mathbb{R}^k$ is smooth then the set \mathcal{G} becomes a *smooth manifold* with *intrinsic dimension* k where continuity and differentiability properties in \mathcal{G} are inherited from the continuity and differentiability of \mathbb{R}^k . In particular, curves can be defined over the set \mathcal{G} by mapping curves in Ω_i to \mathcal{G} using the φ_i^{-1} function.

With the previous concepts a question on to recognize a set as a manifold arises. There is a nice discussion in [Absil et al. 2009] about how to check if a given set is a manifold although a more well-posed question is to ask whether a given set admits to be described and covered by an atlas. Then a manifold results in not simply a set \mathcal{G} but rather in a couple $(\mathcal{G}, \{(\mathcal{D}_i, \varphi_i)\}_i)$, *i.e.* in the set together with a given atlas.

Let \mathcal{G} be a smooth manifold with the atlas $\{(\mathcal{D}_i, \varphi_i)\}_i$ and let \mathbb{V} be a vector space, a function $\mathbf{f}: \mathcal{G} \rightarrow \mathbb{V}$ is smooth if all the compositions $\mathbf{f} \circ \varphi_i^{-1}: \Omega_i \subset \mathbb{R}^k \rightarrow \mathbb{V}$ are smooth functions. Similarly, a function $\mathbf{g}: \mathbb{V} \rightarrow \mathcal{G}$ is smooth if all $\varphi_i \circ \mathbf{g}: \mathbb{V} \rightarrow \Omega_i$ are smooth.

A *Lie group* of spatial transformations is a group of spatial transformations \mathcal{G} which can be endowed with a smooth differentiable manifold structure. Additionally, the composition $(\cdot \bullet \cdot): \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$ must be a smooth map with respect to the topology of the manifold. As a consequence, the inversion operation also results in a smooth function within the topology of the manifold [Kolar et al. 1993]. The *intrinsic dimension* of the Lie group is the dimension of the images of the selected charts k .

In the case of a matrix group $\mathbf{G} \subset \mathbb{M}_n$, the set of matrices \mathbb{M}_n is isomorphic to \mathbb{R}^{n^2} and according to the strong Whitney's embedding Theorem, the manifold \mathbf{G} can be intuitively thought as a smooth “surface” (without boundaries) of dimension k embedded in the Euclidean space \mathbb{R}^{2k} . This “surface” will have an *intrinsic dimension* $k \leq 2n^2$ being the dimension of the Lie group. Each transformation can be parameterized as $X(x^1, \dots, x^k)$ or $X(\bar{x})$ and $\bar{x}(X): \mathbf{D} \subset \mathbf{G} \rightarrow \mathbb{R}^k$ is a local coordinate chart. For some matrix groups the “surface” where the transformations live are disjoint “surfaces” where each one has intrinsic dimension k . For example, the General Linear group $\mathbf{GL}(d)$ is composed by the “surface” corresponding to the matrices with determinant greater than 0, $\mathbf{GL}^+(d)$, and the “surface” corresponding to the matrices with determinant smaller than 0.

As we are interested in building continuous paths between elements of a group, we will avoid transformation groups composed of isolated elements such as Inversion or Dihedral groups and only cases where $k \geq 1$ will be considered. Moreover, only elements belonging to the same *connected component* can be connected by continuous paths.

¹ As in previous chapters, *smooth* refers to functions with derivatives up to any order, *i.e.* C^∞ functions.

4.2.2 Tangent vectors, tangent space and vector fields

A *curve* in a matrix group \mathbf{G} is a smooth function $\gamma: \mathbb{R} \rightarrow \mathbf{G}: t \mapsto \gamma_{(t)}$. The set of all curves in \mathbf{G} will be denoted as $\mathcal{C}(\mathbf{G})$. The *velocity*, alternatively called *tangent vector*, of the curve is defined as the usual limit $\lim_{\epsilon \rightarrow 0} (\gamma_{(t+\epsilon)} - \gamma_{(t)})/\epsilon$ resulting in a matrix² $V_{(t)} = d_s \gamma_{(s)} \Big|_{(s=t)} = \dot{\gamma}_{(t)}$. Like group elements, the velocities $\dot{\gamma}_{(t)}$ can also be represented by matrices in \mathbb{M}_n but in general they do not belong to the set \mathbf{G} . The velocities of all curves passing through the point X at time $t = 0$ define the *tangent space* of \mathbf{G} at X and it is denoted by $T_X \mathbf{G} = \{\dot{\gamma}_{(0)} \mid \text{such that } \gamma \text{ is a curve in } \mathbf{G} \text{ with } \gamma_{(0)} = X\}$. Every tangent space $T_X \mathbf{G}$ is a vector space of dimension k .

A *vector field* \mathcal{H} on the manifold is an assignment from each $X \in \mathbf{G}$ to a velocity in $T_X \mathbf{G}$. \mathcal{H} acts on a smooth scalar function $\mathbf{f}: \mathbf{G} \rightarrow \mathbb{R}$ providing the directional derivative $(\mathcal{H}\mathbf{f})(X) = \mathcal{H}(X)\mathbf{f} = \lim_{\epsilon \rightarrow 0} (\mathbf{f}(X + \epsilon \mathcal{H}(X)) - \mathbf{f}(X))/\epsilon$ (whenever the limit can be evaluated³). A vector field \mathcal{H} is smooth if for any smooth scalar function \mathbf{f} , the scalar function $(\mathcal{H}\mathbf{f}): \mathbf{G} \rightarrow \mathbb{R}$ is smooth. The successive application of two smooth vector fields \mathcal{H} and \mathcal{S} returns a second order derivative which cannot be identified with a vector field. Nonetheless, the second order terms are canceled in the operation $(\mathcal{H} \circ \mathcal{S} - \mathcal{S} \circ \mathcal{H})\mathbf{f}|_X$. This is a bilinear operation and it is called *Jacobi–Lie bracket*, denoted as $[\mathcal{H}, \mathcal{S}]$. Smooth vector fields are closed under the Jacobi–Lie bracket.

4.2.3 Translating elements, curves and velocities

For every group element Y , *left-translation* is the map $\mathbf{L}_Y: \mathbf{G} \rightarrow \mathbf{G}$ defined by $\mathbf{L}_Y(X) = YX$ for all $X \in \mathbf{G}$. Analogously, *right-translation* is defined by $\mathbf{R}_Y(X) = XY$. Note that, both operations are related by

$$\mathbf{R}_Y(X) = XY = (Y^{-1}X^{-1})^{-1} \Rightarrow \mathbf{R}_Y = \mathbf{inv} \circ \mathbf{L}_{Y^{-1}} \circ \mathbf{inv}. \quad (4.2.1)$$

Let γ be a curve through X with velocity V at $t = 0$, then the curve ξ defined by $\xi_{(t)} = \mathbf{L}_Y(\gamma_{(t)})$ passes through YX at $t = 0$ ($\xi_{(0)} = YX$). Its velocity $W = \dot{\xi}_{(0)}$ belongs to $T_{YX} \mathbf{G}$. The corresponding action on tangent vectors of \mathbf{L}_Y is a linear operation denoted by $T_X \mathbf{L}_Y$ given by⁴ $W = d_s(Y\gamma_{(s)}) \Big|_{(s=0)} = T_X \mathbf{L}_Y V = YV$.

As velocities can be translated along the group, it is very useful and convenient to define the properties of velocities at a single point. Algebraically, the most convenient point is the identity

² On an abstract manifold \mathcal{G} without a vector space structure, the usual limit can only be evaluated in terms of coordinates. Alternatively tangent vectors can be defined as differential operators. Let \mathbf{f} be a smooth real valued function defined on \mathcal{G} , then $\mathbf{f} \circ \gamma: \mathbb{R} \rightarrow \mathbb{R}: t \mapsto \mathbf{f}(\gamma_{(t)})$ and the velocity $\dot{\gamma}_{(t)}$ is the linear operator satisfying (see for example [Absil et al. 2010] [Gallier 2013])

$$\dot{\gamma}_{(t)}\mathbf{f} := d_s \mathbf{f}(\gamma_{(s)}) \Big|_{(s=t)}.$$

³ In an abstract group \mathcal{G} where the argument of the first term of the limit cannot be evaluated, the directional derivative is defined as

$$\mathcal{H}(\chi)\mathbf{f} = \lim_{\epsilon \rightarrow 0} (\mathbf{f}(\gamma_{(\epsilon)}) - \mathbf{f}(\gamma_{(0)}))/\epsilon,$$

where γ is a curve in \mathcal{G} such that $\gamma_{(0)} = \chi$ and $\dot{\gamma}_{(0)} = \mathcal{H}(\chi)$.

⁴ For abstract structures, $\mathcal{V} \in T_x \mathcal{G}$ and $\mathcal{W} \in T_{yx} \mathcal{G}$ are related by a linear operator called *tangent-lift*, denoted by $T_x \mathbf{L}_y: T_x \mathcal{G} \rightarrow T_{yx} \mathcal{G}$, or in a more general form, for a function $\mathbf{F}: \mathcal{G} \rightarrow \mathcal{G}$, $T_x \mathbf{F}: T_x \mathcal{G} \rightarrow T_{\mathbf{F}(x)} \mathcal{G}$ (see for example [Holm et al. 2009]). There are different notations for this operator, for example: [Gallier 2013] denotes it as $(d\mathbf{F})_\chi$; [Arnold 1989] and [Arnold and Khesin 1998] use \mathbf{F}_* or $\mathbf{F}_*|_\chi$; [Arsigny 2006] uses $D_\chi \mathbf{F}$; [Cendra et al. 2003] [Kolev 2004] simply denotes \mathbf{F} leaving to the context its meaning; $\mathbf{DF}(\chi)$ were used in [Absil et al. 2009].

I. The tangent space at the identity $T_I \mathbf{G}$ is called *group algebra* and is denoted by \mathfrak{g} . Any velocity V in any tangent space $T_X \mathbf{G}$ can be associated with an element $U \in \mathfrak{g}$ via $V = XU$.

We may identify the whole tangent bundle $T\mathbf{G}$, *i.e.* the collection of all tangent spaces $T\mathbf{G} \equiv \{T_X \mathbf{G} \mid X \in \mathbf{G}\}$, with the set $\mathbf{G} \times \mathfrak{g}$ by means of the mapping that sends (X, U) to $T_I \mathbf{L}_X U \in T_X \mathbf{G}$. The inverse identification of $T\mathbf{G} \rightarrow \mathbf{G} \times \mathfrak{g}$ is called *left-trivialization*. It is defined by the mapping $V \mapsto (X, T_X \mathbf{L}_{X^{-1}} V)$ where $V \in T_X \mathbf{G}$ and the resulting $T_X \mathbf{L}_{X^{-1}} V \in \mathfrak{g}$ is usually called the *left-trivialized velocity*.

The algebra \mathfrak{g} is a k -dimensional vector space, therefore a set of bases $\{B_1, B_2, \dots, B_k\}$, with $B_i \in \mathbb{M}_n$, can be selected such that $\mathfrak{g} \equiv \text{span}(\{B_i\})$. In some disciplines these bases are referred to as *infinitesimal group generators* and they span all group perturbations that can be done at the identity. Matrices representing vectors of \mathfrak{g} are written as $U = \sum_i^k B_i \mu^i \in \mathfrak{g}$, where $\bar{\mu} = (\mu^1, \mu^2, \dots, \mu^k)^T \in \mathbb{R}^k$ are called the *algebra coordinates of U* . Let \bar{U} denote the n^2 dimensional array obtained by stacking the columns of the matrix U . It is practical to write $\bar{U} = \mathbf{B} \bar{\mu}$, where \mathbf{B} is an $n^2 \times k$ matrix built from the vectorization of each basis, $\mathbf{B} = (\bar{B}_1, \bar{B}_2, \dots, \bar{B}_k)$. Moreover, as \mathbf{B} is full column rank, the algebra coordinates can be obtained by $\bar{\mu} = \mathbf{B}^\dagger \bar{U}$, where $(\cdot)^\dagger$ denotes a left generalized inverse, *i.e.* any matrix which satisfies $\mathbf{B}^\dagger \mathbf{B} = I_k$ (for example the Moore–Penrose pseudo-inverse can be used).

The group algebra \mathfrak{g} has a bilinear operation called *Lie bracket*, $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$. By representing the elements of the algebra as matrices⁵, the Lie bracket operation results in the matrix commutation, $[U, W] = UW - WU$ for U and $W \in \mathfrak{g}$. Algebra vectors are closed under the Lie bracket and the algebra is fully characterized by the *structure constants* C_{ijm} [Sattinger and Weaver 1986] given by

$$[B_i, B_j] = \sum_{m=1}^k C_{ijm} B_m. \quad (4.2.2)$$

This closure property can be used to check if a set of matrices in $\{B_i \in \mathbb{M}_n\}_i$ is a valid basis of an algebra. Under the Lie bracket the bases of an algebra must satisfy

$$\text{rank}((\bar{B}_1, \bar{B}_2, \dots, \bar{B}_k)) = \text{rank}\left(\left(\bar{B}_1, \bar{B}_2, \dots, \bar{B}_k, [\bar{B}_i, \bar{B}_j]\right)\right) = k \quad , \text{ for all } i, j.$$

Left-translations of velocities from the group algebra to the whole group allows the generation of *left-invariant vector fields*⁶ $\tilde{u}(X) = XU$, for any $U \in \mathfrak{g}$ and any $X \in \mathbf{G}$. It is remarkable that any smooth vector field (not necessarily invariant) can be written as a point-wise linear combination of left-invariant vector fields, $\mathcal{H}(X) = \sum_i \mathbf{h}_i \tilde{u}_i(X)$ with \mathbf{h}_i smooth scalar functions on \mathbf{G} .

4.2.4 Conjugate action on groups elements, Adjoint and adjoint operations on algebra vectors

Conjugation of a group element X with another group element Y is defined as $(\mathbf{L}_Y \circ \mathbf{R}_{Y^{-1}})X = YXY^{-1}$. Its corresponding action on vectors of the group algebra is called the *Adjoint action* of

⁵ Although not all finite dimensional groups can be represented by a matrix group, the *Ado's Theorem* asserts that every finite dimensional Lie algebra can be viewed as an algebra of some square matrices under the matrix commutation [Hall 2003].

⁶ Equivalently, a left-invariant vector field is a vector field which satisfies $\tilde{\mathcal{H}}(yx) = T_x \mathbf{L}_y \tilde{\mathcal{H}}(x)$ for any $x, y \in \mathcal{G}$. By identifying a left-invariant vector fields with its left-trivialized velocity in the algebra, the group algebra can be alternatively defined as the set of all left-invariant vector fields and the Lie bracket results in the restriction of the Jacobi–Lie bracket to left-invariant vector fields: $[\tilde{u}, \tilde{w}] = \tilde{u} \circ \tilde{w} - \tilde{w} \circ \tilde{u}$.

\mathbf{G} on \mathfrak{g} given by⁷ $\mathbf{Ad}_Y: \mathfrak{g} \rightarrow \mathfrak{g}: U \mapsto YUY^{-1}$ for any $Y \in \mathbf{G}$. For a curve γ passing through I with velocity U at $t = 0$, the *conjugated* curve $\zeta_{(t)} = Y\gamma_{(t)}Y^{-1}$ passes through $\zeta_{(0)} = I$ with velocity $\dot{\zeta}_{(0)} = \mathbf{Ad}_Y U$. The Adjoint operation is a linear operation over \mathfrak{g} . By using the matrix identity $\overline{ABC} = (C^T \otimes A)\overline{B}$ for any matrices A, B and C , where $(\cdot \otimes \cdot)$ is the matrix Kronecker product, $\mathbf{Ad}_Y(\cdot)$ can be represented by a matrix of the form

$$\overline{\mathbf{Ad}_Y U} = (Y^{-T} \otimes Y)\overline{U}.$$

Now, let η be a curve in \mathbf{G} such that $\eta_{(0)} = I$ and $\dot{\eta}_{(0)} = W \in \mathfrak{g}$. For a given $U \in \mathfrak{g}$, the operation given by $d_s(\mathbf{Ad}_{\eta_{(s)}} U)|_{(s=0)}$ is called the *adjoint operator* of W on \mathfrak{g} . It is denoted by $\mathbf{ad}_W: \mathfrak{g} \rightarrow \mathfrak{g}$ and is a linear operation on \mathfrak{g} for any $W \in \mathfrak{g}$. Moreover, it can be proven that the adjoint operator coincides with the Lie bracket on the algebra vectors, *i.e.* $\mathbf{ad}_W(U) = [W, U]$ for any $U, W \in \mathfrak{g}$. In the same fashion than before, $\mathbf{ad}_W(\cdot)$ can be represented by a matrix of the form

$$\overline{\mathbf{ad}_W U} = ((I_n \otimes W) - (W^T \otimes I_n))\overline{U},$$

with $W, U \in \mathfrak{g}$.

4.2.5 Group exponential

Given a smooth vector field \mathcal{H} , a curve γ is called an *integral curve* of \mathcal{H} if $\dot{\gamma}_{(t)} = \mathcal{H}(\gamma_{(t)})$ for $t \in (-\epsilon, \epsilon) \subset \mathbb{R}$, a non-empty interval of \mathbb{R} . For any smooth vector field \mathcal{H} and any $X \in \mathbf{G}$ there is a unique integral curve γ of \mathcal{H} defined on $(-\epsilon, \epsilon)$ such that $\gamma_{(0)} = X$. The vector field \mathcal{H} is called *complete* if, for all initial points X , the integral curves can be extended to all times $t \in \mathbb{R}$.

For matrix groups and in the cases where the vector field is a left-invariant vector field \tilde{u} , integral curves are given by the solution of the following autonomous ordinary differential equation (ODE):

$$\begin{aligned} \dot{X}_{(t)} &= \tilde{u}(X_{(t)}) = X_{(t)}U \\ \text{with initial condition } X_{(t=0)} &= X_0. \end{aligned} \tag{4.2.3}$$

The solution of Eq. (4.2.3) can be computing by using the *matrix exponential* function [Moler and Van Loan 2003] [Al-Mohy 2010] [Higham and Al-Mohy 2010] by $X_{(t)} = X_0 e^{tU} = X_0 \sum_m t^m U^m / m!$. Notice that endowing \mathbb{M}_n with a norm, one has $\|U^k\| \leq \rho \|U\|^k$ for some finite ρ , and the previous power series absolutely converges for all t . Thus, the resulting integral curve is well defined for any $t \in \mathbb{R}$ and therefore left-invariant vector fields are complete.

The *group exponential* is a function $\exp: \mathfrak{g} \rightarrow \mathbf{H} \subseteq \mathbf{G}: U \mapsto X = \exp(U)$ defined by $X = \gamma_{(t=1)}$, where γ is the integral curve of the left-invariant vector field \tilde{u} starting at the identity. The group exponential is a *one-parameter subgroup* of \mathbf{G} , *i.e.* $\exp(\alpha U) \exp(v U) = \exp((\alpha + v)U)$ for any $U \in \mathfrak{g}$. In particular, $\exp(0) = I$ and $\exp(-U) = (\exp(U))^{-1}$. An algebraic expression for the group exponential function on matrix groups is obtained directly from the solution of Eq. (4.2.3) up to time $t = 1$ in terms of the matrix exponential, $\exp(U) = e^U$. Again, it is a well defined expression for any $U \in \mathfrak{g}$ and therefore the group exponential is a mapping from the whole⁸ group algebra \mathfrak{g} .

The group exponential map is a smooth function and it can be shown that its Jacobian $D\exp(U)|_{(U=0)} = \text{id}_{\mathfrak{g}}$. Therefore, $\exp(\cdot)$ is, at least in a non-empty neighborhood of 0,

⁷ In abstract notation: $\mathbf{Ad}_y = (T_{y^{-1}} \mathbf{L}_y \circ T_y \mathbf{R}_{y^{-1}}): \mathfrak{g} \rightarrow \mathfrak{g}$.

⁸ It can be show that on finite dimensional abstract groups, where the matrix exponential expressions are no longer valid, left-invariant vector fields are also complete and the group exponential function is still a well defined map from the whole group algebra (see for example [Gallier 2013]).

an invertible map. The inverse of the group exponential function is given by the *matrix logarithm* [Higham and Al-Mohy 2010], satisfying $\exp(\log(X)) = X$ for any $X \in \mathbf{H} \subseteq \mathbf{G}$, a non-empty neighborhood around I . For some groups, such as Rotation group, the matrix logarithm function is multivalued and in general it cannot be ensured that $\log(\exp(U)) = U$.

The group exponential function is commonly used to parameterize group elements where we can use (\mathbf{H}, \log) as a local chart of \mathbf{G} around the identity. The coordinates of the algebra vectors are called *canonical coordinates of the first kind* of the group element:

$$X(\nu^1, \dots, \nu^k) = \exp(B_1\nu^1 + B_2\nu^2 + \dots + B_k\nu^k).$$

Another parametrization of the group elements, called *canonical coordinates of the second kind*, is given by:

$$X(\theta^1, \dots, \theta^k) = \exp(B_1\theta^1) \exp(B_2\theta^2) \dots \exp(B_k\theta^k).$$

The different conventions of *Euler angles* are examples of transformations parameterized in this way. Note that, due to the non-commutativity of the group elements, this parametrization depends on the choice of the order of the basis. Both coordinates are commonly used as linear parameterizations of the group where classical linear operations can be directly applied. Moreover, the translated group exponentials joining two elements, representing straight rays in the coordinates, are usually used as smooth paths between group elements [Alexa 2002] [Rahman et al. 2005] [Li and Hao 2006].

Remarks

For finite dimensional groups, the group exponential is a diffeomorphism from an open neighborhood of the algebra to an open neighborhood of the group at the identity. Even though the group exponential is defined for all algebra vectors, not all group elements are the group exponential of an algebra vector. For example, it can be checked that the two dimensional spatial transformation obtained by the composition of a rotation of $3/4\pi$ around the origin followed by a non-isotropic scale of 6 along the horizontal axis, cannot be connected to the identity with a one-parameter subgroup. That is, the matrix

$$\begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(3/4\pi) & -\sin(3/4\pi) \\ \sin(3/4\pi) & \cos(3/4\pi) \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} -6 & -6 \\ 1 & -1 \end{pmatrix}$$

does not have a real matrix logarithm (the conditions for a real matrix to have a real logarithm can be found in [Gallier 2011]). Therefore, the canonical coordinates of the first kind defined before, cannot be used as a valid chart for the whole group. Nevertheless, a smooth atlas to cover the whole group can be defined via left-translations of the group exponential function (see Section 4.3.2).

4.2.6 Riemannian metrics, length of curves, Riemannian geodesics and Riemannian exponential

Let \mathbb{V} be a vector space, an *inner product* is a bilinear, symmetric and positive definite form from a pair of vectors to \mathbb{R} and it is denoted by $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$. The inner product is uniquely identified with a linear map to the dual space $\mathcal{M} : \mathbb{V} \rightarrow \mathbb{V}^*$ which is called the *inertia operator*. Let us denote the pairing between \mathbb{V} and its dual \mathbb{V}^* by $\langle \cdot, \cdot \rangle : \mathbb{V}^* \times \mathbb{V} \rightarrow \mathbb{R} : \langle \mathbf{w}, v \rangle = \mathbf{w}(v)$, where $\mathbf{w} \in \mathbb{V}^*$ is a *linear form*, also called *co-vectors*. Then, the inner product and the inertia operator satisfy $\langle w, v \rangle = \langle \mathcal{M}(w), v \rangle$ (see [Arnold 1989] [Kolev 2004]).

A *Riemannian metric* on a manifold \mathbf{G} is a smooth assignment of inner products to the tangent space at every element of \mathbf{G} , or, equivalently, each $X \in \mathbf{G}$ is associated with an inertia operator \mathcal{M}_X such that $\langle\langle W, V \rangle\rangle_X = \langle \mathcal{M}_X(W), V \rangle$ is the inner product on $T_X \mathbf{G}$. With the assignment $(\mathbf{G}, \langle\langle \cdot, \cdot \rangle\rangle)$ the *length* of a curve segment $\gamma: [t_0, t_1] \subset \mathbb{R} \rightarrow \mathbf{G}$ is defined by

$$\text{Length}(\gamma; t_0, t_1) = \int_{t_0}^{t_1} \langle\langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle\rangle_{\gamma(s)}^{1/2} ds. \quad (4.2.4)$$

The *Riemannian distance* between two elements A and $B \in \mathbf{G}$ that can be connected by a curve is

$$\text{distance}(A, B) = \left\{ \begin{array}{l} \inf_{\gamma \in \mathcal{C}(\mathbf{G})} \text{Length}(\gamma; t_0, t_1) \\ \text{subject to } \gamma(t_0) = A \quad ; \quad \gamma(t_1) = B \end{array} \right. \quad (4.2.5)$$

Endowed with the function $\text{distance}(\cdot, \cdot)$, the group \mathbf{G} acquires a *metric space* structure with the same topology as the original topology on \mathbf{G} given by the atlas $\{\mathcal{D}_i, \varphi_i\}_i$ as in Section 4.2.1.

The length of a curve segment is independent of its parametrization, which means that $\text{Length}(\gamma; t_0, t_1) = \text{Length}(\gamma \circ h; t_0, t_1)$ for any smooth and invertible time warping $h: [t_0, t_1] \rightarrow [t_0, t_1]$ with $h(t_0) = t_0$ and $h(t_1) = t_1$. If the *arc-length* parametrization is used, $\|\dot{\gamma}(s)\|_{\gamma(s)} = \langle\langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle\rangle_{\gamma(s)}^{1/2} = \text{constant}$, then the length of the segment is equal to $(t_1 - t_0) \|\dot{\gamma}(0)\|_{\gamma(0)}$.

In order to avoid multiple solutions of minimal curve length (4.2.4), the following *Riemannian energy* (also called *action* in the area of geometric mechanics) is used:

$$\mathbf{E}(\gamma; t_0, t_1) = \frac{1}{2} \int_{t_0}^{t_1} \langle\langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle\rangle_{\gamma(s)} ds = \frac{1}{2} \int_{t_0}^{t_1} \|\dot{\gamma}(s)\|_{\gamma(s)}^2 ds, \quad (4.2.6)$$

where the factor $1/2$ is included for later algebraic convenience. It is easy to observe that $\mathbf{E}(\gamma)$ does depend on the time parametrization of γ and its value is minimal when $\|\dot{\gamma}(s)\|_{\gamma(s)}$ is constant along s . Therefore,

$$\left. \begin{array}{l} \inf_{\gamma \in \mathcal{C}(\mathbf{G})} \mathbf{E}(\gamma; t_0, t_1) \\ \text{subject to } \gamma(t_0) = A \quad ; \quad \gamma(t_1) = B \end{array} \right\} = \frac{\text{distance}(A, B)^2}{2(t_1 - t_0)}. \quad (4.2.7)$$

Finally, solving Eq. (4.2.7) is equivalent to minimizing **Length** (Eq. (4.2.5)) with the advantage that the solution is an arc-length parameterized curve.

The extremal curves of Eq. (4.2.6) are called *Riemannian geodesics*. In general, given two elements A, B belonging to the same connected component of \mathbf{G} , it may exist several geodesic segments γ satisfying $\gamma(0) = A$ and $\gamma(1) = B$, even having the same length or energy. The minimal length among all geodesic segments is the distance between the elements. Furthermore, given a geodesic segment $\gamma(t)$ passing through A and B , an affine reparametrization $\gamma_{(\alpha t + \tau)}$ is also a geodesic segment connecting A and B . The lengths and energies of all affine reparameterized segments from A to B are the same. As a general rule, when considering geodesics between two points we will restrict our attention to those which at $t = 0$ pass through the first point and at $t = 1$ pass through the second one.

Note that a curve is a smooth function from an open time-interval, not the geometrical set of contiguous points which is the image of a curve. Any two bounded paths with the same image have the same length, however to compute their Riemannian energies, their functional time dependencies are mandatory.

Given a Riemannian metric $(\mathbf{G}, \langle\cdot, \cdot\rangle.)$, for every $X \in \mathbf{G}$ and every $V \in T_X \mathbf{G}$, there exist an open interval $(-\varepsilon, \varrho) \subset \mathbb{R}$ around 0 and a unique geodesic $\gamma(t; X, V)$ for t in the interval, such that $\gamma_{(t=0)} = X$ is its *starting point* and $\dot{\gamma}_{(t=0)} = V$ is its *initial velocity*. This construction of a geodesic satisfies the homogeneity property $\gamma(t; X, V) = \gamma(\alpha t; X, V/\alpha)$ for $\alpha \in \mathbb{R}^+$. With the previous construction, if $t = 1$ is included in the interval $(-\varepsilon, \varrho)$, the *Riemannian exponential* function is defined as the element of the resulting geodesic at $t = 1$

$$\text{Exp}_X : T_X \mathbf{G} \rightarrow \mathbf{G} : V \mapsto \text{Exp}_X(V) = \gamma(1; X, V).$$

The Riemannian exponential function provides geodesics in terms of a starting point and an initial velocity instead of a starting and an ending element: $\gamma(t; X, V) = \text{Exp}_X(tV)$ is the unique geodesic such that $\gamma_{(t=0)} = X$ and $\dot{\gamma}_{(t=0)} = V$. As it is arc-length parameterized, its length under the given metric is $\text{Length}(\text{Exp}_X(tV); 0, 1) = \langle\langle V, V \rangle\rangle_X^{1/2} = \|V\|_X$.

In the cases where $\text{Exp}_X(\cdot)$ can be defined for all $V \in T_X \mathbf{G}$, the Riemannian metric $(\mathbf{G}, \langle\cdot, \cdot\rangle.)$ is called *geodesically complete*, which means that geodesics $\gamma(t; X, V)$ are defined for all $t \in \mathbb{R}$.

The set $\mathfrak{c}_X = \{V \in T_X \mathbf{G} \text{ such that } \|V\|_X = \text{distance}(X, \text{Exp}_X(V))\}$ is star-shaped and its boundary is called the *tangential cut locus*. The $\text{Exp}_X(\cdot)$ function is a diffeomorphism from any open 0-neighborhood $\mathfrak{s}_X \subseteq \mathfrak{c}_X$ onto an open X -neighborhood $\mathbf{H}_X \subseteq \mathbf{G}$. Therefore, the inverse function $\text{Exp}_X^{-1}(\cdot)$ is well defined on \mathbf{H}_X . In computer vision, $\text{Exp}_X^{-1}(\cdot)$ is usually called the *Riemannian logarithm function* [Pennec 2006a] [Fletcher and Joshi 2007] [Fletcher et al. 2009] [Subbarao and Meer 2009] [Sommer et al. 2013] and is denoted by $\text{Log}_X(\cdot)$. The function $\text{Log}_X(\cdot) : \mathbf{H}_X \rightarrow T_X \mathbf{G}$, together with a set of bases $\{Z_1, Z_2, \dots, Z_k\}$ such that $T_X \mathbf{G} \equiv \text{span}(\{Z_i\})$, allows the definition of a local chart of \mathbf{G} around X . Its coordinates are called *normal coordinates* and every element of \mathbf{H}_X can be identified by its normal coordinates as

$$X(\eta^1, \dots, \eta^k) = \text{Exp}_X(Z_1\eta^1 + Z_2\eta^2 + \dots + Z_k\eta^k).$$

Let $A, B \in \mathbf{G}$ such that $B \in \mathbf{H}_A$, then a geodesic connecting both elements is given by $\gamma_t = \text{Exp}_A(t \text{Log}_A(B))$, being $\gamma_{(0)} = A$ and $\gamma_{(1)} = B$. Notice that, while there can exist several velocities $V \in T_X \mathbf{G}$ such that $\text{Exp}_A(V) = B$, only that V belonging to the interior of \mathfrak{c}_X is considered as $\text{Log}_A(B)$. The corresponding length of the curve segment from $t = 0$ to $t = 1$ is $\text{Length}(\gamma; 0, 1) = \|\text{Log}_A(B)\|_A$. Moreover, this geodesic segment is the shortest one and therefore its length is the Riemannian distance between A and B . Note that, while $\text{Exp}(\cdot)$ identifies geodesics in terms of starting points and velocities, $\text{Log}(\cdot)$ allows the identification of geodesics in terms of their initial and ending points.

4.2.7 Isometries

Given two groups with assigned Riemannian metrics $(\mathbf{G}_a, \langle\cdot, \cdot\rangle^a)$ and $(\mathbf{G}_c, \langle\cdot, \cdot\rangle^c)$ and let $\Phi : \mathbf{G}_a \rightarrow \mathbf{G}_c$ be a smooth function. The tangent-lift of the function Φ is the linear function $T_X \Phi : T_X \mathbf{G}_a \rightarrow T_{\Phi(X)} \mathbf{G}_c$ which translate the velocity of a curve γ through $X \in \mathbf{G}_a$ to the velocity of the curve $\Phi(\gamma)$ through $\Phi(X) \in \mathbf{G}_c$.

It is say that both groups are *isometric* if there exist an invertible and differentiable function Φ such that

$$\langle\langle V_1, V_2 \rangle\rangle_X^a = \langle\langle T_X \Phi V_1, T_X \Phi V_2 \rangle\rangle_{\Phi(X)}^c,$$

for all $X \in \mathbf{G}_a$ and for all $V_1, V_2 \in T_X \mathbf{G}_a$.

Under an isometry function $\Phi(\cdot)$, the following identities are satisfied

$$\begin{aligned} \text{distance}^a(X, Y) &= \text{distance}^c(\Phi(X), \Phi(Y)) \\ \Phi(\text{Exp}_X^a(V)) &= \text{Exp}_{\Phi(X)}^c(T_X \Phi V), \end{aligned}$$

for any $X \in \mathbf{G}_a$ and any $V \in T_X \mathbf{G}_a$, where $\mathbf{distance}^{a/c}$ and $\text{Exp}^{a/c}$ correspond to the Riemannian metrics $\langle\langle \cdot, \cdot \rangle\rangle^a$ and $\langle\langle \cdot, \cdot \rangle\rangle^c$, respectively.

4.2.8 Invariant Riemannian metrics

In the following the group structure and group operations will be considered on the manifold \mathbf{G} leading to metrics that preserve the group symmetries and these symmetries will be inherited by the Riemannian exponential $\text{Exp}(\cdot)$. Two families of Riemannian metrics on Lie groups arise naturally: *left-invariant* and *right-invariant* metrics.

In this exposition, without loss of generality, left-invariant metrics are analyzed, but right-invariant ones can be likewise considered or eventually solved by means of a change of parametrization (see later Eqs. (4.2.12) and (4.2.13)). Besides, Section 4.8 summarizes most expressions to the right-invariant case.

A left-invariant metric is an isometry under left-translations, *i.e.* the inner products fulfil

$$\langle\langle V_1, V_2 \rangle\rangle_X = \langle\langle T_X \mathbf{L}_Y V_1, T_X \mathbf{L}_Y V_2 \rangle\rangle_{YX} = \langle\langle YV_1, YV_2 \rangle\rangle_{YX}$$

for any $X, Y \in \mathbf{G}$ and any $V_1, V_2 \in T_X \mathbf{G}$. In particular,

$$\langle\langle V_1, V_2 \rangle\rangle_X = \langle\langle X^{-1}V_1, X^{-1}V_2 \rangle\rangle_I = \langle\langle U_1, U_2 \rangle\rangle_I, \quad (4.2.8)$$

where U_i are the velocities V_i translated from the left to the algebra \mathfrak{g} (V_i are left-trivialized to U_i). Any inner product at the algebra can be propagated to the whole group using Eq. (4.2.8) defining a left-invariant Riemannian metric. Correspondingly, an inertia operator $\mathcal{M}_I: \mathfrak{g} \rightarrow \mathfrak{g}^*$ (henceforth referred to as \mathcal{M}) can be defined and properly propagated⁹ to \mathcal{M}_X . In terms of algebra coordinates, \mathcal{M} can be represented by a $k \times k$ symmetric and positive definite matrix, denoted by \mathbf{m} , such that

$$\langle\langle U_1, U_2 \rangle\rangle_I = \langle \mathcal{M}(U_1), U_2 \rangle = \sum_{i,j}^k \mu_1^i \mu_2^j \langle \mathcal{M}(B_i), B_j \rangle = \bar{\mu}_1^T \mathbf{m} \bar{\mu}_2, \quad (4.2.9)$$

where $U_1 = \sum_i B_i \mu_1^i$ and $U_2 = \sum_i B_i \mu_2^i$ are matrix representations of the algebra elements. In particular, the induced norm results

$$\|U\|_I^2 = \langle\langle U, U \rangle\rangle_I = \bar{\mu}^T \mathbf{m} \bar{\mu}. \quad (4.2.10)$$

The entries $\mathbf{m}^{i,j}$ measure the energy of differential transformations around the identity. These values are application dependent and some heuristic strategies for their selection are given in [Park 1995] and [Woods 2003, Appendix 3].

Following Section 4.2.7, the corresponding Riemannian distance is invariant to left-translations

$$\mathbf{distance}(X, Y) = \mathbf{distance}(ZX, ZY) = \mathbf{distance}(I, X^{-1}Y)$$

for any $X, Y, Z \in \mathbf{G}$. The same invariance property applies to geodesics, yielding

$$Y \text{Exp}_X(tV) = \text{Exp}_{YX}(tYV), \quad (4.2.11)$$

whose length is

$$\mathbf{Length}(\text{Exp}_X(tV); 0, 1) = \langle\langle V, V \rangle\rangle_X^{1/2} = \langle\langle U, U \rangle\rangle_I^{1/2} = \|U\|_I,$$

⁹ In abstract notation, \mathcal{M}_e is left-invariantly propagated to the whole group as $\mathcal{M}_x = ((T_x \mathbf{L}_{x^{-1}})^* \circ \mathcal{M}_e \circ T_x \mathbf{L}_{x^{-1}}) : T_x \mathcal{G} \rightarrow T_x^* \mathcal{G}$, where $(T_x \mathbf{L}_{x^{-1}})^*$ is the cotangent-lift of the \mathbf{L}_x which translate linear functions from \mathfrak{g}^* to $T_x^* \mathcal{G}$, see [Arnold 1989, Appendix B].

with $U = X^{-1}V \in \mathbf{g}$. Similarly, $Z \operatorname{Log}_X(Y) = \operatorname{Log}_{ZX}(ZY)$.

Whenever right-invariance is required, all previous concepts can be derived using the same procedure but replacing left-translations $\mathbf{L}(\cdot)$ with right-translations $\mathbf{R}(\cdot)$. Alternatively, it can be shown that left-invariant and right-invariant metrics are isometric under the group inverse function $\operatorname{inv}(\cdot)$ which can be seen from Eq. 4.2.1. Thus, from Section 4.2.7, the following relations can be used¹⁰:

$$\operatorname{Exp}_I^{right}(U) = \left(\operatorname{Exp}_I^{left}(-U) \right)^{-1} \quad (4.2.12)$$

$$\operatorname{distance}^{right}(X, Y) = \operatorname{distance}^{left}(X^{-1}, Y^{-1}). \quad (4.2.13)$$

When the Riemannian exponential function is developed around I the subscript is usually omitted, $\operatorname{Exp}(U) \equiv \operatorname{Exp}_I(U)$, and it is a mapping from \mathbf{g} to \mathbf{G} (correspondingly $\operatorname{Log}(X) \equiv \operatorname{Log}_I(X) : \mathbf{G} \rightarrow \mathbf{g}$). Riemannian exponential functions starting at I will be usually considered, and Eq. (4.2.11) can be used to get $\operatorname{Exp}(\cdot)$ developed at a different point.

Remarks

For a Lie group endowed with an invariant Riemannian metric, geodesics are complete, *i.e.* Riemannian geodesics can be extended to the whole real line and $\operatorname{Exp}_I(\cdot)$ maps the whole algebra onto the group [Gallier 2013]. Moreover, according to the Hopf–Rinow Theorem, any two elements $A, B \in \mathbf{G}$ belonging to the same connected component of the group can be connected with a Riemannian geodesic, *i.e.* there exist at least one velocity V such that $\operatorname{Exp}_A(V) = B$.

Unlike the group exponential function, the Riemannian exponential function from the identity under an invariant metric is a surjective mapping from algebra vectors to the identity component of the group.

4.2.9 Totally geodesic subgroups

Given a group equipped with a Riemannian metric $(\mathbf{G}, \langle \cdot, \cdot \rangle)$ and a proper subgroup \mathbf{H} with the induced metric, then \mathbf{H} is called a *totally geodesic subgroup* of \mathbf{G} with respect to $\langle \cdot, \cdot \rangle$ if each geodesic in \mathbf{H} is also a geodesic in \mathbf{G} .

A sufficient and necessary condition to be a totally geodesic subgroup with respect to an invariant Riemannian metric is

$$\langle\langle U_h, [U_h, U_n] \rangle\rangle_I = 0$$

for all $U_h \in \mathbf{h}$ (the algebra of \mathbf{H}) and $U_n \in \mathbf{h}^\perp$, where \mathbf{h}^\perp is the orthogonal complement of \mathbf{h} with respect to $\langle \cdot, \cdot \rangle_I$. For more details and extended properties see [Modin et al. 2011].

The concept of totally geodesic subgroup is very useful in those scenarios where geodesics of the group \mathbf{G} can be easily computed. It will be used in Section 5.1 to give explicit solutions of geodesics in some groups.

4.2.10 Bi-invariant metrics

When a Riemannian metric is both left- and right-invariant, it is called *bi-invariant* and the Riemannian exponential function developed around the identity coincides with the group exponential function. Therefore, geodesics are translated one-parameter subgroups through the identity

$$\operatorname{Exp}_X(V) = X \exp(X^{-1}V) = \exp(VX^{-1})X. \quad (4.2.14)$$

¹⁰ In general and for abstract groups: $\operatorname{Exp}_x^{right}(\mathcal{V}) = \operatorname{inv}\left(\operatorname{Exp}_{x^{-1}}^{left}(-(T_{x^{-1}}\mathbf{L}_x)^{-1} \circ T_x\mathbf{R}_{x^{-1}}\mathcal{V})\right)$.

For *commutative groups*, where \mathbf{L}_X coincides with \mathbf{R}_X , any left-invariant metric is bi-invariant and geodesics can be computed from Eq. (4.2.14) by using the matrix exponential function.

Some non-commutative groups admit a bi-invariant metric, although not all left-invariant metrics are bi-invariant. Compact groups always admit bi-invariant metrics [Arsigny 2006]. Some non-compact group also admit bi-invariant metrics, but this is not the general case. For example, no bi-invariant metric exists for the group of rigid transformations $\mathcal{SE}(d)$ for $d \geq 2$ [Žefran et al. 1999].

A sufficient and necessary condition for a metric to be a bi-invariant is [Milnor 1976] [Do Carmo 1992] [Gallier 2013]

$$\langle\!\langle [U_1, W], U_2 \rangle\!\rangle_I = -\langle\!\langle U_1, [U_2, W] \rangle\!\rangle_I$$

for any $U_1, U_2, W \in \mathfrak{g}$.

Finally, if the group does not admit a bi-invariant Riemannian metric, then no bi-invariant C^2 (with continuous derivative up to order 2) distance metric can be defined on the group [Park 1995].

4.2.11 Geodesics as “straightest” curves

For the sake of completeness, this section introduces geodesics defined via the concept of “straightest” curve in a manifold. While there is a natural way to move forward keeping a direction fixed in a Euclidean space, an additional structure is needed in a manifold to define what keep a direction means. This added structure must allow to map velocities between nearby tangent spaces and is defined by an affine connection which quantifies the derivative of vector fields on the manifold. Most textbooks about differential geometry define geodesics in terms of the affine connection structure.

We will outline in the following that, for a given Riemannian metric there always exist affine connections whose “straightest” geodesics coincide with Riemannian (“shortest”) geodesics defined in Section 4.2.6. Analogously, for a given group there always exist affine connections whose geodesics coincide with the translated group exponential (Section 4.2.5 and Section 4.2.10).

Intuitive notions

Let us start by describing an intuitive description of the derivative of a vector field with respect to another vector field, both defined on a Riemannian manifold. In a Euclidean space, two lines are parallel provided they make equal angles with a transversal. Generalizing this concept, it can be said that a vector field on a Riemannian manifold is parallel provided that any two integral curves from it make equal angles with a geodesic joining their base points. It allows the definition of a *transport* of a tangent vector along a geodesic by smoothly translate the vector along the geodesic while its inner product with the velocity of the geodesic is maintained. Then, given a geodesic γ we can choose a collection of linear maps $\{\Gamma_{s \rightarrow t} : T_{\gamma(s)} \mathbf{G} \rightarrow T_{\gamma(t)} \mathbf{G} \mid s, t \text{ in the domain of } \gamma\}$ such that

- $\langle\!\langle V, \dot{\gamma}_{(s)} \rangle\!\rangle_{\gamma(s)} = \langle\!\langle \Gamma_{s \rightarrow t} V, \dot{\gamma}_{(t)} \rangle\!\rangle_{\gamma(t)}$, for any $V \in T_{\gamma(s)} \mathbf{G}$.
- $\Gamma_{s \rightarrow s}$ is the identity map on $T_{\gamma(s)} \mathbf{G}$.
- $\Gamma_{u \rightarrow t} \circ \Gamma_{s \rightarrow u} = \Gamma_{s \rightarrow t}$.
- $\Gamma_{s \rightarrow t}$ depends smoothly on s and t .

Using this transport it can be defined a new kind of derivative of a smooth vector field. Let $\mathcal{R}: \gamma_{(t)} \rightarrow T_{\gamma(t)} \mathbf{G}$ be a vector field defined on γ , then the derivative of \mathcal{R} with respect to $\Gamma_{s \rightarrow t}$ is defined by

$$\frac{D_\Gamma}{dt}(\mathcal{R}) \Big|_{\gamma(t)} = \lim_{s \rightarrow t} \frac{\Gamma_{s \rightarrow t} \mathcal{R}(\gamma(s)) - \mathcal{R}(\gamma(t))}{s - t} \in T_{\gamma(t)} \mathbf{G}.$$

It can be shown that, once the transport $\Gamma_{s \rightarrow t}$ is chosen, the previous derivative depends only on the velocity of the geodesic at t . Although $\Gamma_{s \rightarrow t}$ allows to take derivatives along the geodesic γ , a consistent rule have to be specified to select the maps $\Gamma_{s \rightarrow t}$ such that depends smoothly between nearby geodesics. When this consistency is achieved, the derivative $\frac{D_\Gamma}{dt}$ can be naturally extended to a new operation called *affine connection*, that locally takes the derivative of a smooth vector field with respect to another smooth vector field.

The entire process can be reversed such that, given an affine connection, one can uniquely define the transport $\Gamma_{s \rightarrow t}$ that agrees with the affine connection. Furthermore, the construction of an affine connection will allow to define transports on manifolds without a Riemannian structure.

Affine connections, covariant derivative

An *affine connection* on a manifold, denoted by $\nabla_{\mathcal{H}} \mathcal{S} = \mathcal{Z}$, is a smooth map which takes two smooth vector fields \mathcal{H} and \mathcal{S} and returns another smooth vector field \mathcal{Z} , satisfying the following properties [Burns and Gidea 2005]:

- $\nabla_{(\mathcal{H} + \mathbf{f}\mathcal{P})} \mathcal{S} = \nabla_{\mathcal{H}} \mathcal{S} + \mathbf{f} \nabla_{\mathcal{P}} \mathcal{S}$
- $\nabla_{\mathcal{H}} (\mathcal{S} + \mathbf{f}\mathcal{P}) = \nabla_{\mathcal{H}} \mathcal{S} + \mathbf{f} \nabla_{\mathcal{H}} \mathcal{P} + (\mathcal{H}\mathbf{f})\mathcal{P}$

for $\mathcal{H}, \mathcal{S}, \mathcal{P}$ smooth vector fields and \mathbf{f} a smooth scalar function on the manifold (in particular, by taking \mathbf{f} constant, ∇ results bilinear). From the previous properties follows that $\mathcal{Z}(X) = (\nabla_{\mathcal{H}} \mathcal{S})(X)$ depends on the value $\mathcal{H}(X)$ and on the values of \mathcal{S} in a neighborhood of X .

In terms of a coordinate chart, the affine connection ∇ can be expressed by

$$\nabla_{\partial_i} \partial_j = \sum_{c=1}^k \Gamma_{ij}^c \partial_c,$$

such that, if $\mathcal{H} = \sum h^c \partial_c$ and $\mathcal{S} = \sum s^c \partial_c$ where h^c and s^c are the coordinates of \mathcal{H} and \mathcal{S} respectively, then

$$\nabla_{\mathcal{H}} \mathcal{S} = \sum_{c=1}^k \left(\sum_{i,j}^k h^i s^j \Gamma_{ij}^c + \sum_{i=1}^k h^i (\partial_i s^c) \right) \partial_c,$$

where the functions $\Gamma_{i,j}^c$ are called the *Christoffel symbols* of ∇ . Conversely, any set of k^3 smooth functions $\Gamma_{i,j}^c$ can be chosen in a chart to define an affine connection.

The affine connection ∇ defines univocally the *covariant derivative along a curve* γ , denoted by $\frac{D_\gamma}{dt}$ (the subscript γ is usually omitted when the curve is clear from the context), that assigns to every smooth vector field \mathcal{R} defined on γ another smooth vector field along γ . While the operation $\nabla(\cdot)$ is defined for vector fields defined on the whole manifold (or in an open set of the manifold), $\frac{D_\gamma}{dt}(\cdot)$ can be applied to vector fields defined only on the curve. The covariant derivative along the curve $\frac{D_\gamma}{dt}$ is compatible with the affine connection ∇ in the sense that $\frac{D_\gamma}{dt}(\mathcal{S} \circ \gamma) = (\nabla_{\dot{\gamma}} \mathcal{S}) \circ \gamma$ for any smooth vector field \mathcal{S} defined on the whole manifold (or in an open set containing the curve γ). A vector field \mathcal{R} along a curve γ is called *parallel* if $\frac{D_\gamma}{dt}(\mathcal{R}) = 0$, similarly, a vector field \mathcal{S} on the whole manifold is called *parallel with respect to the curve* γ if $\frac{D_\gamma}{dt}(\mathcal{S} \circ \gamma) = 0$.

In terms of an affine connection ∇ defined on the manifold, a ∇ -geodesic is a curve γ with acceleration $\frac{D\dot{\gamma}}{dt}\dot{\gamma}$ equals to zero¹¹, which means that *the velocity of the curve is parallel to the curve*. In coordinates of a chart, where $\bar{q}_{(t)}$ are the coordinates of $\gamma_{(t)}$, the previous condition results

$$\ddot{q}^c + \Gamma_{i,j}^c(\bar{q})\dot{q}^i\dot{q}^j = 0. \quad (4.2.15)$$

Given a point X in the manifold and a velocity V in the corresponding tangent space there exist an interval $(-\varepsilon, \varrho) \subset \mathbb{R}$ containing 0 and a unique ∇ -geodesic γ defined for t in the interval that satisfies $\gamma_{(0)} = X$ and $\dot{\gamma}_{(0)} = V$. In the case that the interval includes 1, then the ∇ -exponential function is defined by $\text{EXP}_X(V) = \gamma_{(t=1)}$.

Relation with Riemannian geodesics

The previous concept of parallelism does not take into account the concept of angles between tangent vectors which is a concept introduced in the manifold once it is endowed with a Riemannian metric structure. Given a Riemannian metric $(\mathbf{G}, \langle\cdot, \cdot\rangle)$ it is said that the affine connection ∇ is *compatible with the Riemannian metric* when, for any curve γ and two parallel vector fields along it \mathcal{R} and \mathcal{P} , the inner products $\langle\langle \mathcal{R}(\gamma_{(t)}), \mathcal{P}(\gamma_{(t)}) \rangle\rangle_{\gamma_{(t)}}$ remain constant along time. This condition is satisfied if and only if $\mathcal{H}\langle\langle \mathcal{R}, \mathcal{P} \rangle\rangle = \langle\langle \nabla_{\mathcal{H}}\mathcal{R}, \mathcal{P} \rangle\rangle + \langle\langle \mathcal{R}, \nabla_{\mathcal{H}}\mathcal{P} \rangle\rangle$ for all smooth vector fields $\mathcal{R}, \mathcal{P}, \mathcal{H}$ defined on the manifold.

The *Fundamental Theorem of Riemannian geometry* states that: for a given (pseudo-)Riemannian metric¹² $(\mathbf{G}, \langle\cdot, \cdot\rangle)$ there exist a unique affine connection on \mathbf{G} , denoted by $\nabla^{<\cdot>}$, that is compatible with the metric $\langle\cdot, \cdot\rangle$ and fulfills that¹³ $\nabla_{\mathcal{H}}^{<\cdot>} \mathcal{S} - \nabla_{\mathcal{S}}^{<\cdot>} \mathcal{H} = \mathcal{H} \circ \mathcal{S} - \mathcal{S} \circ \mathcal{H}$. The resulting affine connection is called the *Levi-Civita connection* for the given (pseudo-)metric. With respect to a coordinate chart, the Christoffel symbols Γ_{ij}^c corresponding to the Levi-Civita connection $\nabla^{<\cdot>}$ satisfy at each point

$$2 \sum_{c=1}^k \Gamma_{ij}^c \langle\langle \partial_c, \partial_m \rangle\rangle = \partial_i \langle\langle \partial_j, \partial_m \rangle\rangle + \partial_j \langle\langle \partial_i, \partial_m \rangle\rangle - \partial_m \langle\langle \partial_i, \partial_j \rangle\rangle. \quad (4.2.16)$$

It can be shown that $\nabla^{<\cdot>}$ -geodesics are local extremals of the corresponding Riemannian energy given in Eq. (4.2.6), leaving the endpoints fixed [Postnikov 1967]. That is, given a $\nabla^{<\cdot>}$ -geodesic $\gamma: [t_0, t_1] \subset \mathbb{R} \rightarrow \mathbf{G}$, then the variation of the Riemannian energy

$$\delta \frac{1}{2} \int_{t_0}^{t_1} \langle\langle \dot{\gamma}_{(s)}, \dot{\gamma}_{(s)} \rangle\rangle_{\gamma_{(s)}} ds$$

vanishes under perturbations $\delta\gamma_{(s)}$ of the curve with $\delta\gamma_{(t_0)} = \delta\gamma_{(t_1)} = 0$. Therefore, the curve γ is a Riemannian geodesic. Thus, under the $\nabla^{<\cdot>}$ affine connection $\text{EXP}_X(V) \equiv \text{Exp}_X(V)$ for the metric $\langle\cdot, \cdot\rangle$.

Left-invariant affine connections

On a group \mathbf{G} , an affine connection ∇ , is said to be left-invariant [Postnikov 2001] if for any two left-invariant vector fields $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{S}}$, the resulting $\nabla_{\tilde{\mathcal{H}}} \tilde{\mathcal{S}}$ is also a left-invariant vector field¹⁴.

¹¹ The notation $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ is also common in many textbooks and papers.

¹² A *pseudo-Riemannian metric* is also a smooth assignment of non-degenerate bilinear forms between tangent vectors but without the requirement to be positive-definite.

¹³ Given an affine connection ∇ , the bilinear form defined by $T(\mathcal{H}, \mathcal{S}) = \nabla_{\mathcal{H}}\mathcal{S} - \nabla_{\mathcal{S}}\mathcal{H} - (\mathcal{H} \circ \mathcal{S} - \mathcal{S} \circ \mathcal{H})$ is called the *torsion tensor*. When the torsion tensor vanishes, the affine connection ∇ is called *symmetric*.

¹⁴ Alternatively, left-invariant connections are also defined [Pennec and Arsigny 2012] [Lorenzi and Pennec 2013b] [Saccon et al. 2013] as those satisfying $\nabla_{(T\mathbf{L}_Y\mathcal{H})}(T\mathbf{L}_Y\mathcal{S}) = T\mathbf{L}_Y(\nabla_{\mathcal{H}}\mathcal{S})$ for any $Y \in \mathbf{G}$ and any smooth vector fields \mathcal{H} and \mathcal{S} .

As any smooth vector field can be described as a linear combination of left-invariant vector fields (see Section 4.2.3), it can be shown that any left-invariant affine connection is completely defined by its action on algebra vectors. Therefore, ∇ is uniquely defined in terms of a bilinear multiplication operation known as *left connection function* for ∇ [Saccon et al. 2013]

$$\alpha(\cdot, \cdot) : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g} : (U, W) \mapsto \alpha(U, W) = (\nabla_{\tilde{u}} \tilde{w})(I).$$

In terms of algebra coordinates, $U = \sum_i^k B_i \mu^i$ and $W = \sum_i^k B_i \omega^i$, the multiplication α is determined by a $k \times k \times k$ array α_{ij}^c such that

$$(\alpha(U, W))^c = \sum_{i,j} \alpha_{ij}^c \mu^i \omega^j$$

and any array of k^3 numbers defines univocally a left-invariant affine connection. The array α_{ij}^c is equal to $\Gamma_{ij}^c(I)$ when the Christoffel symbols are written for the group exponential chart.

For a left-invariant affine connection ∇ , the corresponding ∇ -geodesics are left-invariant in the sense that if γ is a ∇ -geodesic passing through X with velocity $V \in T_X \mathbf{G}$ at $t = 0$, then the curve $\mathbf{L}_Y \gamma$ is also a ∇ -geodesic but passing through YX with velocity $T_X \mathbf{L}_Y V \in T_{YX} \mathbf{G}$ at $t = 0$. Correspondingly, the ∇ -exponential function is left-invariant, $Y \text{EXP}_X(tV) = \text{EXP}_{YX}(tYV)$. In terms of the array α_{ij}^c a ∇ -geodesic depends only on the symmetric part $(\alpha_{ij}^c + \alpha_{ji}^c)/2$. It is interesting to point that, under a left- (right-)invariant affine connection, the domain of ∇ -geodesics can be extended to the whole \mathbb{R} .

For a left-invariant Riemannian metric $\langle \cdot, \cdot \rangle$ on \mathbf{G} , its corresponding Levi-Civita connection $\nabla^{\langle \cdot, \cdot \rangle}$ is a left-invariant affine connection [Arnold and Khesin 1998] [Kolev 2007] and satisfies¹⁵

$$2\alpha(U, W) = [U, W] - \mathcal{M}^{-1}(\text{ad}_U^*(\mathcal{M}(W))) - \mathcal{M}^{-1}(\text{ad}_W^*(\mathcal{M}(U))),$$

where ad^* is the *dual* of the adjoint operator (see later Section 4.3.3).

Bi-invariant affine connections

Left-invariant connections ∇^λ defined by $\alpha(U, W) = \lambda[U, W]$, where λ is any real number, are also right-invariant connections and are called *Cartan connections* [Laquer 1992] [Postnikov 2001] [Pennec and Arsigny 2012] [Pennec 2013]. Their symmetric part vanish for all λ and therefore their ∇^λ -geodesics coincide. Moreover, ∇^λ -geodesics departing from the identity coincide with the one-parameter subgroup curves defined in Section 4.2.5. Therefore, under a Cartan connection, $\text{EXP}_I(U) \equiv \exp(U)$ for any $U \in \mathfrak{g}$ and, more generally $\text{EXP}_X(V) = X \exp(X^{-1}V)$.

Final remarks

To summarize this background section, a few remarks about the definition of geodesics are given below. Roughly speaking, a geodesic on a manifold is a curve, *i.e.* a time-dependent smooth function, which may be specified by nothing more than a point passing through and its velocity in that point. A more elaborated and formal definition requires additional information about the structures defined on the given manifold. In the case of a manifold endowed with an affine connection, a geodesic is a curve whose velocity is parallel and the curve goes forward and

¹⁵ In some works such as [Lorenzi and Pennec 2013b] [Hinkle et al. 2013] it can be found the following equivalent expression: $2\alpha(U, W) = [U, W] - \text{ad}_U^T(W) - \text{ad}_W^T(U)$, where $\text{ad}_{(\cdot)}^T = (\mathcal{M}^{-1} \circ \text{ad}_{(\cdot)}^* \circ \mathcal{M})$ is the *transpose* of ad with respect to the inner product $\langle \cdot, \cdot \rangle$ satisfying, $\langle U_1, \text{ad}_W U_2 \rangle = \langle \text{ad}_W^T U_1, U_2 \rangle$. The distinction between ad_W^* and ad_W^T is that the former is defined for a “raw” vector space \mathbb{V} while the latter makes sense only when an inner product is assigned to \mathbb{V} .

backwards in the “straightest” way. By endowing the manifold with a Riemannian metric, a geodesic is an arc-length parameterized curve which is formed by concatenating local shortest segments. Instead, if a pseudo-Riemannian metric is defined on the manifold, the concept of length has no sense and a geodesic is defined as a curve connecting two fixed points with an extreme value of its Riemannian energy. This last description encompasses the second one and is compatible with the first one when the corresponding Levi-Civita connection is used. For more complex structures assigned to a manifold, such as *Finsler metrics*, different definitions of geodesics on the manifold are obtained, but these topics are beyond the scope of this thesis.

Three different exponential functions were defined in this section: the ∇ -exponential, upper case $\text{EXP}(\cdot)$, is a function defined by an affine connection; the Riemannian exponential, capitalized $\text{Exp}(\cdot)$, is a function on a manifold equipped with a (pseudo-)Riemannian metric; and the group exponential, lower case $\exp(\cdot)$, which is defined for groups and only group operations are involved in its definition. On a Lie group, the three curves, $\text{EXP}_I(tU)$, $\text{Exp}_I(tU)$, and $\exp(tU)$, start at the identity with the same velocity $U \in \mathfrak{g}$. The former is the “straightest” curve, the second is guided by a metric yielding *least action* curves, and the latter is guided by the group composition law generating one-parameter subgroups.

In a spatial transformation group (or, in general, for any group) exponential maps and geodesics can be defined taking into account the symmetries of the group. The symmetries of the group can be given to the affine connection structure resulting in left- (right-)invariant connections. Also, a Riemannian metric can inherit the symmetries of the group operations obtaining left- (right-)invariant metrics. A nice property of these invariant Riemannian metrics is that $\text{Exp}_X(\cdot)$ is a surjective mapping from $T_X\mathbf{G}$ to the connected component of the group. In general, the surjectivity of ∇ -exponential function defined in terms of an affine connection is only guaranteed for those connections which are compatible with a Riemannian metric. In addition, a Riemannian metric provides a metric space structure to the manifold which is a key requirement for many applications, for example those involving statistical analysis on Lie groups.

4.3 Computing left-invariant Riemannian geodesics

Several descriptions of a geodesic were overviewed in the previous section, but from now on only Riemannian geodesics will be considered and will be called geodesics for short.

Besides the theoretical aspects of geodesics presented before, the next step is to know how to compute invariant geodesics and the Riemannian exponential and logarithm functions. Unfortunately, explicit expressions of them are only known for few matrix groups and under only a few particular inertia operators and most solutions must be faced by numerical algorithms.

The manifold formed by the elements of an $n \times n$ matrix group can be considered as an algebraic embedded manifold of \mathbb{R}^{n^2} . There exist several proposals to compute geodesics in algebraic embedded manifolds, embedded implicit functions [Dedieu and Nowicki 2005], or also for discretized spaces [Sethian 1999] and surfaces [Peyré and Cohen 2006]. However, the Riemannian metric is almost always considered as the restriction of the Euclidean metric in the large space (\mathbb{R}^{n^2} in our case) on the corresponding tangent planes. Excepting some simple cases, invariant metrics on a matrix group do not correspond to the projection of the \mathbb{R}^{n^2} embedding metric¹⁶.

In the following, several approaches for computing left-invariant geodesics are presented. For the cases where right-invariant metrics are considered, Eqs. (4.2.12) and (4.2.13) can be used to

¹⁶ Regarding this point, it is interesting to remark the *Nash embedding Theorems* [Han and Hong 2006] which assert that every k -dimensional Riemannian manifold can be isometrically embedded into a Euclidean space of dimension no larger than $k(k+1)(3k+11)/2$. However, follow this approach becomes rather complicated [Verma 2013] and will be not treated here.

compute right-invariant geodesics in terms of the left-invariant Riemannian exponential function. Alternatively, the right-invariant versions of the following approaches are gathered in Section 4.8.

4.3.1 Computing geodesic in coordinate charts

Let (\mathcal{D}, φ) be a chart of a k -dimensional group \mathcal{G} and let $\Omega \subseteq \mathbb{R}^k$ be the image of φ , as was defined in Section 4.2.1. Then, every element $\chi \in \mathcal{D}$ can be described by its coordinates $\bar{x} = \varphi(\chi) = (x^1, x^2, \dots, x^k)^T$. In the following it will be assumed that all variables involved in the expressions can be represented in the same chart (\mathcal{D}, φ) . Whenever it is not possible, the variables must be switched to be in another overlapping chart.

The left-translation operator, \mathbf{L}_y , was introduced in Section 4.2.3. In chart coordinates the operation $\mathsf{L}_{\bar{y}}$ can be defined as $\mathsf{L}_{\bar{y}}\bar{x} = \varphi(\mathbf{L}_{\varphi^{-1}(\bar{y})}\varphi^{-1}(\bar{x})) = \varphi(\varphi^{-1}(\bar{y}) \bullet \varphi^{-1}(\bar{x}))$. Analogously, the operation $(\bar{x})^{-1}$ can be defined as the coordinates of the element χ^{-1} , i.e. $(\cdot)^{-1}: \Omega \rightarrow \Omega: \bar{x} \mapsto (\bar{x})^{-1} = (\varphi \circ \mathbf{inv} \circ \varphi^{-1})(\bar{x})$.

A curve $\gamma: \mathbb{R} \rightarrow \mathcal{G}$ with $\gamma_{(0)} = \chi$ and $\dot{\gamma}_{(0)} = \nu \in T_\chi \mathcal{G}$ can be described in chart coordinates as $\varphi \circ \gamma = \bar{\gamma}: \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}^k$ passing through $\bar{x} = \varphi(\chi)$ at $t = 0$ with velocity $\bar{v} = d_s \bar{\gamma}(s)|_{(s=0)} \in T_{\bar{x}} \Omega \equiv \mathbb{R}^k$. We define the function $d\varphi: T\mathcal{G} \rightarrow T\Omega \equiv \mathbb{R}^k$ as the linear operator such that $d\varphi(\nu) = \bar{v}$.

Finally, the tangent-lift of the operation \mathbf{L}_y , $T_\chi \mathbf{L}_y \nu$ which translates the velocity $\nu \in T_\chi \mathcal{G}$ to $T_{y\chi} \mathcal{G}$, has a corresponding operation in coordinates $T_{\bar{x}} \mathsf{L}_{\bar{y}} \bar{v} = (d\varphi \circ T_{\varphi^{-1}(\bar{x})} \mathbf{L}_{\varphi^{-1}(\bar{y})} \circ (d\varphi)^{-1}) \bar{v} = \bar{w} \in T_{(\mathsf{L}_{\bar{y}}\bar{x})} \Omega \equiv \mathbb{R}^k$, which is the velocity of the curve $\mathsf{L}_{\bar{y}}\bar{\gamma}$ at $t = 0$.

In the Riemannian framework, each tangent space $T_\chi \mathcal{G}$ is endowed with an inner product. In addition, the local coordinate chart $\varphi(\chi) = (x^1, x^2, \dots, x^k)^T$ induces a coordinate basis $\partial_{\bar{x}} = (\partial_{x^1}, \partial_{x^2}, \dots, \partial_{x^k})^T$ on each tangent space. The corresponding inner product can then be expressed by a $k \times k$ symmetric and positive-definite matrix of the form $\mathbf{m}_{\bar{x}}^{i,j} = \langle\langle \partial_{x^i}, \partial_{x^j} \rangle\rangle_\chi$, such that given curves γ and $\xi \in \mathcal{C}(\mathcal{G})$ passing through χ at $t = 0$ with coordinates $\bar{\gamma}$ and $\bar{\xi}$ respectively, then the inner product $\langle\langle d_s \gamma(s)|_{(s=0)}, d_s \xi(s)|_{(s=0)} \rangle\rangle_\chi = (\partial_{\bar{x}} \bar{\gamma})^T \mathbf{m}_{\bar{x}} (\partial_{\bar{x}} \bar{\xi})$. The matrix $\mathbf{m}_{\bar{x}}$ is the local representation the inertia operator \mathbf{M}_χ (and therefore of the Riemannian metric $(\mathcal{G}, \langle\langle \cdot, \cdot \rangle\rangle)$) in the chart (\mathcal{D}, φ) .

According to Section 4.2.8, a left-invariant metric fulfils $\langle\langle \cdot, \cdot \rangle\rangle_\chi = \langle\langle T_\chi \mathbf{L}_{\chi^{-1}} \cdot, T_\chi \mathbf{L}_{\chi^{-1}} \cdot \rangle\rangle_e$ which is written in coordinates as

$$\begin{aligned} \bar{v}^T \mathbf{m}_{\bar{x}} \bar{v} &= \langle\langle d\varphi^{-1}(\bar{v}), d\varphi^{-1}(\bar{v}) \rangle\rangle_{\varphi^{-1}(\bar{x})} \\ &= \langle\langle \nu, \nu \rangle\rangle_\chi = \langle\langle T_\chi \mathbf{L}_{\chi^{-1}} \nu, T_\chi \mathbf{L}_{\chi^{-1}} \nu \rangle\rangle_e \\ &= (T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \bar{v})^T \mathbf{m}_{\bar{e}} (T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \bar{v}) \\ &= \bar{u}^T \mathbf{m}_{\bar{e}} \bar{u}, \end{aligned} \tag{4.3.1}$$

where \bar{e} are the coordinates of the identity element and $\bar{u} = T_{\bar{x}} \mathsf{L}_{(\bar{x})^{-1}} \bar{v}$ are the coordinates resulting from left-trivialize the velocity \bar{v} .

By gathering all the above coordinate descriptions, the Riemannian geodesic problem written in coordinates results in the following variational problem:

$$\begin{aligned} \text{extremize}_{\bar{q}(t) \in \Omega \subset \mathbb{R}^k} \quad & \frac{1}{2} \int_0^1 \dot{\bar{q}}_{(t)}^T \mathbf{m}_{\bar{q}} \dot{\bar{q}}_{(t)} \\ \text{subject to} \quad & \bar{q}_{(0)} = \bar{q}_0, \quad \bar{q}_{(1)} = \bar{q}_1 \end{aligned} \tag{varChart}$$

as long as \bar{q}_0, \bar{q}_1 and the whole path $\bar{q}_{(t)}$ belong to Ω . This is a classical variational problem with Lagrangian function $\mathcal{L}(t, \bar{q}, \dot{\bar{q}}) = \frac{1}{2} (\dot{\bar{q}}_{(t)}^T \mathbf{m}_{\bar{q}} \dot{\bar{q}}_{(t)})$.

It is widely known that the extremals of the functional are given by the curves $\bar{q}(t)$ which satisfy the endpoint constraints and the Euler–Lagrange equation [Van Brunt 2004]

$$\partial_{\bar{q}} \mathcal{L}(t, \bar{q}, \dot{\bar{q}}) - d_t \partial_{\dot{\bar{q}}} \mathcal{L}(t, \bar{q}, \dot{\bar{q}}) = 0. \quad (\text{E-L})$$

In matrix notation and using the matrix calculus rules described in [Abadir and Magnus 2005, Chapter 13] [Magnus and Neudecker 1988], it can be shown that

$$\begin{aligned} \partial_{\bar{q}} \mathcal{L}(t, \bar{q}, \dot{\bar{q}}) &= \frac{1}{2} (\dot{\bar{q}}^T \otimes \dot{\bar{q}}^T) (\mathbf{D}\mathbf{m}_{\bar{q}}) \\ \partial_{\dot{\bar{q}}} \mathcal{L}(t, \bar{q}, \dot{\bar{q}}) &= \dot{\bar{q}}^T \mathbf{m}_{\bar{q}} \\ d_t \partial_{\dot{\bar{q}}} \mathcal{L}(t, \bar{q}, \dot{\bar{q}}) &= \ddot{\bar{q}}^T \mathbf{m}_{\bar{q}} + \dot{\bar{q}}^T (\mathbf{D}\mathbf{m}_{\bar{q}})^T (\dot{\bar{q}} \otimes I_k), \end{aligned}$$

where $(\mathbf{D}\mathbf{m}_{\bar{q}})$ is the Fréchet derivative of the matrix $\mathbf{m}_{\bar{q}}$ with respect to \bar{q} as a $k^2 \times k$ matrix¹⁷.

By rearranging terms, it is obtained that the extremality conditions can be written in coordinates as

$$\ddot{\bar{q}} = \mathbf{m}_{\bar{q}}^{-1} \left(\frac{1}{2} (\mathbf{D}\mathbf{m}_{\bar{q}})^T (\dot{\bar{q}} \otimes \dot{\bar{q}}) - (\dot{\bar{q}}^T \otimes I_k) (\mathbf{D}\mathbf{m}_{\bar{q}}) \dot{\bar{q}} \right). \quad (\text{odeChart})$$

The Riemannian exponential function $\text{Exp}_x(\mathcal{V})$ is obtained by solving the ordinary differential equation (ODE) (odeChart) with initial conditions $\bar{q}(0) = \varphi(x)$ and $\dot{\bar{q}}(0) = d\varphi(\mathcal{V})$ up to time $t = 1$, $\text{Exp}_x(\mathcal{V}) = \varphi^{-1}(\bar{q}(1))$.

The expression given in (odeChart) is a matrix expression corresponding to the classical geodesic equation $\ddot{q}^c + \Gamma_{i,j}^c(\bar{q})\dot{q}^i\dot{q}^j = 0$ given in Eq. (4.2.15) using $\Gamma_{i,j}^c$ from Eq. (4.2.16)

$$\Gamma_{i,j}^c(\bar{q}) = \frac{1}{2} \sum_{m=1}^k (\mathbf{m}_{\bar{q}}^{-1})^{m,c} \left(\partial_{q^i} \mathbf{m}_{\bar{q}}^{j,m} + \partial_{q^j} \mathbf{m}_{\bar{q}}^{i,m} - \partial_{q^m} \mathbf{m}_{\bar{q}}^{i,j} \right).$$

The left-invariant Riemannian exponential can be solved in a general way by performing the following steps: (a) choose a chart (\mathcal{D}, φ) ; (b) translate the initial velocity to a convenient element of the chart domain \mathcal{D} by using Eq. (4.2.11); (c) evolve the ODE (odeChart) up to a time t ensuring the solution remains in $\Omega \equiv \varphi^{-1}(\mathcal{D})$; (d) eventually re-translate the solution at time t and continue the evolution up to time $t = 1$.

Note that, all operations written in coordinates, also the matrices $\mathbf{m}_{\bar{q}}$, are uniquely determined once the function φ and $\mathbf{m}_{\bar{e}}$ are chosen.

Example: scale and translation group in 1D in local coordinates

In order to illustrate this method to compute geodesics for a given chart, a simple group of spatial transformations will be used: scaling and translation in the one-dimensional space, $\mathcal{ST}(1)$ group, also known as 1D-affine transformations (Section 2.5.2). The elements in this group can be described with two parameters $h(s, t)$ where $s \in \mathbb{R}^+$ is a scaling factor and $t \in \mathbb{R}$ represents a translation.

An obvious chart for this group is $\varphi(h(s, t)) = \begin{pmatrix} s \\ t \end{pmatrix}$ whose domain \mathcal{D} is the whole group and its image is $\Omega \equiv \mathbb{R}^+ \times \mathbb{R}$. The coordinates of the identity element are $\varphi(e) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Left-translations

¹⁷ For a matrix function of matrix argument $\mathbf{F}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q}$, $\mathbf{DF}(M)$ is the $(pq) \times (mn)$ linear operator fulfilling $(\mathbf{DF}(M))^{i,j} = \partial_{\overline{M}^j} \overline{\mathbf{F}(M)}^i$ such that, $\overline{\mathbf{F}(M + \varepsilon P)} = \overline{\mathbf{F}(M)} + \varepsilon (\mathbf{DF}(M)) \overline{P} + O(\varepsilon)$. Several useful rules to compute derivatives of matrix functions can be found in [Abadir and Magnus 2005, Chapter 13] [Magnus and Neudecker 1988].

are written in coordinates as $\mathsf{L}_{(d)}^{(r)} \left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right) = \left(\begin{smallmatrix} rs \\ rt+d \end{smallmatrix} \right)$ and the inverse as $\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)^{-1} = \left(\begin{smallmatrix} 1/s & 0 \\ -t/s & 1 \end{smallmatrix} \right)$. Regarding velocities, they are identified by $\left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right) = d\varphi(h(\dot{s}, \dot{t}))$ and $T_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} \mathsf{L}_{(d)}^{(r)} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right) = \left(\begin{smallmatrix} r\dot{s} \\ r\dot{t} \end{smallmatrix} \right)$.

To compute the 2×2 matrix $\mathbf{m}_{\bar{q}}$ for $\bar{q} = \left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)$ note that

$$\begin{aligned} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right)^T \mathbf{m}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right) &= \left(T_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} \mathsf{L}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)^{-1}} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right) \right)^T \mathbf{m}_{\bar{e}} \left(T_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} \mathsf{L}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)^{-1}} \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right) \right) \\ &= \left(\begin{smallmatrix} \dot{s}/s \\ \dot{t}/s \end{smallmatrix} \right)^T \mathbf{m}_{\bar{e}} \left(\begin{smallmatrix} \dot{s}/s \\ \dot{t}/s \end{smallmatrix} \right) \end{aligned}$$

and therefore $\mathbf{m}_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} = \frac{1}{s^2} \mathbf{m}_{\bar{e}}$.

Choosing $\mathbf{m}_{\bar{e}} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} \right)$, its derivative is given by $(\mathcal{D}\mathbf{m}_{\bar{q}})_{\left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)} = -\frac{2}{s^3} \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{smallmatrix} \right)$. Replacing in Eq. (odeChart) the following IVP system is obtained for a geodesic:

$$\begin{aligned} \ddot{s}_{(t)} &= \frac{\dot{s}_{(t)}^2 - \dot{t}_{(t)}^2}{s_{(t)}} \\ \ddot{t}_{(t)} &= 2 \frac{\dot{s}_{(t)} \dot{t}_{(t)}}{s_{(t)}} \\ \left(\begin{smallmatrix} s \\ t \end{smallmatrix} \right)_{(0)} &= \varphi(x) \\ \left(\begin{smallmatrix} \dot{s} \\ \dot{t} \end{smallmatrix} \right)_{(0)} &= d\varphi(\mathcal{V}). \end{aligned} \tag{4.3.2}$$

It is worth mentioning that computing geodesics in chart coordinates for a very simple transformation group yielded an IVP for which is not straightforward to find a solution.

The solution of the system (4.3.2) involves a tedious algebraic and analytic work (see for example [Vaccaro 2012]). We will relegate the explicit expression of its solution to the Section 5.1.4. Notice that, there was no need to use a matrix representation of the group.

This method to compute geodesics has two limitations. First, different charts (\mathcal{D}, φ) must be selected for each group, and each choice of the function φ involves different ODEs to solve. This makes difficult to define a general computing method or algorithm valid for any group. Second, the selected chart is usually not valid for the whole group and it is required to test when the solution approaches the boundary of the domain. On the other hand, smart selections of the chart yield particular ODEs for which closed-form solutions are known. Unfortunately, this happens for very few cases.

4.3.2 Solving geodesics in the group exponential chart

Some of the limitations of using the ODE (odeChart) can be circumvented by selecting a generic function φ valid for any group. The group exponential function $\exp(\cdot)$, and the canonical coordinates of the first kind defined in Section 4.2.5, is a possible choice. Elements of a matrix group \mathbf{G} can be expressed by $\exp(\sum_i B_i \mu^i) \in \mathbf{G}$ being $\{B_i\}$ the bases of the algebra \mathbf{g} . Moreover, this choice allows the definition of a domain where the chart is valid. Using this chart, geodesics can be computed for any matrix group, where only the algebra basis $\{B_i\}_i$ and the inertia operator \mathbf{m} must be specified.

For an $n \times n$ matrix group \mathbf{G} of dimension k , an open set of the algebra coordinates can be mapped to an open set of the group via the group exponential function $\varphi^{-1}: \Omega \subset \mathbb{R}^k \rightarrow \mathbf{D} \subset \mathbf{G}: \bar{x} \mapsto X = \exp(\sum_i B_i x^i)$ by specifying proper k matrices of \mathbb{M}_n , where Ω is an open set around 0 and \mathbf{D} is an open set of the group around the identity. The inverse of this mapping defines the group exponential chart denoted by (\mathbf{D}, φ) , where $\varphi: \mathbf{D} \subset \mathbf{G} \rightarrow \Omega: X \mapsto \bar{x} = \mathbf{B}^\dagger \log(X)$. It

was shown in [Rossmann 2002] that an admissible set Ω such that $(\varphi \circ \varphi^{-1})(\bar{x}) = \bar{x}$ is given by $\|\sum_i B_i x^i\|_F < \log(2)$, for $\|\cdot\|_F$ the Frobenius norm.

Using the group exponential chart, the group identity $I \in \mathbf{G}$ is mapped to $0 \in \Omega$. The left-translation operation in chart coordinates is given by

$$\mathsf{L}_{\bar{y}}(\bar{x}) = \varphi(\varphi^{-1}(\bar{y})\varphi^{-1}(\bar{x})) = \mathbf{B}^\dagger \overline{\log(\exp(B_j y^j) \exp(B_i x^i))},$$

where $B_i x^i$ is used as a shorthand for $\sum_i B_i x^i$. In this chart, the inverse operation results in $(\bar{x})^{-1} = -\bar{x}$.

Let γ be a curve in \mathbf{D} and let $\bar{\gamma}_{(t)} = \mathbf{B}^\dagger \overline{\log(\gamma_{(t)})}$ be its coordinates in Ω . The velocities $\bar{v} = \dot{\bar{\gamma}} \in T\Omega$ can be computed from the velocities $V = \dot{\gamma} \in T\mathbf{G}$ by the expression $\bar{v}_{(t)} = \mathbf{B}^\dagger \mathcal{D} \log(\gamma_{(t)}) \overline{V}_{(t)} = d\varphi(V_{(t)})$, where \mathcal{D} is the Fréchet derivative operator. The chart coordinates of the velocity of left translated curve $\xi = \mathbf{L}_Y \gamma$ are

$$\begin{aligned} d\varphi(\xi_{(t)}) &= d\varphi(T_{\gamma_{(t)}} \mathbf{L}_Y V_{(t)}) \\ &= \mathbf{B}^\dagger \mathcal{D} \log(\exp(B_j y^j) \exp(B_i \gamma_{(t)}^i)) (I_n \otimes \exp(B_j y^j)) \mathcal{D} \exp(B_i \gamma_{(t)}^i) \mathbf{B} \dot{\gamma}_{(t)} \\ &= T_{\bar{\gamma}_{(t)}} \mathsf{L}_{\bar{y}}(\dot{\bar{\gamma}}_{(t)}). \end{aligned} \quad (4.3.3)$$

Given a left-invariant inner product defined by the matrix \mathbf{m} in the algebra, the inertia operator in coordinates of the chart $\mathbf{m}_{\bar{x}}$ is defined by, (see Eq. (4.3.1))

$$\dot{x}_{(t)}^T \mathbf{m}_{\bar{x}} \dot{\bar{x}}_{(t)} = (T_{\bar{x}} \mathsf{L}_{(\bar{x})}^{-1} \dot{\bar{x}}_{(t)})^T \mathbf{m}_{\bar{x}} (T_{\bar{x}} \mathsf{L}_{(\bar{x})}^{-1} \dot{\bar{x}}_{(t)}).$$

Note that in this particular chart $\mathbf{m}_{\bar{x}} = \mathbf{m}$ because is where it was defined in Eq. (4.2.9). By using $\bar{\gamma}_{(t)} = \bar{x}$ and $\bar{y} = (\bar{x})^{-1}$ in Eq. (4.3.3), it simplifies to

$$T_{\bar{x}} \mathsf{L}_{(\bar{x})}^{-1} \dot{\bar{x}} = \underbrace{\mathbf{B}^\dagger (I_n \otimes \exp(-B_i x^i)) \mathcal{D} \exp(B_i x^i)}_{\mathbf{t}_{\bar{x}}} \dot{\bar{x}} = \mathbf{t}_{\bar{x}} \dot{\bar{x}},$$

where $\mathbf{t}_{\bar{x}}$ is a $k \times k$ matrix. Therefore, the matrix $\mathbf{m}_{\bar{x}} = \mathbf{t}_{\bar{x}}^T \mathbf{m} \mathbf{t}_{\bar{x}}$.

The geodesic problem written in chart coordinates is solved by a curve $\bar{q}_{(t)}$ satisfying Eq. (odeChart) given in Section 4.3.1. The Fréchet derivative of the metric, $(\mathcal{D}\mathbf{m}_{\bar{q}})$, must be computed, which is given by

$$\mathcal{D}\mathbf{m}_{\bar{q}} = (K_{kk} + I_{k^2}) (I_k \otimes \mathbf{t}_{\bar{q}}^T \mathbf{m}) \mathcal{D} \mathbf{t}_{\bar{q}}, \quad (4.3.4)$$

where K_{mm} is the *commutation matrix* [Abadir and Magnus 2005] defined by $K_{mm} \overline{M} = \overline{M^T}$ for an $m \times m$ matrix M , and

$$\begin{aligned} \mathcal{D} \mathbf{t}_{\bar{q}} &= (\mathbf{B}^T \otimes \mathbf{B}^\dagger) \left((I_{n^3} \otimes \exp(-B_i q^i)) \mathcal{D}^2 \exp(B_i q^i) - \right. \\ &\quad \left. (\mathcal{D} \exp(B_i q^i)^T \otimes I_{n^2}) (K_{nn} \otimes I_{n^2}) (I_n \otimes \overline{I_n} \otimes I_n) \mathcal{D} \exp(-B_i q^i) \right) \mathbf{B}. \end{aligned}$$

To evaluate the expression in Eq. (4.3.4) it is needed to compute the derivatives $\mathcal{D} \exp(M)$ and $\mathcal{D}^2 \exp(M)$. There are different attempts to compute $\mathcal{D} \exp(M)$ [Mathias 1992] [Najfeld and Havel 1995] [Al-Mohy and Higham 2009] [Al-Mohy 2010] [Li et al. 2010]. In our computations we have used the approach given in [Mathias 1996], \mathbf{F} being an analytic matrix function, then

$$\mathbf{F} \left(\begin{pmatrix} M & P \\ 0 & M \end{pmatrix} \right) = \begin{pmatrix} \mathbf{F}(M) & \mathbf{d}_\varepsilon \mathbf{F}(M + \varepsilon P) \\ 0 & \mathbf{F}(M) \end{pmatrix},$$

and, if P is the j -th canonical perturbation, then the vectorization of the upper-right submatrix $d_\varepsilon \mathbf{F}(M + \varepsilon P)$ is the j -th column of the matrix $\mathcal{D}\mathbf{F}(M)$. Additionally [Higham and Relton 2013],

$$\mathbf{F} \left(\begin{array}{c|c|c|c} M & P_1 & P_2 & 0 \\ \hline 0 & M & 0 & P_2 \\ \hline 0 & 0 & M & P_1 \\ \hline 0 & 0 & 0 & M \end{array} \right) = \left(\begin{array}{c|c|c|c} \mathbf{F}(M) & d_\varepsilon \mathbf{F}(M + \varepsilon P_1) & d_\varrho \mathbf{F}(M + \varrho P_2) & d_\varrho d_\varepsilon \mathbf{F}(M + \varepsilon P_1 + \varrho P_2) \\ \hline 0 & \mathbf{F}(M) & 0 & d_\varrho \mathbf{F}(M + \varrho P_2) \\ \hline 0 & 0 & \mathbf{F}(M) & d_\varepsilon \mathbf{F}(M + \varepsilon P_1) \\ \hline 0 & 0 & 0 & \mathbf{F}(M) \end{array} \right).$$

Rearranging the elements of the upper-right block in an $n^4 \times n^2$ matrix for the n^2 possible canonical perturbations P_1 and P_2 it is possible to build the matrix $\mathcal{D}^2\mathbf{F}(M)$ such that, $\mathcal{D}\mathbf{F}(M + \varepsilon P) = \mathcal{D}\mathbf{F}(M) + \varepsilon(\mathcal{D}^2\mathbf{F}(M))\bar{P} + O(\varepsilon)$.

Regarding $\mathcal{D}\log(M)$, using the inverse function theorem it can be obtained

$$\mathcal{D}\log(M) = (\mathcal{D}\exp(\log(M)))^{-1}.$$

Finally, $\text{Exp}_{Q_0}(V_0) = \varphi^{-1}(\bar{q}_{(1)})$ is computed solving the second order ODE defined in Eq. (odeChart) with initial conditions $\bar{q}_{(0)} = \varphi(Q_0)$ and $\dot{\bar{q}}_{(0)} = d\varphi(V_0)$ up to time $t = 1$.

The previous framework gives a generic solution valid for any matrix group and under any left-invariant metric. However this approach requires computation of the derivatives of the matrix exponential function, $\mathcal{D}\exp(\cdot)$ and $\mathcal{D}^2\exp(\cdot)$, and of the matrix logarithm function, $\mathcal{D}\log(\cdot)$, which are computationally rather expensive to evaluate.

4.3.3 Evolution of geodesics described in the algebra

The description of geodesics in local coordinates given above usually involves highly non-linear differential equations, which require numerical solutions with high computational cost. Alternatively, the symmetries of the group and metrics allow to describe the evolution of geodesics from a single tangent space. For convenience the algebra is chosen. Accordingly, the following variational problem is proposed

$$\begin{aligned} \text{extremize}_{Q_{(t)} \in \mathbf{G}} \quad & \mathbf{E}(Q) = \frac{1}{2} \int_0^1 \left\langle \dot{Q}_{(t)}, \dot{Q}_{(t)} \right\rangle_{Q_{(t)}} dt \\ \text{subject to} \quad & Q_{(t=0)} = Q_0 \quad ; \quad Q_{(t=1)} = Q_1 \end{aligned} \quad (\text{varQ})$$

with the corresponding Lagrangian $\mathcal{L}(t, Q, \dot{Q}) = \frac{1}{2} \left\langle \dot{Q}_{(t)}, \dot{Q}_{(t)} \right\rangle_{Q_{(t)}}$. For a smooth family of curves $Q_{(\cdot, \epsilon)} \in \mathcal{C}(\mathbf{G})$, such that $\partial_\epsilon Q_{(t, \epsilon)}|_{(\epsilon=0)} = \delta Q_{(t)} \in T_{Q_{(t)}} \mathbf{G}$ (which can be interpreted as locally $Q_{(\epsilon, t)} \approx Q_{(t)} + \epsilon \delta Q_{(t)}$) and with $\delta Q_{(0)} = \delta Q_{(1)} = 0$, the curve $Q_{(\cdot, 0)}$ is a geodesic if $d_\epsilon \mathbf{E}(Q_{(\cdot, \epsilon)})|_{(\epsilon=0)}$ vanishes.

Unlike the Problem (varChart), the variations δQ must be restricted to keep curves $Q_{(\cdot, \epsilon)}$ in the group \mathbf{G} which, in general, is not a vector space and no classical variational methods can be applied.

Using the left-invariance property given in Eq. (4.2.8), the Lagrangian can be rewritten as $\mathcal{L}(t, Q, \dot{Q}) = \frac{1}{2} \left\langle Q_{(t)}^{-1} \dot{Q}_{(t)}, Q_{(t)}^{-1} \dot{Q}_{(t)} \right\rangle_I$. Let $U_{(t)} = Q_{(t)}^{-1} \dot{Q}_{(t)} \in \mathfrak{g}$ be the left-trivialized velocities of the curve, then, problem (varQ) is equivalent to

$$\begin{aligned} \text{extremize}_{U_{(t)} = Q_{(t)}^{-1} \dot{Q}_{(t)}} \quad & \frac{1}{2} \int_0^1 \left\langle U_{(t)}, U_{(t)} \right\rangle_I dt \\ \text{subject to} \quad & Q_{(t=0)} = Q_0 \quad ; \quad Q_{(t=1)} = Q_1, \end{aligned} \quad (\text{varU})$$

where the Lagrangian is *trivialized* to $\ell(U) = \frac{1}{2}\langle\langle U_{(t)}, U_{(t)} \rangle\rangle_I$, *i.e.* the restriction to the identity and to the algebra $\ell = \mathcal{L}|_{I,\mathfrak{g}} : \mathfrak{g} \rightarrow \mathbb{R}$.

Let $U_{(\cdot,\epsilon)} = Q_{(\cdot,\epsilon)}^{-1} \dot{Q}_{(\cdot,\epsilon)}$ be the family of left-trivialized velocities from curves $Q_{(\cdot,\epsilon)}$. The extremal curves solving (varU) are those fulfilling

$$d_\epsilon \int_0^1 \ell(U_{(\cdot,\epsilon)}) dt = \int \langle \partial_U \ell, \partial_\epsilon U_{(\cdot,\epsilon)} \rangle = 0.$$

It is shown in¹⁸ [Holm et al. 2009] that, $\partial_\epsilon U_{(\cdot,\epsilon)} = \dot{W}_{(\cdot,\epsilon)} + \mathbf{ad}_{U_{(\cdot,\epsilon)}} W_{(\cdot,\epsilon)}$, where $W_{(\cdot,\epsilon)} = Q_{(\cdot,\epsilon)}^{-1} \delta Q_{(\cdot)}$ is a curve in \mathfrak{g} . Therefore,

$$\begin{aligned} \left(d_\epsilon \int_0^1 \ell(U_{(\cdot,\epsilon)}) \right) \Big|_{(\epsilon=0)} &= \left(\int \langle \partial_U \ell, \dot{W}_{(\cdot,\epsilon)} + \mathbf{ad}_{U_{(\cdot,\epsilon)}} W_{(\cdot,\epsilon)} \rangle \right) \Big|_{(\epsilon=0)} \\ &\stackrel{(a)}{=} \int \left\langle -\frac{d}{dt}(\partial_U \ell), W \right\rangle + \int \left\langle \mathbf{ad}_U^*(\partial_U \ell), W \right\rangle \\ &= \int \left\langle -\frac{d}{dt}(\partial_U \ell) + \mathbf{ad}_U^*(\partial_U \ell), W \right\rangle. \end{aligned}$$

The first term in (a) was obtained by integrating by parts and using that $W_{(0)} = W_{(1)} = 0$. For the second term, notice that $\mathbf{ad}_U : \mathfrak{g} \rightarrow \mathfrak{g}$ is a linear operator and any linear operator $\mathbf{f} : \mathbb{V} \rightarrow \mathbb{V}$ has an associated *dual* operator $\mathbf{f}^* : \mathbb{V}^* \rightarrow \mathbb{V}^*$ such that $\langle \mathbf{w}, \mathbf{f}(u) \rangle = \langle \mathbf{f}^*(\mathbf{w}), u \rangle$. Finally, $\partial_U \ell = \mathcal{M}(U) \in \mathfrak{g}^*$ and since the above variations must vanish for any W , it follows that extremal curves of (varU) satisfy the *Euler–Poincaré equation*¹⁹

$$\frac{d}{dt} \mathcal{M}(U_{(t)}) = \mathbf{ad}_{U_{(t)}}^* (\mathcal{M}(U_{(t)})). \quad (\text{E–P})$$

It is also common to find [Marsden 1992] a Hamiltonian version of the (E–P) equation where instead to parameterize the dynamical system in terms of generalized variables q and \dot{q} , it is parameterized in terms of conjugated variables. In terms of the *momentum variable* $\mathbf{m} = \partial_U \ell(U)$, the Euler–Poincaré equation takes the form

$$\dot{\mathbf{m}}_{(t)} = \mathbf{ad}_{\mathcal{M}^{-1}(\mathbf{m}_{(t)})}^* \mathbf{m}_{(t)}.$$

In the literature is sometimes also found [Modin et al. 2011] [Hinkle et al. 2013]

$$\dot{U}_{(t)} = \mathbf{ad}_{U_{(t)}}^\top U_{(t)},$$

where $\mathbf{ad}_{(\cdot)}^\top = (\mathcal{M}^{-1} \circ \mathbf{ad}_{(\cdot)}^* \circ \mathcal{M})$ is the transpose of \mathbf{ad} with respect to the inner product (see footnote 15 on page 91).

The Euler–Poincaré equation (E–P) together with the *reconstruction equation* $\dot{Q}_{(t)} = Q_{(t)} U_{(t)}$ define the evolution of a geodesic.

The major benefit of (E–P) is that the geodesic evolution is described in a single vector space, avoiding the non-linearities in the metric when using a coordinate chart.

Recalling Section 4.2.3, the velocity $U_{(t)} \in \mathfrak{g}$ can be written as $\sum_i B_i \mu^i$. The corresponding $\mathbf{w} = \mathcal{M}(U) \in \mathfrak{g}^*$ can be written as $\sum_j \mathfrak{D}_j \omega^j = \sum_j \mathfrak{D}_j (\mathbf{m} \bar{\mu})^j$, where $\bar{\omega} = \mathbf{m} \bar{\mu}$ are the co-algebra coordinates of \mathbf{w} and $\{\mathfrak{D}_j\}$ is a basis of \mathfrak{g}^* such that $\text{span}(\{\mathfrak{D}_1, \mathfrak{D}_2, \dots, \mathfrak{D}_k\}) \equiv \mathfrak{g}^*$ and $\langle \mathfrak{D}_j, B_i \rangle = \delta_{ji}$. A schematic illustration of a curve in the group \mathbf{G} and many involved concepts used for describing the evolution of a geodesic are given in Figure 4.3.1.

¹⁸ For a formal proof valid for any abstract group without the particularization to be a matrix group, see [Bloch et al. 1996, Appendix].

¹⁹ In some textbooks it is called *basic Euler–Poincaré equation* to strengthen that it only includes the *kinetic energy of a free-body*.

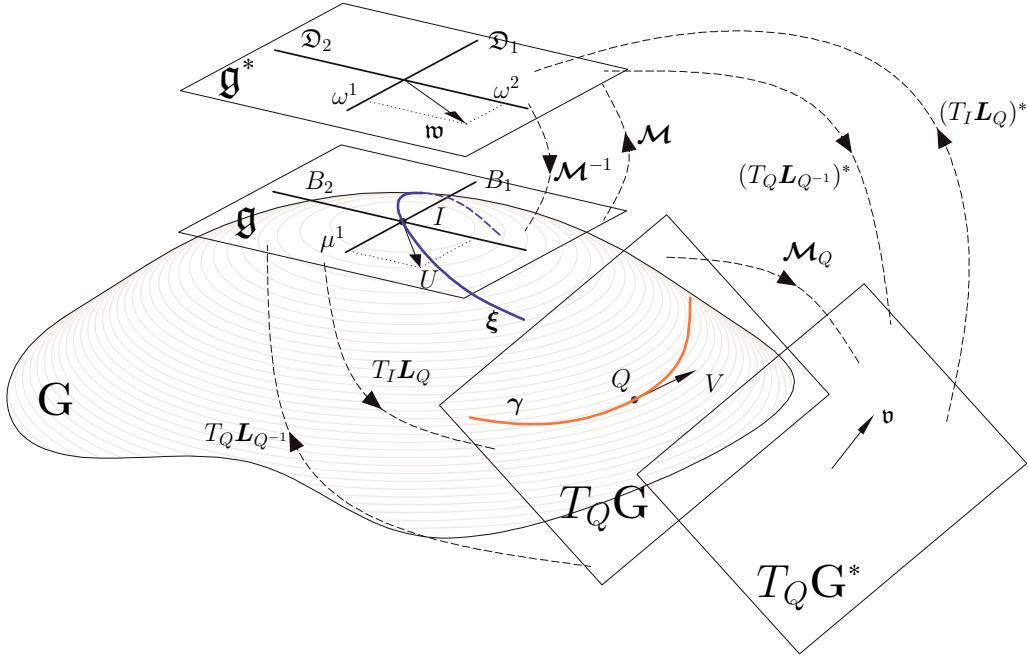


Figure 4.3.1. Schematic illustration of a curve in the group \mathbf{G} and many involved concepts used for describing the evolution of a geodesic. The group \mathbf{G} , the algebra \mathfrak{g} as the tangent space at I and its dual space \mathfrak{g}^* are shown. The algebra \mathfrak{g} has basis $\{B_1, B_2, \dots\}$ while \mathfrak{g}^* has dual basis $\{\mathfrak{D}_1, \mathfrak{D}_2, \dots\}$ such that $\langle \mathfrak{D}_j, B_i \rangle = \delta_{ji}$. A curve on the group passing through $Q \in \mathbf{G}$ with velocity $V \in T_Q \mathbf{G}$ is denoted by γ . Translating the curve with $\mathbf{L}_{Q^{-1}}$ the curve ξ is obtained. It passes through I with left-trivialized velocity $U = T_Q \mathbf{L}_{Q^{-1}} V \in \mathfrak{g}$ having algebra coordinates $\bar{\mu} = (\mu^1, \mu^2, \dots)^T$. This vector can be identified with $\mathfrak{w} = \mathcal{M}(U) \in \mathfrak{g}^*$ with co-algebra coordinates $(\omega^1, \omega^2, \dots)^T = \mathbf{m}\bar{\mu}$. Conversely $U = \mathcal{M}^{-1}(\mathfrak{w})$. The inner product $\langle\langle U, U \rangle\rangle_I = \langle \mathfrak{w}, U \rangle$ defines a Riemannian metric which computes the *instantaneous kinetic energy* at \mathfrak{g} . By the left-invariance requirement of the Riemannian metric $\langle\langle U, U \rangle\rangle_I = \langle\langle V, V \rangle\rangle_Q$ and therefore $V \in T_Q \mathbf{G}$ is identified with $\mathfrak{v} = \mathcal{M}_Q(V) \in T_Q^* \mathbf{G}$ such that $\langle\langle V, V \rangle\rangle_Q = \langle \mathfrak{v}, V \rangle_{(T_Q \mathbf{G})}$ gives the same kinetic energy as $\langle\langle U, U \rangle\rangle_I$.

In terms of the bases, both the primal adjoint $\mathbf{ad}_{(\cdot)}(\cdot)$ and dual adjoint $\mathbf{ad}_{(\cdot)}^*(\cdot)$ operators are described by the *structure constants* C_{ijm} (see Section 4.2.2 and [Sattinger and Weaver 1986] [Cendra et al. 1998])

$$\mathbf{ad}_{B_i} B_j = \sum_{m=1}^k B_m C_{ijm} \quad \text{and therefore} \quad \mathbf{ad}_{B_i}^* \mathfrak{D}_j = \sum_{m=1}^k \mathfrak{D}_m C_{imj} .$$

These structure constants C_{ijm} can be computed in the primal space using the matrix representations of the bases of the algebra. Recalling that $\mathbf{ad}_{B_i} B_j = [B_i, B_j]$ where $[\cdot, \cdot]$ is the matrix commutator, the elements C_{ijm} can be calculated as

$$C_{ijm} = (\mathbf{B}^\dagger (\overline{B_i B_j} - \overline{B_j B_i}))^m .$$

Using the bi-linearity property of $\mathbf{ad}_{(\cdot)}^*(\cdot)$,

$$\mathbf{ad}_U^* \mathfrak{w} = \sum_{m=1}^k \mathfrak{D}_m \left(\sum_{i,j}^k C_{imj} \mu^i \omega^j \right) = \sum_{m=1}^k \mathfrak{D}_m \nu^j ,$$

being $\bar{\nu} = (\nu^1, \nu^2, \dots, \nu^k)^T$ the co-algebra coordinates of $\text{ad}_U^* \mathfrak{w}$. The ad^* operator in terms of the algebra and co-algebra coordinates will be denoted as $\text{ad}_{\bar{\mu}}^* \bar{\omega} = \bar{\nu}$.

The problem of computing the geodesic $Q_{(t)}$ passing through X with velocity V at $t = 0$ is resolved first by obtaining the curve $\bar{\mu}_{(t)} \in \mathbb{R}^k$ given by the ODE

$$\begin{aligned}\dot{\bar{\mu}}_{(t)} &= \mathbf{m}^{-1}(\text{ad}_{\bar{\mu}_{(t)}}^*(\mathbf{m}\bar{\mu}_{(t)})) \\ \text{with initial condition } \bar{\mu}_{(0)} &= \mathbf{B}^\dagger(\overline{X^{-1}V}),\end{aligned}\tag{4.3.5}$$

and afterwards by solving the reconstruction equation given by the ODE

$$\begin{aligned}\dot{Q}_{(t)} &= Q \left(\sum_{i=1}^k B_i \mu_{(t)}^i \right) \\ \text{with initial condition } Q_{(0)} &= X.\end{aligned}\tag{recons}$$

Additional numerical details to solve the reconstruction step in Eq. (recons) will be given in Section 5.2. Once the geodesic is solved, the corresponding Riemannian exponential is $\text{Exp}_X(V) = Q_{(t=1)}$. An illustration of the geodesic evolution, its symmetries and its relation with the Riemannian exponential function is shown in Figure 4.3.2.

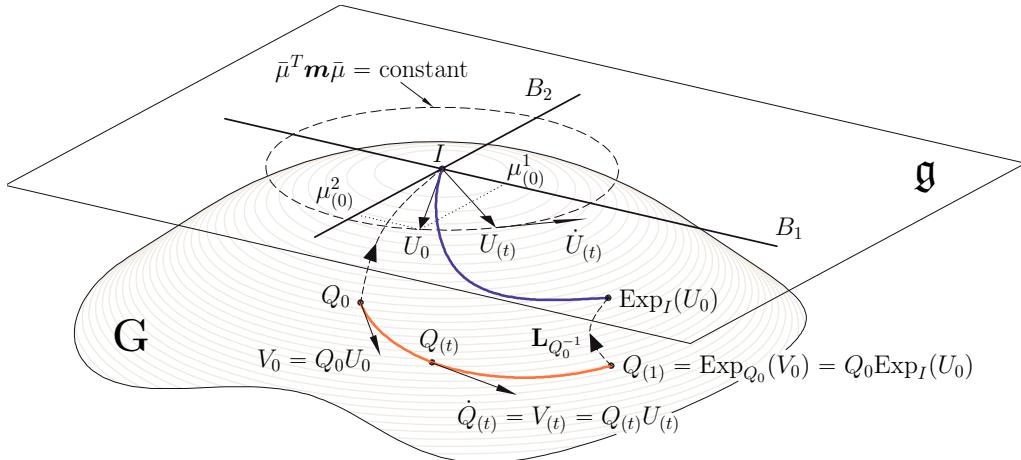


Figure 4.3.2. Illustration of the geodesic evolution. The *red* geodesic segment $Q_{(t)}$ in the group \mathbf{G} departs from the point Q_0 with initial velocity $V_0 = Q_0 U_0$. Every instantaneous velocity $V_{(t)} = \dot{Q}_{(t)}$ can be left-trivialized, providing a curve of algebra vectors $U_{(t)} = Q_{(t)}^{-1} V_{(t)} \in \mathfrak{g}$ with coordinates $\bar{\mu}_{(t)} = (\mu_{(t)}^1, \mu_{(t)}^2, \dots)^T$ in the algebra bases $\{B_1, B_2, \dots\}$. The *red* geodesic segment can be left-translated by Q_0^{-1} resulting in the *blue* geodesic segment from I to $\text{Exp}_I(U_0)$. Both segments have the same length.

4.3.4 A coordinate-free solution based on (E–P)

While the elements of the algebra have been numerically represented by $n \times n$ matrices (embedding \mathfrak{g} in \mathbb{M}_n), their corresponding dual elements (or co-vectors) living in \mathfrak{g}^* were considered so far as abstract entities. In this section a numerical representation of co-algebra elements as matrices will provide a coordinate-free expression for (E–P).

In a finite dimensional vector space, it is common to interpret primal elements as *column-vectors* and the co-vectors as *row-vectors* such that the pairing operation results in the ordinary matrix multiplication. In the case we are concerned with, a primal element $U \in \mathfrak{g}$ is numerically

represented by a matrix of \mathbb{M}_n which is obtained by reshaping the array resulting from $\mathbf{B}\bar{\mu}$ (see Section 4.2.3). Elements of the co-algebra, $\mathbf{w} \in \mathbf{g}^*$, are naturally represented by matrices of the form $\sum_j D_j \omega^j$, where $\bar{\omega} = (\omega^1, \omega^2, \dots, \omega^k)^T$ are the coordinates of \mathbf{w} and D_j are matrix representations of the dual basis \mathcal{D}_j . The matrices D_j must satisfy that the pairing $\langle D_j, B_i \rangle = \delta_{ij}$. In general, there are many possible ways to both, choose a matrix representation of D_j and make explicit the pairing operation. A convenient way is by reshaping the rows of a matrix \mathbf{B}^\dagger , any generalized left inverse of \mathbf{B} . Analogously to the primal case, $\mathbf{D} = \mathbf{B}^{\dagger T} = (\overline{D}_1, \overline{D}_2, \dots, \overline{D}_k)$ is the matrix built from the vectorization of the dual bases D_j . Then, a co-vector $\mathbf{w} \in \mathbf{g}^*$ with coordinates $\bar{\omega}$ is numerically represented by a matrix $W \in \mathbb{M}_n$ such that $\overline{W} = \mathbf{B}^{\dagger T} \bar{\omega}$. In addition, under this representation, the pairing operation $\langle \cdot, \cdot \rangle : \mathbf{g}^* \times \mathbf{g} \rightarrow \mathbb{R}$ takes the form

$$\langle W, U \rangle = \sum_m \mu^m \omega^m = (\mathbf{D}^\dagger \overline{W})^T (\mathbf{B}^\dagger \overline{U}) = \text{trace}(W^T U),$$

where $\bar{\mu}$ and $\bar{\omega}$ are the coordinates of U and W , and \mathbf{D}^\dagger and \mathbf{B}^\dagger are any generalized left inverse of \mathbf{D} and \mathbf{B} , respectively.

The first step towards a coordinate-free expression for (E–P) is to write the \mathbf{ad}^* operator in terms of the matrix representations defined above. As $\mathbf{ad}_Z(\cdot) = [Z, \cdot]$ is a linear operator on $\mathbb{M}_n \equiv \mathbb{R}^{n^2}$, it can be represented by a matrix \mathbf{A}_Z of \mathbb{M}_{n^2} (see Eq. (4.2.4) in Section 4.2.4) such that, $\mathbf{ad}_Z \overline{U} = \mathbf{A}_Z \overline{U}$, with $\mathbf{A}_Z = (I_n \otimes Z) - (Z^T \otimes I_n)$ for $Z, U \in \mathbf{g}$. The pairing $\langle W, \mathbf{ad}_Z U \rangle$ can be written using the matrix representations as $(\mathbf{D}^\dagger \overline{W})^T (\mathbf{B}^\dagger \overline{\mathbf{ad}_Z U})$. Using $\overline{U} = \mathbf{B} \mathbf{B}^\dagger \overline{U}$ (which is fulfilled only when $U \in \mathbf{g}$) and $\mathbf{D}^T \mathbf{D}^{\dagger T} = I_k$ follows

$$\begin{aligned} \langle W, \mathbf{ad}_Z U \rangle &= (\mathbf{D}^\dagger \overline{W})^T (\mathbf{B}^\dagger \overline{\mathbf{ad}_Z U}) \\ &= \overline{W}^T \mathbf{D}^{\dagger T} \mathbf{B}^\dagger \mathbf{A}_Z \overline{U} \\ &= \overline{W}^T \mathbf{D}^{\dagger T} \mathbf{B}^\dagger \mathbf{A}_Z (\mathbf{B} \mathbf{B}^\dagger) \overline{U} \\ &= \overline{W}^T \mathbf{D}^{\dagger T} \mathbf{B}^\dagger \mathbf{A}_Z \mathbf{B} (\mathbf{D}^T \mathbf{D}^{\dagger T}) \mathbf{B}^\dagger \overline{U} \\ &= (\mathbf{D}^\dagger (\mathbf{D} \mathbf{B}^T \mathbf{A}_Z^T \mathbf{B}^{\dagger T} \mathbf{D}^\dagger) \overline{W})^T (\mathbf{B}^\dagger \overline{U}) \\ &= (\mathbf{D}^\dagger \overline{\mathbf{ad}_Z^* W})^T (\mathbf{B}^\dagger \overline{U}) \\ &= \langle \mathbf{ad}_Z^* W, U \rangle. \end{aligned}$$

Therefore, \mathbf{ad}_Z^* is identified with the matrix operation given by $\overline{\mathbf{ad}_Z^* W} = \mathbf{D} \mathbf{B}^T \mathbf{A}_Z^T \mathbf{D} \mathbf{D}^\dagger \overline{W}$. In order to get a more compact expression, from now on $(\cdot)^\dagger$ will be chosen as the Moore–Penrose pseudo-inverse, then $\mathbf{B}^T = \mathbf{D}^\dagger$.

It is easy to show that for any $Z \in \mathbb{M}_n$, $\mathbf{A}_Z^T = \mathbf{A}_{Z^T}$ holds, i.e. \mathbf{A}_Z^T is the linear operator associated to the matrix commutation $[Z^T, \cdot]$, therefore

$$\overline{\mathbf{ad}_Z^* W} = \mathbf{D} \mathbf{D}^\dagger \mathbf{A}_{Z^T} \overline{W} \Rightarrow \mathbf{ad}_Z^* W = \pi[Z^T, W],$$

where $\pi : \mathbb{M}_n \rightarrow \mathbf{g}^* : \overline{\pi(X)} = \mathbf{D} \mathbf{D}^\dagger \overline{X} = \mathbf{B}^{\dagger T} \mathbf{B}^T \overline{X}$ is a projector ($\pi \circ \pi = \pi$), Z and W are matrix representations of elements of \mathbf{g} and \mathbf{g}^* , respectively, and $\mathbf{D} \mathbf{D}^\dagger \overline{W} = \overline{W}$ for $W \in \mathbf{g}^*$. In this expression, $[\cdot, \cdot]$ is the matrix commutation and should not be confused with the Lie bracket. The matrix commutation coincides with the Lie bracket only when its both operands belong to the algebra. In general, the resulting matrix $[Z^T, W]$ does not represent elements belonging to either \mathbf{g} or \mathbf{g}^* . Therefore, the projector π is mandatory in the expression of the \mathbf{ad}^* operator. In [Crouch et al. 2010] [Bloch et al. 2011], the authors obtained the same coordinate-free expression but by endowing to the space \mathbb{M}_n with the Frobenius inner product $\langle\langle A, B \rangle\rangle_F = \text{trace}(A^T B)$ (also called *Hilbert–Schmidt inner product*).

The second step is to consider the inertia operator \mathcal{M} . Using the previous numerical representations, the identification between \mathbf{g} and \mathbf{g}^* carried out by the operation $W = \mathcal{M}(U)$ is written as $\overline{W} = (\mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger) \overline{U}$. The inverse, $\mathcal{M}^{-1}: \mathbf{g}^* \rightarrow \mathbf{g}: W \mapsto U = \mathcal{M}^{-1}(W)$ can be expressed by $\overline{U} = (\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T) \overline{W}$. To alleviate algebraic expressions, it is useful to define the following linear operators²⁰

$$\begin{aligned}\chi: \mathbb{M}_n &\rightarrow \mathbb{M}_n & \text{such that} & \quad \overline{\chi(X)} = (\mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger) \overline{X} \\ \sigma: \mathbb{M}_n &\rightarrow \mathbb{M}_n & \text{such that} & \quad \overline{\sigma(X)} = (\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T) \overline{X},\end{aligned}\tag{4.3.6}$$

such that $\mathcal{M}(U) = \chi U$ and $\mathcal{M}^{-1}(W) = \sigma W$. It is straightforward to verify $\sigma(\chi(U)) = U$ for $U = \sum B_i \mu^i \in \mathbf{g}$. Moreover, $\chi(\sigma(W)) = W$ for $W = \sum D_i \omega^i \in \mathbf{g}^*$.

Recalling the Euler–Poincaré equation (E–P) and considering the matrix representations, it is obtained

$$\begin{aligned}\frac{d}{dt}(\overline{\mathcal{M}(U)}) &= \overline{\text{ad}_U^* \mathcal{M}(U)} \\ &= \overline{\mathbf{B}^{\dagger T} \mathbf{B}^T \mathbf{A}_U^T \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger \overline{U}} \\ \Rightarrow \dot{\overline{U}} = \mathcal{M}^{-1} \left(\overline{\text{ad}_U^* \mathcal{M}(U)} \right) &= \overline{\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T \mathbf{B}^{\dagger T} \mathbf{B}^T \mathbf{A}_U^T \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger \overline{U}} \\ &= (\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T) \mathbf{A}_{U^T} (\mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger) \overline{U}.\end{aligned}$$

Then the equation of the velocity guiding a geodesic can be compactly written as

$$\dot{U}_{(t)} = \sigma[U_{(t)}^T, \chi U_{(t)}],\tag{4.3.7}$$

which together with the reconstruction step given in (recons) provide a coordinate-free expression for left-invariant geodesics. Equivalent expressions can be found in [Crouch et al. 2010] [Bloch et al. 2011] [Nordkvist et al. 2011].

Note that in the case of Rotation groups $\mathbf{SO}(d)$ (where the elements of the algebra are skew-symmetric matrices, see later Section 5.1.5), it is obtained that $\chi \dot{U} = [\chi U, U]$ which is well known as the *generalized rigid body equation* [Bloch et al. 2002].

When $U_{(t)}$ remains constant, *i.e.* $\sigma[U_0^T, \chi U_0] = 0$, the corresponding geodesic is a translated one-parameter subgroup. Note that for some particular initial velocities, $\text{Exp}_I(t U_0)$ may coincide with $\exp(t U_0)$ even for a non bi-invariant metric. These cases are described for the particular case of Rotation group in [Arnold 1989, Appendix 2.E] and were called *stationary rotations*. Moreover, the reverse can be proven, if $\text{Exp}_I(t U_0)$ coincides with $\exp(t U_0)$ then $\sigma[U_{(t)}^T, \chi U_{(t)}]$ vanishes for all t . In [Pripoae 2006] it was shown that for a given $U_0 \in \mathbf{g}$ there always exists a left-invariant Riemannian metric such that the group exponential $\exp(t U_0)$ coincides with its Riemannian exponential $\text{Exp}_I(t U_0)$.

In general, Eq. (4.3.7) is a second-order polynomial ODE and no closed-form solution is known except for a few cases. An interesting particular case of Eq. (4.3.7) occurs when $\sigma = \chi = \text{id}_{\mathbb{M}_n} = I_{n^2}$, and the solution is given in closed-form by [Andruchow et al. 2011] [Vandereycken et al. 2013]

$$U_{(t)} = \exp(-t(U_0 - U_0^T)) U_0 \exp(t(U_0 - U_0^T)).$$

²⁰ For brevity reasons, when there is no possible misunderstanding, $\chi(X)$ and $\sigma(X)$ will be written as χX and σX , respectively.

4.3.5 Left-invariant geodesics on matrix groups using Optimal Control

The computation of left-invariant geodesics and the Riemannian exponential function on matrix groups can be formulated as the solution of an Optimal Control problem. Optimal Control also provides a general framework valid for any matrix group and is developed for any left-invariant Riemannian metric. The added value of this approach is that the co-state variable will allow to find conserved quantities (see Section 4.4).

Optimal Control is a technique to minimize an integral cost functional that depends on the evolution of the states of a dynamical system [Jurdjevic 1996, Chapters 11-12] [Sachkov 2009] [Lewis et al. 2012]. In a first stage, the dynamical system and the functional are described in terms of *control variables*. Then, the optimal time-behavior of the control variables is solved and later the dynamical system variables are obtained.

Recalling the variational problem (varQ), the curve $Q_{(t)}$ can be considered as a dynamical system governed or guided by a *control variable* $V_{(t)} = \dot{Q}_{(t)}$, the velocity of the curve. Accordingly, (varQ) is reformulated as a control problem

$$\begin{aligned} & \text{extremize}_{V_{(t)} \in T_{Q(t)}\mathbf{G}} \quad \frac{1}{2} \int_0^1 \langle\!\langle V_{(t)}, V_{(t)} \rangle\!\rangle_{Q_{(t)}} dt \\ & \text{subject to} \quad \dot{Q}_{(t)} = V_{(t)} \\ & \quad Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1. \end{aligned} \tag{ocpV}$$

For a left-invariant metric Eq. (4.2.8) holds, $\langle\!\langle V_1, V_2 \rangle\!\rangle_Q = \langle\!\langle U_1, U_2 \rangle\!\rangle_I$, where $U_i = Q^{-1}V_i$ are left-trivialized velocities. Then, (ocpV) can be written as

$$\begin{aligned} & \text{extremize}_{U_{(t)} \in \mathfrak{g}} \quad \frac{1}{2} \int_0^1 \langle\!\langle U_{(t)}, U_{(t)} \rangle\!\rangle_I dt \\ & \text{subject to} \quad \dot{Q}_{(t)} = Q_{(t)}U_{(t)} \\ & \quad Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1. \end{aligned} \tag{ocpU}$$

Three different constraints can be identified in (ocpU): the *fixed bound constraints* specify the initial and final elements, Q_0 and Q_1 , respectively; the *dynamic law* $\dot{Q}_{(t)} = Q_{(t)}U_{(t)}$ constrains the evolution of the dynamical system; finally, the *control variable domain constraint* $U_{(t)} \in \mathfrak{g}$ guarantees that $Q_{(t)} \in \mathbf{G}$ for all t , since $Q_{(t)}U_{(t)} \in T_{Q(t)}\mathbf{G}$.

Writing velocities in terms of the algebra basis, $\bar{U}_{(t)} = \mathbf{B}\bar{\mu}_{(t)}$, and using Eq. (4.2.9) to express the instantaneous energy, the Optimal Control problem can be reformulated without constraints on the control variable:

$$\begin{aligned} & \text{extremize}_{\bar{\mu}_{(t)} \in \mathbb{R}^k} \quad \frac{1}{2} \int_0^1 \bar{\mu}_{(t)}^T \mathbf{m} \bar{\mu}_{(t)} dt = \mathbf{E}(\bar{\mu}) \\ & \text{subject to} \quad \dot{Q}_{(t)} = Q_{(t)} \sum B_i \mu_{(t)}^i \\ & \quad Q_{(0)} = Q_0 \quad ; \quad Q_{(1)} = Q_1. \end{aligned} \tag{ocp\bar{\mu}}$$

Optimal Control solves (ocp\bar{\mu}) by including the dynamic constraint in an *augmented functional* with the corresponding Lagrange multiplier. This Lagrange multiplier enforces the constraint $\dot{Q}_{(t)} - Q_{(t)}U_{(t)} = 0$, which can equivalently be written as $\langle\!\langle \mathfrak{p}, \dot{Q} - QU \rangle\!\rangle_Q = 0$ for all $\mathfrak{p} \in T_Q^*\mathbf{G}$, where $\langle\cdot, \cdot\rangle_Q$ is the pairing in the vector space $T_Q\mathbf{G}$. Again, matrix representations are used to

express the pairing $\langle \mathbf{p}, V \rangle_Q$ as²¹ $\text{trace}(P^T V)$. Accordingly, $(\text{ocp} \bar{\mu})$ is equivalent to extremizing

$$\begin{aligned}\mathbf{E}'(\bar{\mu}) &= \int_0^1 \left(\frac{1}{2} \bar{\mu}_{(t)}^T \mathbf{m} \bar{\mu}_{(t)} + \text{trace} \left(P_{(t)}^T \left(\dot{Q}_{(t)} - Q_{(t)} \sum_i^k B_i \mu_{(t)}^i \right) \right) \right) dt \\ &= \int_0^1 \left(\mathcal{H}(Q, \bar{\mu}, P, t) + \text{trace}(P_{(t)}^T \dot{Q}_{(t)}) \right) dt,\end{aligned}$$

where

$$\mathcal{H}(Q, \bar{\mu}, P, t) = \frac{1}{2} \bar{\mu}_{(t)}^T \mathbf{m} \bar{\mu}_{(t)} - \text{trace} \left(P_{(t)}^T Q_{(t)} \sum_i^k B_i \mu_{(t)}^i \right). \quad (4.3.8)$$

Following [Lewis et al. 2012], variations $\delta Q_{(t)}$, $\delta \bar{\mu}_{(t)}$ and $\delta P_{(t)}$ on the variables Q , $\bar{\mu}$ and P respectively, generate a perturbation of the augmented functional \mathbf{E}'

$$\begin{aligned}d\mathbf{E}' &= \int \left\langle \partial_Q \mathcal{H}, \delta Q \right\rangle + \left\langle P, \delta \dot{Q} \right\rangle + \left\langle \partial_{\bar{\mu}} \mathcal{H}, \delta \bar{\mu} \right\rangle + \left\langle \partial_P \mathcal{H}, \delta P \right\rangle + \left\langle \dot{Q}, \delta P \right\rangle \\ &= \int \left\langle \partial_Q \mathcal{H} - \dot{P}, \delta Q \right\rangle + \left\langle \partial_{\bar{\mu}} \mathcal{H}, \delta \bar{\mu} \right\rangle + \left\langle \partial_P \mathcal{H} + \dot{Q}, \delta P \right\rangle,\end{aligned}$$

where integration by parts was used to eliminate the term $\delta \dot{Q}$.

The extremality conditions are obtained by setting to zeros the coefficients of the independent perturbations

$$\partial_{\bar{\mu}} \mathcal{H} = \mathbf{m} \bar{\mu}_{(t)} - \mathbf{B}^T (\overline{Q_{(t)}^T P_{(t)}}) = 0 \quad (4.3.9a)$$

$$-\partial_P \mathcal{H} = Q_{(t)} \sum_i^k B_i \mu_{(t)}^i = \dot{Q}_{(t)} \quad (4.3.9b)$$

$$-\partial_Q \mathcal{H} = P_{(t)} \sum_i^k B_i^T \mu_{(t)}^i = -\dot{P}_{(t)} \quad (4.3.9c)$$

Equation (4.3.9a) relates the *costate variables* $P_{(t)}$ and the *control variables* $\bar{\mu}_{(t)}$ via an algebraic relation with the *state variables* $Q_{(t)}$. This equation can be rewritten in terms of the matrix representation as

$$\overline{U}_{(t)} = \mathbf{B} \bar{\mu}_{(t)} = (\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T) (\overline{Q_{(t)}^T P_{(t)}}) \Rightarrow U_{(t)} = \boldsymbol{\sigma}(Q_{(t)}^T P_{(t)}), \quad (4.3.10)$$

where $\boldsymbol{\sigma}(\cdot)$ was defined in Eq. (4.3.6).

The Riemannian geodesic starting at Q_0 with velocity V satisfy the IVP defined by

$$U_{(t)} = \boldsymbol{\sigma}(Q_{(t)}^T P_{(t)}) \quad (4.3.11a)$$

$$\dot{Q}_{(t)} = Q_{(t)} U_{(t)} \quad (4.3.11b)$$

$$\dot{P}_{(t)} = -P_{(t)} U_{(t)}^T \quad (4.3.11c)$$

$$Q_{(t=0)} = Q_0. \quad (4.3.11d)$$

While the initial condition for Q is specified, $P_{(0)}$ remains unspecified. However, $U_{(0)} = Q_0^{-1} V$ is given by the specifications of the geodesic. In addition, $P_{(0)}$ must fulfil $U_{(0)} = \boldsymbol{\sigma}(Q_0^T P_{(0)})$. It can be shown²² that $P_{(0)} = Q_0^{-T} \boldsymbol{\chi} U_{(0)}$.

²¹ Let $\{B_1, B_2, \dots\}$ be a set of bases a matrix representation of \mathbf{g} , then for a group element Q , the set $\{QB_1, QB_2, \dots\}$ can be used as a set of matrix bases for $T_Q \mathbf{G}$. Moreover, for $\{D_1, D_2, \dots\}$ a set of matrix bases for \mathbf{g}^* (see Section 4.3.4), then $\{Q^{-T} D_1, Q^{-T} D_2, \dots\}$ can be used as matrix bases for $T_Q^* \mathbf{G}$ which satisfies $\langle Q^{-T} D_j, QB_i \rangle_Q = \delta_{ji}$. Under these representations $P = \sum_j Q^{-T} D_j \rho^j \in T_Q^* \mathbf{G}$, $V = \sum_i QB_i \nu^i \in T_Q \mathbf{G}$, and $\langle P, V \rangle_Q = \sum \rho^m \nu^m = \text{trace}(P^T V)$.

²² For a matrix P representing an element of $T_x^* \mathcal{G}$, the cotangent lift $(T_x \mathbf{L}_q)^* : T_x^* \mathcal{G} \rightarrow T_x^* \mathcal{G}$ is represented by $Q^T P$. Therefore $Q^T P \in \mathbf{g}^*$ and applying from the left the operator $\boldsymbol{\chi}$ defined in Eq. (4.3.6) it is obtained $Q^T P = \boldsymbol{\chi} U$.

Note that this IVP involves two coupled evolutions ($Q_{(t)}$ and $P_{(t)}$), while in the methodologies described in Section 4.3.3 and Section 4.3.4, the evolution of $U_{(t)}$ is independent of $Q_{(t)}$.

4.3.6 Symmetric representation of the geodesics evolution

Some authors have proposed a symmetric representation of the geodesic equations for particular matrix groups and more recently for a general matrix case. A symmetric formulation is given below starting from IVP (4.3.11). $P_{(t)}$ can be written as $P_{(t)} = P_0 Y_{(t)}$, with $Y_{(0)} = I$, then from Eq. (4.3.11c) $\dot{P}_{(t)} = P_0 \dot{Y}_{(t)} = -P_0 Y_{(t)} U_{(t)}^T$. Although P_0 is not necessarily invertible, the previous equation is fulfilled if $\dot{Y}_{(t)} = Y_{(t)} (-U_{(t)}^T)$. Therefore, the IVP (4.3.11) can be rewritten in terms of $Y_{(t)}$ as

$$U_{(t)} = \sigma(Q_{(t)}^T P_0 Y_{(t)}) \quad (4.3.12a)$$

$$\dot{Q}_{(t)} = Q_{(t)} U_{(t)} \quad (4.3.12b)$$

$$\dot{Y}_{(t)} = Y_{(t)} (-U_{(t)}^T) \quad (4.3.12c)$$

$$Q_{(t=0)} = Q_0 \quad (4.3.12d)$$

$$Y_{(t=0)} = I \quad (4.3.12e)$$

$$P_0 = Q_0^{-T} \chi U_0. \quad (4.3.12f)$$

It can be observed that $Y_{(t)}$ starts at the identity and follows a left-invariant field defined on the transposed algebra. Therefore, $Y_{(t)}$ belongs to the *transposed group*²³ $\mathbf{G}^T = \{ X \mid X^T \in \mathbf{G} \subset \mathbb{M}_n \}$, or equivalently the group whose algebra is given by $\text{span}(\{B_1^T, B_2^T, \dots, B_k^T\})$. Then, the IVP (4.3.12) defines a curve in the space $\mathbf{G} \times \mathbf{G}^T$.

In [Bloch et al. 2002] a similar approach was performed on Rotation group $\mathbf{SO}(n)$: where in this case the group $\mathbf{SO}(n)^T$ coincides with $\mathbf{SO}(n)$ and the authors called to this framework the *symmetric representation of the rigid body*. Later the framework was extended to a similar symmetric representation of some quadratic groups, which include some interesting groups like Indefinite Rotation groups and Symplectic groups [Bloch et al. 2008]. A general framework has been recently proposed in [Bloch et al. 2011] from the view-point of embedded Optimal Control problems in \mathbb{M}_n .

4.4 Conserved quantities along geodesics

Noether's Theorem states that for every continuous symmetry of a dynamical system there is a corresponding conserved quantity. In the following, two conserved quantities are identified for a geodesic starting at Q_0 with initial velocity V_0 : E corresponds to the invariance of the time origin; and S corresponds to the invariance under the group composition.

- Geodesics are arc-length parameterized curves and therefore $\|\dot{Q}_{(t)}\|_{Q_{(t)}}$ is constant, as was illustrated in Figure 4.3.2, where $\bar{\mu}_{(t)}$ is constrained to evolve along the elliptical surface defined by \mathbf{m} . This means that $\langle\langle U_{(t)}, U_{(t)} \rangle\rangle_I = \langle \mathcal{M}(U_{(t)}), U_{(t)} \rangle = \langle \mathcal{M}(U_{(0)}), U_{(0)} \rangle$. Following Section 4.3.4, $\mathcal{M}(U)$ is represented by χU and the pairing is expressed in terms of the matrix trace function. Then, $\langle\langle U_{(t)}, U_{(t)} \rangle\rangle_I$, can be compactly written as $\text{trace}((\chi U_{(t)})^T U_{(t)})$.

²³ This definition makes sense only in a matrix group which is embedded in the set \mathbb{M}_n .

Finally, in terms of the states of a geodesic, the following quantity is conserved:

$$\begin{aligned} E &:= \langle\langle U_{(t)}, U_{(t)} \rangle\rangle_I \\ &= \text{trace}(\chi(Q_{(t)}^{-1}\dot{Q}_{(t)})^T Q_{(t)}^{-1}\dot{Q}_{(t)}) \\ &= \text{trace}(\chi(Q_0^{-1}V_0)^T Q_0^{-1}V_0) = \text{constant}. \end{aligned} \quad (4.4.1)$$

- Another conserved quantity is $P_{(t)}Q_{(t)}^T$, because its time-derivative (see Eq. (4.3.11b) and Eq. (4.3.11c)) vanishes

$$\begin{aligned} (P_{(t)}Q_{(t)}^T)' &= \dot{P}Q^T + P\dot{Q}^T \\ &= -PU^TQ^T + PU^TQ^T = 0. \end{aligned}$$

Using Eq. (4.3.10) and the fact that $P_{(t)} \in T_{Q(t)}^* \mathbf{G}$, in a similar way as was done to obtain $P_{(0)}$ in Section 4.3.5, it is obtained that $P_{(t)} = Q_{(t)}^{-T}\chi U_{(t)}$.

Finally, in terms of the states of a geodesic, the following quantity is also conserved:

$$\begin{aligned} S &:= P_{(t)}Q_{(t)}^T \\ &= Q_{(t)}^{-T}\chi(U_{(t)})Q_{(t)}^T \\ &= Q_{(t)}^{-T}\chi(Q_{(t)}^{-1}\dot{Q}_{(t)})Q_{(t)}^T \\ &= Q_0^{-T}\chi(Q_0^{-1}V_0)Q_0^T = \text{constant}. \end{aligned} \quad (4.4.2)$$

From a *physical* view-point the quantity E corresponds to the *kinetic energy* and the quantity S corresponds to the *spatial angular momentum* [Marsden 1992], which are preserved for a free-body.

It is worthy to mention that the conserved quantity S in Eq. (4.4.2) can be written in terms of intrinsic operations of the group. The expression $Q^{-T}WQ^T$ can be interpreted as the dual of the Adjoint action defined in Section 4.2.3

$$\langle Q^{-T}WQ^T, U \rangle = \text{trace}(QW^TQ^{-1}U) = \text{trace}(W^TQ^{-1}UQ) = \langle W, \mathbf{Ad}_{Q^{-1}}U \rangle,$$

for $W \in \mathfrak{g}^*$ and $U \in \mathfrak{g}$, therefore²⁴ $(\mathbf{Ad}_{Q^{-1}})^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^* : W \mapsto Q^{-T}WQ^T$. Accordingly the conserved quantity S is written as²⁵ $(\mathbf{Ad}_{Q(t)^{-1}})^*\chi U_{(t)} = \text{constant}$. The quantity S can also be obtained using a different approach, where additional geometrical structures and concepts are used, such as *Jacobi brackets* and *symplectic forms* [Abraham and Marsden 1978] [Marsden 1992] [Holm et al. 2009].

4.5 Reducing to a first-order equation

The methods for geodesic computation revisited so far include the following:

- Second order differential equation systems in local coordinates (\mathbb{R}^k) , see Sections 4.3.1 and 4.3.2.

²⁴ Two different notations are found in the literature: the most prevalent is $\mathbf{Ad}_\chi^* = (\mathbf{Ad}_{\chi^{-1}})^*$ satisfying $\mathbf{Ad}_\chi^* \circ \mathbf{Ad}_y^* = \mathbf{Ad}_{xy}^*$, see [Holm et al. 2009, p.215] [Kolev 2007]; the second one is $\mathbf{Ad}_\chi^* = (\mathbf{Ad}_\chi)^*$, which fulfils $\mathbf{Ad}_\chi^* \circ \mathbf{Ad}_y^* = \mathbf{Ad}_{yx}^*$ (note the change of the order in the foot) [Arnold 1989].

²⁵ In abstract notation: $((\mathbf{Ad}_{q(t)^{-1}})^* \circ \mathcal{M} \circ T_{q(t)} \mathbf{L}_{q(t)^{-1}})(\dot{q}_{(t)}) = \text{constant}$.

- A first order equation system in algebra coordinates followed by the reconstruction of the group curve $(\mathbb{R}^k \times \mathbf{G})$, see Section 4.3.3.
- A first order equation in the co-algebra representation followed by the reconstruction $(\mathfrak{g}^* \times \mathbf{G})$, see Section 4.3.4.
- A first order equation in the representation of co-tangent spaces coupled with the group curve evolution $(T^*\mathbf{G} \times \mathbf{G})$, see Section 4.3.5.
- A first order equation system of coupled evolutions in $(\mathbf{G}^T \times \mathbf{G})$, see Section 4.3.6.

All of them are either second order or first order systems but with a double number of degrees of freedom.

A direct consequence of the conservation of the quantity S is that geodesic evolution equations can be reduced to a first order equation exclusively in terms of the current element of the curve. From Eq. (4.4.2), multiplying from the left by Q^T and from the right by $Q_{(t)}^{-T}$

$$Q_{(t)}^T S Q_{(t)}^{-T} = \chi(Q_{(t)}^{-1} \dot{Q}_{(t)}) , \quad (4.5.1)$$

where $Q_{(t)}^{-1} \dot{Q}_{(t)} \in \mathfrak{g}$. Now, it can be noticed that $\overline{\sigma(\chi U)} = \mathbf{B}\mathbf{B}^\dagger \bar{U} = \bar{U}$, for all $U \in \mathfrak{g}$. Then, by applying σ to both sides of Eq. (4.5.1) and multiplying from the left by $Q_{(t)}$

$$\dot{Q}_{(t)} = Q_{(t)} \sigma \left(Q_{(t)}^T S Q_{(t)}^{-T} \right) \quad (\text{dotQ})$$

yielding a first order equation for the geodesic evolution without performing the trivial trick of doubling the dimension. The initial velocity of the curve V_0 is encoded in the value of $S = Q_0^{-T} \chi(Q_0^{-1} V_0) Q_0^T$. In particular, when $Q_0 = I$, then $S = \chi U_0$.

It is easy to see that $\exp(U_0)$ is the solution of $\text{Exp}_I(U_0)$ if $[\chi U_0, \exp(U_0^T)] = 0$. This is satisfied when $[\chi U_0, U_0^T] = 0$ which, following Eq. (4.3.7), means that $\dot{U} = 0$.

It is also interesting to compare Eq. (dotQ) with the symmetric representation given by IVP (4.3.12). Notice that if $Q_{(t)}$ is an integral curve of the equation $\dot{X} = XU$, then $Q_{(t)}^{-T}$ is an integral curve of $\dot{X} = -XU^T$, which is precisely the equation that guides the evolution of the variable $Y_{(t)}$.

As it was done in the previous section, the appearance of the transposed elements can be replaced by²⁶

$$\dot{Q}_{(t)} = Q_{(t)} \sigma \left((\mathbf{Ad}_{Q_{(t)}})^* S \right) ,$$

which is an expression intrinsically described by group operations.

Equation (dotQ) is the extension of a similar equation given in [Bloch et al. 2002] which was derived only for Rotation groups, $\mathbf{SO}(d)$. For this particular case the right-hand side turns out to be a cubic polynomial expression in the elements of Q .

²⁶ For abstract groups it can be shown that the corresponding ODE is

$$\dot{q}_{(t)} = \left(T_I \mathbf{L}_{q(t)} \circ \mathcal{M}^{-1} \circ (\mathbf{Ad}_{q(t)})^* \circ (\mathbf{Ad}_{q_0^{-1}})^* \circ \mathcal{M} \circ T_{q_0} \mathbf{L}_{q_0^{-1}} \right) \dot{q}(0) .$$

4.6 Sensitivity with respect to initial velocity

Given initial conditions $Q_0 = I \in \mathbf{G}$ and $\dot{Q}_{(0)} = U_0 \in \mathfrak{g}$, $\text{Exp}_I(U)$ is obtained by integrating Eq. (dotQ) (or any of the methodologies described through Section 4.3.1–4.3.6) up to $t = 1$. In this section we are interested in computing the Gâteaux derivative of the geodesic $Q_{(t)}$ with respect to U in a direction $W \in \mathfrak{g}$, i.e. $d_\epsilon \text{Exp}_I(t(U + \epsilon W))|_{(\epsilon=0)}$. Let $Q_{(t,\epsilon)}$ be the geodesic starting at I with a perturbed initial velocity $U_0 + \epsilon W$, such that $Q_{(t,0)} = Q_{(t)} = \text{Exp}_I(tU_0)$. The evolution of $Q_{(t,\epsilon)}$ is governed by Eq. (dotQ)

$$\dot{Q}_{(t,\epsilon)} = Q_{(t,\epsilon)} \boldsymbol{\sigma} \left(Q_{(t,\epsilon)}^T \boldsymbol{\chi}(U_0 + \epsilon W) Q_{(t,\epsilon)}^{-T} \right). \quad (4.6.1)$$

Taking derivatives with respect to ϵ on both sides of Eq. (4.6.1) and evaluating at $\epsilon = 0$, it is obtained

$$\begin{aligned} \dot{J}_{(t)} &= J_{(t)} \boldsymbol{\sigma} \left(Q_{(t)}^T \boldsymbol{\chi}(U_0) Q_{(t)}^{-T} \right) + \\ &Q_{(t)} \boldsymbol{\sigma} \left(J_{(t)}^T \boldsymbol{\chi}(U_0) Q_{(t)}^{-T} + Q_{(t)}^T \boldsymbol{\chi}(W) Q_{(t)}^{-T} - Q_{(t)}^T \boldsymbol{\chi}(U_0) Q_{(t)}^{-T} J_{(t)}^T Q_{(t)}^{-T} \right), \end{aligned} \quad (4.6.2)$$

where J is defined as $J_{(t)} = \partial_\epsilon Q_{(t,\epsilon)}|_{(\epsilon=0)} \in T_{Q_{(t)}} \mathbf{G}$ and the identity $\partial_\epsilon(Q^{-T}) = -Q^{-T}(\partial_\epsilon Q)^T Q^{-T}$ is used.

Integrating Eq. (4.6.2) with initial condition $J_{(0)} = 0$ results in a vector field along the geodesic which is known as *Jacobi field* [Do Carmo 1992, Chapter 5] [Sommer et al. 2013]. In these works the Jacobi field evolution equation is usually expressed by a second order equation which takes into account the curvature of the manifold. Similarly to Section 4.5, the second order equation can be reduced to a first order equation by using the symmetries of the group and metric.

Likewise the evolution of $Q_{(t)}$ was described in terms of algebra elements (see Section 4.3.3 and Section 4.3.4), it is convenient to describe $J_{(t)}$ as the left-trivialized variable $R_{(t)} = Q_{(t)}^{-1} J_{(t)} \in \mathfrak{g}$. The time evolution of $R_{(t)}$ is

$$\begin{aligned} \dot{R}_{(t)} &= (Q_{(t)}^{-1} J_{(t)})^\cdot = -Q_{(t)}^{-1} \dot{Q}_{(t)} Q_{(t)}^{-1} J_{(t)} + Q_{(t)}^{-1} \dot{J}_{(t)} \\ &= -UR + R \boldsymbol{\sigma} (Q^T \boldsymbol{\chi}(U_0) Q^{-T}) + \\ &\quad \boldsymbol{\sigma} (J^T \boldsymbol{\chi}(U_0) Q^{-T} + Q^T \boldsymbol{\chi}(W) Q^{-T} - Q^T \boldsymbol{\chi}(U_0) Q^{-T} J^T Q^{-T}). \end{aligned} \quad (4.6.3)$$

From Eq. (dotQ), $\boldsymbol{\sigma}(Q^T \boldsymbol{\chi}(U_0) Q^{-T}) = U$ and in consequence $Q^T \boldsymbol{\chi}(U_0) Q^{-T} = \boldsymbol{\chi}U$. Also note that $R^T = J^T Q^{-T}$ and therefore $J^T = R^T Q^T$. Finally, after simple algebraic manipulations, the expression in Eq. (4.6.3) simplifies to²⁷

$$\dot{R}_{(t)} = \boldsymbol{\sigma} \left(Q_{(t)}^T \boldsymbol{\chi}(W) Q_{(t)}^{-T} \right) + \boldsymbol{\sigma} [R_{(t)}^T, \boldsymbol{\chi}(U_{(t)})] + [R_{(t)}, U_{(t)}]. \quad (4.6.4)$$

With the previous results, it can be computed derivatives of functions defined on the group in terms of its normal coordinates. Given a function $\mathbf{f}: \mathbf{G} \rightarrow \mathbb{R}: Q \mapsto \mathbf{f}(Q)$ and let Q be

²⁷ In abstract notation:

$$\dot{R}_{(t)} = (\mathcal{M}^{-1} \circ (\text{Ad}_{q_{(t)}})^* \circ \mathcal{M}) \mathcal{W} + (\mathcal{M}^{-1} \circ \text{ad}_{\mathcal{R}_{(t)}}^* \circ \mathcal{M}) u_{(t)} + \text{ad}_{\mathcal{R}_{(t)}} u_{(t)}.$$

parameterized in terms of an initial velocity $U_0 = \sum \mu^i B_i \in \mathfrak{g}$, such that $Q = \text{Exp}_I(U_0)$, then its derivative with respect to the coordinates of U_0 is given by

$$\partial_{\mu^i} \mathbf{f}(\text{Exp}_I(U_0)) = \text{trace} \left(\left(\partial_X \mathbf{f}(X)|_{X=Q} \right)^T Q R_{(1)}^{B_i} \right),$$

where $R_{(1)}^{B_i}$ is the value of the solution of Eq. (4.6.4) at $t = 1$ with $W = B_i$.

4.7 Riemannian exponential functions under different left-invariant metrics

It was noted in [Milnor 1976] that the selection of “*different metrics on the same Lie group may exhibit drastically different curvatures properties*” of the manifold and different geodesics. Each k -dimensional Lie group possesses a $k(k - 1)/2$ -dimensional family of different inertia operators yielding distinct left-invariant metrics. It is possible to find a relationship between a smaller family of left-invariant Riemannian metrics and the family of their corresponding geodesics. Thus, given known the Riemannian exponential function for an inertia operator \mathcal{M} it is possible to find a family of Riemannian exponential functions whose inertia operators are given by \mathcal{M} composed with the Adjoint action of the group.

Let $(\mathbf{G}, \langle\langle \cdot, \cdot \rangle\rangle)$ be a Riemannian metric which is the left-invariant extension of $\langle\langle \cdot, \cdot \rangle\rangle_I$ and let γ be a geodesic with respect to this metric. As it was shown in Section 4.2.4, the velocities of the conjugated curve, $\xi_{(t)} = Y\dot{\gamma}_{(t)}Y^{-1}$, are $\dot{\xi}_{(t)} = Y\ddot{\gamma}_{(t)}Y^{-1} = \xi_{(t)}\mathbf{Ad}_Y(\gamma_{(t)}^{-1}\dot{\gamma}_{(t)})$. If a second left-invariant Riemannian metric is chosen such that²⁸ $\langle\langle (\mathbf{Ad}_Y)^{-1}(\cdot), (\mathbf{Ad}_Y)^{-1}(\cdot) \rangle\rangle_I$ then it is straightforward to verify that the velocities $\dot{\xi}_{(t)}$ are extremes of the Riemannian energy under the new metric and therefore, ξ is a geodesic for the second metric.

Given an inertia operator represented by \mathbf{m} which defines the inner product $\langle\langle U_1, U_2 \rangle\rangle_I = \overline{U_1^T \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger} \overline{U_2}$, and the operation $\overline{\mathbf{Ad}_{Y^{-1}} U} = \overline{Y^{-1} U Y} = (Y^T \otimes Y^{-1}) \overline{U}$, then the inner product $\langle\langle \mathbf{Ad}_{Y^{-1}} U_1, \mathbf{Ad}_{Y^{-1}} U_2 \rangle\rangle_I = \overline{U_1^T (Y \otimes Y^{-T}) \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger (Y^T \otimes Y^{-1})} \overline{U_2}$. Therefore, if the representation $\tilde{\mathbf{m}}$ of an inertia operator can be written as

$$\tilde{\mathbf{m}} = \mathbf{B}^T (Y \otimes Y^{-T}) \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^\dagger (Y^T \otimes Y^{-1}) \mathbf{B} \quad (4.7.1)$$

for some $Y \in \mathbf{G}$, and if there is a known solution of $\text{Exp}_I(\cdot)$ for the metric defined by \mathbf{m} , then the Riemannian exponential function under the metric defined by $\tilde{\mathbf{m}}$ is given by

$$\text{Exp}_I^{\tilde{\mathbf{m}}}(U) = Y \text{Exp}_I^{\mathbf{m}}(Y^{-1}UY) Y^{-1}. \quad (4.7.2)$$

This can be formally expressed as both left-invariant Riemannian metrics are isometric under the conjugacy by Y .

Besides, for a metric defined by $\hat{\mathbf{m}} = \alpha \mathbf{m}$, where $\alpha \in \mathbb{R}^+$, $\text{Exp}_I^{\hat{\mathbf{m}}}(U) = \text{Exp}_I^{\mathbf{m}}(U)$, i.e. the scaling of the inertia operator does not change the Riemannian exponential function, although it does change velocity norms, curve lengths, and distances between group elements.

However, the use of the previous conjugate and Adjoint actions does not allow to relate the whole family of inertia operators. In the best case, given \mathbf{m} for a k -dimensional group, the inertia operators represented by $\tilde{\mathbf{m}}$ and $\hat{\mathbf{m}}$ generate a $(k + 1)$ -dimensional family of left-invariant

²⁸ This metric corresponds to the left-translation of the inertia operator given by $\widetilde{\mathcal{M}} = (\mathbf{Ad}_{Y^{-1}})^* \circ \mathcal{M} \circ \mathbf{Ad}_{Y^{-1}}$.

metrics. Moreover, for some groups even a smaller family is generated. For example, let us consider $\mathbf{SO}(3)$ group with $\mathbf{m} = I_3$. The resulting $\tilde{\mathbf{m}}$ given by Eq. (4.7.1) is

$$\begin{aligned}\tilde{\mathbf{m}} &= \mathbf{B}^T(Y \otimes Y^{-T})(Y^T \otimes Y^{-1})\mathbf{B} \\ &= \mathbf{B}^T(YY^T \otimes Y^{-T}Y^{-1})\mathbf{B} \\ &= I_3 = \mathbf{m}\end{aligned}$$

for any $Y \in \mathbf{SO}(3)$. Therefore, in this case, the conjugation trick does not help to solve the Riemannian exponential for different non-trivial inertia operators.

4.8 Expressions for the right-invariant metric case

Through the previous sections, different expressions for geodesics under left-invariant metrics were derived. The most important relations corresponding to *right-invariant metrics version* are given below.

The Euler–Poincaré equation for a right-invariant metric is given by changing the sign of (E–P)

$$\frac{d}{dt} \mathcal{M}(U_{(t)}) = -\text{ad}_{U_{(t)}}^*(\mathcal{M}(U_{(t)})) \quad (\text{E–P R})$$

and the problem of computing a geodesic $Q_{(t)}$ passing through X with velocity V at $t = 0$ becomes in

$$\begin{aligned}\dot{\bar{\mu}}_{(t)} &= -\mathbf{m}^{-1}(\text{ad}_{\bar{\mu}_{(t)}}^*(\mathbf{m}\bar{\mu}_{(t)})) \\ \text{with initial conditions } \bar{\mu}_{(0)} &= \mathbf{B}^\dagger(\overline{VX^{-1}}),\end{aligned}\quad (4.3.5 \text{ R})$$

and afterwards solve the reconstruction equation given by the ODE

$$\begin{aligned}\dot{Q}_{(t)} &= \left(\sum_{i=1}^k B_i \mu_{(t)}^i \right) Q \\ \text{with initial conditions } Q_{(0)} &= X.\end{aligned}\quad (\text{recons R})$$

In terms of matrix representation of the algebra/co-algebra vectors

$$\dot{U}_{(t)} = -\boldsymbol{\sigma}[U_{(t)}^T, \boldsymbol{\chi}U_{(t)}]. \quad (4.3.7 \text{ R})$$

The IVP (4.3.11) derived from Optimal Control problem, (ocp $\bar{\mu}$), is given by

$$U_{(t)} = \boldsymbol{\sigma}(P_{(t)}Q_{(t)}^T) \quad (4.3.11a \text{ R})$$

$$\dot{Q}_{(t)} = U_{(t)}Q_{(t)} \quad (4.3.11b \text{ R})$$

$$\dot{P}_{(t)} = -U_{(t)}^T P_{(t)} \quad (4.3.11c \text{ R})$$

$$Q_{(t=0)} = Q_0 \quad (4.3.11d \text{ R})$$

and the initial value $P_{(0)} = \boldsymbol{\chi}(U_{(0)})Q_0^{-T}$.

The IVP (4.3.12) derived from the symmetric representation results for a right-invariant

metric in

$$U_{(t)} = \sigma(Y_{(t)}P_0Q_{(t)}^T) \quad (4.3.12a \text{ R})$$

$$\dot{Q}_{(t)} = U_{(t)}Q_{(t)} \quad (4.3.12b \text{ R})$$

$$\dot{Y}_{(t)} = -U_{(t)}^TY_{(t)} \quad (4.3.12c \text{ R})$$

$$Q_{(t=0)} = Q_0 \quad (4.3.12d \text{ R})$$

$$Y_{(t=0)} = I \quad (4.3.12e \text{ R})$$

$$P_0 = \chi(U_{(0)})Q_0^{-T}. \quad (4.3.12f \text{ R})$$

Regarding the conserved quantities, the expression for E does not change

$$E := \text{trace}(\chi(\dot{Q}_{(t)}Q_{(t)}^{-1})^T\dot{Q}_{(t)}Q_{(t)}^{-1}) = \text{trace}(\chi(U_0)^TU_0), \quad (4.4.1 \text{ R})$$

while the expression for S becomes

$$S := Q_{(t)}^TP_{(t)} = Q_{(t)}^T\chi(\dot{Q}_{(t)}Q_{(t)}^{-1})Q_{(t)}^{-T} = Q_{(0)}^T\chi(V_0Q_0^{-1})Q_{(0)}^{-T}, \quad (4.4.2 \text{ R})$$

corresponding to $(\text{Ad}_{Q(t)})^*\chi U_{(t)} = \text{constant}$.

The evolution of the geodesics under a right-invariant metric described by a first-order differential equation results in

$$\dot{Q}_{(t)} = \sigma(Q_{(t)}^{-T}S Q_{(t)}^T) Q_{(t)} \quad (\text{dotQ R})$$

which can be written for abstract structures as

$$\dot{q}_{(t)} = \left(T_I \mathbf{R}_{q(t)} \circ \mathcal{M}^{-1} \circ (\text{Ad}_{q(t)^{-1}})^* \circ (\text{Ad}_{q_0})^* \circ \mathcal{M} \circ T_{q_0} \mathbf{R}_{q_0^{-1}} \right) \dot{q}_{(0)}.$$

Finally, the sensitivity with respect to the initial velocity under a right-invariant metric can be computed by

$$\begin{aligned} \dot{J}_{(t)} &= \sigma \left(Q_{(t)}^{-T} \chi(U_0) Q_{(t)}^T \right) J_{(t)} + \\ &\quad \sigma \left(Q_{(t)}^{-T} \chi(U_0) J_{(t)}^T + Q_{(t)}^{-T} \chi(W) Q_{(t)}^T - Q_{(t)}^{-T} J_{(t)}^T Q_{(t)}^{-T} \chi(U_0) Q_{(t)}^T \right) Q_{(t)} \end{aligned} \quad (4.6.2 \text{ R})$$

and it is described from the algebra by

$$\dot{R}_{(t)} = \sigma \left(Q_{(t)}^{-T} \chi(W) Q_{(t)}^T \right) - \sigma[R_{(t)}^T, \chi(U_{(t)})] - [R_{(t)}, U_{(t)}], \quad (4.6.4 \text{ R})$$

which, in abstract notation, is written as

$$\dot{\mathcal{R}}_{(t)} = (\mathcal{M}^{-1} \circ \text{Ad}_{q_{(t)}^{-1}}^* \circ \mathcal{M}) \mathcal{W} - (\mathcal{M}^{-1} \circ \text{ad}_{\mathcal{R}_{(t)}}^* \circ \mathcal{M}) \mathcal{U}_{(t)} - \text{ad}_{\mathcal{R}_{(t)}} \mathcal{U}_{(t)}.$$

CHAPTER

5

Solutions for left-invariant geodesics and applications

5.1	Closed-form solutions of $\text{Exp}(\cdot)$	114
5.1.1	Translation group, $\mathcal{T}(d)$	114
5.1.2	Isotropic Scale group, $\mathcal{S}^+(1)$	115
5.1.3	Non-Isotropic Scale group, $\mathcal{S}^+(d)$	115
5.1.4	Isotropic Scale + Translation group, $\mathcal{ST}(d)$	115
5.1.5	Rotation group, $\mathcal{SO}(d)$	116
5.1.6	Special Euclidean group, $\mathcal{SE}(d)$	116
5.1.7	Similarity group, $\mathcal{SIM}(d)$	117
5.1.8	Orientation preserving General Linear group, $\mathcal{GL}^+(d)$	117
5.1.9	Special Linear group, $\mathcal{SL}(d)$	117
5.1.10	Centered Transformation group, $(\mathcal{G}(d) \times \mathcal{T}(d))$	118
5.1.11	Projective group, $\mathcal{PGL}(d)$	118
5.1.12	Möbius Transformation group, $\mathcal{MO}(d)$	119
5.2	Numerical solutions for the $\text{Exp}()$ function	121
5.2.1	Explicit additive Euler method	121
5.2.2	Explicit Lie–Euler method	122
5.2.3	Taylor series	122
5.2.4	Fixed step-size Taylor method	125
5.2.5	Adaptive Taylor method	125
5.2.6	Computing the Riemannian logarithm function	126
5.3	Results	128
5.3.1	Accuracy	128
5.3.2	Timing performance of Adaptive Taylor method	130
5.3.3	Departures from the group set	131
5.4	Application examples	131
5.4.1	Interpolating spatial transformations	133
5.4.2	Intrinsic sample mean of spatial transformations	136
5.4.3	Intrinsic filtering	140
5.4.4	Extension to any metric for $\mathcal{ST}(1)$ group	142

While, throughout the previous Chapter 4 the evolution of geodesics was described and formulated as ordinary differential equations, in the current chapter we will investigate the solutions of those ODEs and the resulting Riemannian exponential functions. Some few of them can be expressed as closed-forms of known algebraic expressions. However, most cases must be solved by numerical procedures. Simple numerical schemes for computation of the Riemannian exponential function and its inverse are presented. These numerical schemes are given only as basic and illustrative examples. In addition, some illustrative applications of invariant Riemannian geodesics on spatial transformation groups are shown: invariant interpolation of spatial transformations; intrinsic sample means of spatial transformations; Gaussian filtering of a field of spatial transformations; and an extension from closed-form solutions for a particular metric to left-invariant geodesics for $\mathcal{ST}(1)$ group under any invariant metric.

5.1 Closed-form solutions of $\text{Exp}(\cdot)$ for commonly used transformation groups

For a few groups, and under particular invariant metrics, the Riemannian exponential function, and therefore geodesics, can be expressed in closed-form¹. We take the opportunity to gather, summarize and present some new closed-form solutions of left-invariant Riemannian exponential functions for common matrix groups used in computer vision and imaging applications.

For each of the following spatial transformation models, transformation elements can be written in terms of some parameters \bar{p} such that $h^{\bar{p}}$ unambiguously identifies the transformation (see also Section 2.3). A matrix representation of the group element is usually denoted by $H(\bar{p})$ and the corresponding representation of algebra vectors by $U(\bar{\mu})$. Also, a faithful matrix representation of a basis of the algebra is given. Finally, a closed-form expression for the Riemannian exponential function developed around the identity under a specified left-invariant metric is given. Riemannian exponential functions developed around another group element can be computed by using Eq. (4.2.11). For the case of right-invariant Riemannian metrics, Eq. (4.2.12) gives a closed-form expression to make the change from left- to right-invariant. Finally, most of the results given in this section are valid only for some particular inertia operators, but the conjugation trick explained in Section 4.7 can be used to extend the closed-form expressions to some other metrics.

5.1.1 Translation group, $\mathcal{T}(d)$

A matrix representation of group elements and algebra vectors is given, respectively, by $(d+1)\times(d+1)$ matrices of the form

$$H(\bar{t}) = \begin{pmatrix} I_d & \bar{t} \\ 0_d^T & 1 \end{pmatrix} \quad \text{and} \quad U(\bar{\tau}) = \begin{pmatrix} 0_{d\times d} & \bar{\tau} \\ 0_d^T & 0 \end{pmatrix},$$

where $\bar{t}, \bar{\tau} \in \mathbb{R}^d$.

The Riemannian exponential under any left-invariant metric is

$$\text{Exp}_I^{\mathcal{T}(d)}(U(\bar{\tau})) = H(\bar{\tau}).$$

¹ For some groups, the closed-form expression is written in terms of the matrix exponential function. Although there is no global consensus about what a “closed-form” expression means [Borwein and Crandall 2013], the matrix exponential function, $\exp(\cdot)$, can be computed accurately and efficiently by most scientific software packages like **Mathematica** and **MATLAB** [Moler and Van Loan 2003] [Al-Mohy 2010]. Moreover, pure algebraic and explicit formulas (although complicated) in terms of *elementary operations* can be given for matrices of size smaller than 5 and for some larger structured matrices.

5.1.2 Isotropic Scale group, $\mathcal{S}^+(1)$

A matrix representation is given by 1×1 matrices $H(s) = s \in \mathbb{R}^+$, and the algebra vectors are of the form $U(\sigma) = \sigma$, with $\sigma \in \mathbb{R}$.

The $\text{Exp}(\cdot)$ function under any left-invariant metric is

$$\text{Exp}_I^{\mathcal{S}^+(1)}(U(\sigma)) = H(e^\sigma).$$

5.1.3 Non-Isotropic Scale group, $\mathcal{S}^+(d)$

Group elements and algebra vectors can be represented by

$$H(\bar{s}) = \begin{pmatrix} s^1 & 0 & \dots & 0 \\ 0 & s^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & s^d \end{pmatrix} \quad \text{and} \quad U(\bar{\sigma}) = \begin{pmatrix} \sigma^1 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^d \end{pmatrix},$$

where $s^i \in \mathbb{R}^+$ and $\sigma^i \in \mathbb{R}$.

The $\text{Exp}_I^{\mathcal{S}^+(d)}$ function, for any left-invariant metric, is

$$\text{Exp}_I^{\mathcal{S}^+(d)}\left(U\left(\left(\sigma^1, \sigma^2, \dots, \sigma^d\right)^T\right)\right) = H\left(\left(e^{\sigma^1}, e^{\sigma^2}, \dots, e^{\sigma^d}\right)^T\right).$$

Remark

The three previous groups are commutative and therefore any left-invariant metric is also right-invariant and the corresponding Riemannian exponential function, $\text{Exp}_I(\cdot)$, coincides with the group exponential function, $\exp(\cdot)$ (see Section 4.2.10).

5.1.4 Isotropic Scale + Translation group, $\mathcal{ST}(d)$

Matrix representations are

$$H(s, \bar{t}) = \begin{pmatrix} sI_d & \bar{t} \\ 0_d^T & 1 \end{pmatrix} \quad \text{and} \quad U(\sigma, \bar{\tau}) = \begin{pmatrix} \sigma I_d & \bar{\tau} \\ 0_d^T & 0 \end{pmatrix},$$

with $s \in \mathbb{R}^+$, $\sigma \in \mathbb{R}$ and $\bar{t}, \bar{\tau} \in \mathbb{R}^d$.

The following set of $(d+1) \times (d+1)$ matrices can be used as bases of the algebra:

$$\left\{ \begin{pmatrix} I_d & 0_d \\ 0_d^T & 0 \end{pmatrix}, \begin{pmatrix} 0_{d \times d} & \mathbf{e}_1 \\ 0_d^T & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0_{d \times d} & \mathbf{e}_d \\ 0_d^T & 0 \end{pmatrix} \right\},$$

where \mathbf{e}_j is the j -th canonical basis of \mathbb{R}^d .

With these bases and for $\mathbf{m} = I_{d+1}$, the left-invariant Riemannian exponential has the following closed-form expression:

$$\text{Exp}_I^{\mathcal{ST}(d)}(U(\sigma, \bar{\tau})) = \begin{cases} H\left(\frac{(b^2+1)e^w}{b^2e^{2w}+1}, \frac{b(e^{2w}-1)}{b^2e^{2w}+1}\|\bar{\tau}\|\right), & \text{if } \|\bar{\tau}\| \neq 0 \\ H(e^\sigma, 0), & \text{else} \end{cases}, \quad (5.1.1)$$

where $\|\cdot\|$ is the Euclidean norm in \mathbb{R}^d , $w = \sqrt{\sigma^2 + \|\bar{\tau}\|^2}$ and $b = (w - \sigma)/\|\bar{\tau}\|$.

The expressions in Eq. (5.1.1) can be derived by noticing that $\mathcal{ST}(1)$ group with a left-invariant metric is isomorphic to the Poincaré-half plane [Arsigny 2006] [Stahl 2007] [Vaccaro

2012]. These can be derived from the IVP obtained in Section 4.3.1 after an involved algebraic work.

Note that the closed-form in Eq. (5.1.1) is valid for any spatial dimension d , while in most of the literature only the 1D case is usually found.

5.1.5 Rotation group, $\mathcal{SO}(d)$

Let R_θ be a matrix representing a rotation on a d -dimensional space, then $R_\theta R_\theta^T = R_\theta^T R_\theta = I_d$ and $\det(R_\theta) = 1$ (see Section 2.5.2).

The matrix representation of an element of the algebra $\mathfrak{so}(d)$ can be obtained by taking time derivatives of a curve passing through the identity. Noticing that $\dot{R}_\theta R_\theta^T + R_\theta \dot{R}_\theta^T = 0$ and particularizing to the identity it is obtained that $\dot{R} = -\dot{R}^T$. Therefore, the algebra is composed by the set of skew-symmetric $d \times d$ matrices which is a $d(d-1)/2$ dimensional space [Moakher 2002] [Gallier and Xu 2002] [Hartley et al. 2013]

$$U(\bar{\omega}) = \Sigma(\bar{\omega}) = \Sigma \left(\left(\omega^1, \omega^2, \dots, \omega^{d(d-1)/2} \right)^T \right),$$

where $\Sigma(\bar{\omega})$ is a $d \times d$ skew-symmetric matrix built by rearranging the entries of the vector $\bar{\omega}$. For example, in the 3D case where $\bar{\omega} = ((\omega^x, \omega^y, \omega^z)^T$

$$\Sigma(\bar{\omega}) = \begin{pmatrix} 0 & -\omega^z & \omega^y \\ \omega^z & 0 & -\omega^x \\ -\omega^y & \omega^x & 0 \end{pmatrix}.$$

It is well known that $\mathcal{SO}(d)$ group admits a bi-invariant metric [Park and Ravani 1997] [Arsigny 2006] [Huynh 2009] [Hartley et al. 2013]. It corresponds to the inertia operator represented by the identity matrix, $\mathbf{m} = I_d$, when the algebra bases are given by $\{\Sigma(\mathbf{e}_i)\}_{i=1}^{d(d-1)/2}$. For example, in the 3D case

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

Therefore, the inner product in the algebra results in $\langle\langle U_1, U_2 \rangle\rangle_I = \text{trace}(U_1^T U_2)/2$ and the corresponding Riemannian exponential function is given by the group exponential,

$$\text{Exp}_I^{\mathcal{SO}(d)}(\Sigma(\bar{\omega})) = \exp(\Sigma(\bar{\omega})).$$

5.1.6 Special Euclidean group, $\mathcal{SE}(d)$

Matrix representations of elements and algebra vectors are

$$H(\bar{\theta}, \bar{t}) = \begin{pmatrix} R_\theta & \bar{t} \\ 0_d^T & 1 \end{pmatrix} \quad \text{and} \quad U(\bar{\omega}, \bar{\tau}) = \begin{pmatrix} \Sigma(\bar{\omega}) & \bar{\tau} \\ 0_d^T & 0 \end{pmatrix},$$

where R_θ is the rotation parameterized by $\theta, \bar{t}, \bar{\tau} \in \mathbb{R}^d$ and $\Sigma(\bar{\omega})$ is the skew-symmetric $d \times d$ matrix with entries ω^i , analogously to the case of $\mathcal{SO}(d)$ group.

For a left-invariant metric defined by $\mathbf{m} = I_{d+1}$, the Riemannian exponential function results in [Park 1995]

$$\text{Exp}_I^{\mathcal{SE}(d)}(U(\bar{\omega}, \bar{\tau})) = \begin{pmatrix} \exp(\Sigma(\bar{\omega})) & \bar{\tau} \\ 0_d^T & 1 \end{pmatrix}.$$

5.1.7 Similarity group, $\mathcal{SIM}(d)$

To compute the Riemannian exponential for this group it can be use of the following mapping: let $\Phi: \mathcal{SIM}(d) \rightarrow \mathcal{SO}(d) \times \mathcal{ST}(d)$ be a smooth map [Bossa et al. 2011]

$$\Phi\left(h^{(s,\bar{\theta},\bar{t})}\right) = \left(r^{\bar{\theta}}, q^{(s,\bar{t})} \right),$$

where $r^{\bar{\theta}} \in \mathcal{SO}(d)$ and $q^{(s,\bar{t})} \in \mathcal{ST}(d)$.

Under a block-scalar metric, *i.e.* a metric under which the contributions from both spaces do not interact, the mapping Φ results in an isometry [Gallot et al. 2004], thus under this metric, geodesics on $\mathcal{SIM}(d)$ are obtained by lifting geodesics on $\mathcal{SO}(d)$ and $\mathcal{ST}(d)$, respectively

$$\begin{aligned} \text{Exp}_I^{\mathcal{SIM}(d)}\left(\left(\begin{array}{c|c} \sigma I_d + \Sigma(\bar{\omega}) & |\bar{\tau} \\ \hline 0_d^T & |0 \end{array}\right)\right) = \\ \Phi^{-1}\left(\left(\text{Exp}_I^{\mathcal{SO}(d)}(\Sigma(\bar{w})), \text{Exp}_I^{\mathcal{ST}(d)}\left(\left(\begin{array}{c|c} \sigma I_d & |\bar{\tau} \\ \hline 0_d^T & |0 \end{array}\right)\right)\right)\right). \end{aligned}$$

5.1.8 Orientation preserving General Linear group, $\mathcal{GL}^+(d)$

For this group, the algebra $\mathfrak{gl}(d)$ is the set of all $d \times d$ real matrices \mathbb{M}_d . For the canonical bases of \mathbb{M}_d and under a left-invariant metric defined by $\mathbf{m} = I_{d^2}$, a closed-form solution of the Riemannian exponential function was derived in [Andruchow et al. 2011] [Vandereycken et al. 2013]

$$\text{Exp}_I^{\mathcal{GL}^+(d)}(U) = \exp(U^T) \exp(U - U^T). \quad (5.1.2)$$

The corresponding inner product at the algebra is $\langle\langle U_1, U_2 \rangle\rangle_I = \text{trace}(U_1^T U_2)$.

Note that the group exponential $\exp(\cdot)$ is not surjective (Section 4.2.5) while the Riemannian exponential $\text{Exp}(\cdot)$ is (Section 4.2.8). However, there is no contradiction between these statements and Eq. (5.1.2), where a composition of two different group exponentials appears at the r.h.s. This is in agreement with the results of [Wüstner 2003], where it is proved that the composition of two group exponentials is enough to fully cover the identity component of any group.

5.1.9 Special Linear group, $\mathcal{SL}(d)$

A matrix representation of an element is given by a matrix in $\mathbf{GL}^+(d)$ with determinant equals to 1. An algebra representation consists of the set of $d \times d$ traceless matrices [Tapp 2005] [Li et al. 2009] which is a $d^2 - 1$ dimensional space.

Although in [Li et al. 2009] a closed-form expression was derived from scratch, it is interesting to obtain it from using the concept of totally geodesic subgroups.

It is easy to shown that $\mathcal{SL}(d)$ is a totally geodesic subgroup of $\mathcal{GL}^+(d)$ under the metric defined by $\langle\langle U_1, U_2 \rangle\rangle_I = \text{trace}(U_1^T U_2)$. Therefore, Eq. (5.1.2) is also a closed-form solution valid for $\mathcal{SL}(d)$ group.

Although it is beyond the scope of this thesis, it can be noticed that $\mathcal{SL}(d)$ is a *semisimple group* and therefore its algebra admits a *Cartan decomposition* into skew-symmetric matrices and traceless-symmetric matrices [Helgason 1979, Chapter IV]. It is shown in [Helgason 1979, p.277] [Wang 1969] that, under the particular left-invariant metric given by minus the product of the Killing form and the Cartan involution, the Riemannian exponential function have the closed-form solution $\text{Exp}_I(U) = \exp(X - T) \exp(2T)$, where $T = (U - U^T)/2$ is the projection to skew-symmetric matrices and $X = U - T$ is the projection to the space of traceless-symmetric

matrices. Moreover, expressing $\mathcal{GL}^+(d)$ as the direct product $\mathcal{SL}(d) \times \mathcal{S}^+(1)$, the closed-form of $\text{Exp}_I^{\mathcal{GL}^+(d)}$ in Eq. (5.1.2) is easily recovered.

5.1.10 Centered Transformation group, $(\mathcal{G}(d) \times \mathcal{T}(d))$

In groups of centered transformations the action on an object with center if performed by first transform with respect to the center (instead of about the origin of the coordinate system) and then the object and its center are translated together (see Section 2.5.2).

Matrix representations of elements from these groups elements and algebra vectors are

$$H(\bar{p}, \bar{t}) = \left(\begin{array}{c|c|c} H^{\mathcal{G}}(\bar{p}) & 0_{n \times d} & 0_n \\ \hline 0_{d \times n} & I_d & \bar{t} \\ \hline 0_n^T & 0_d^T & 1 \end{array} \right) \quad \text{and} \quad U(\bar{\rho}, \bar{\tau}) = \left(\begin{array}{c|c|c} U^{\mathcal{G}}(\bar{\rho}) & 0_{n \times d} & 0_n \\ \hline 0_{d \times n} & 0_{d \times d} & \bar{\tau} \\ \hline 0_n^T & 0_d^T & 0 \end{array} \right) .$$

The algebra basis for this group is the direct sum $\mathfrak{g}(d) \oplus \mathfrak{t}(d)$, where $\mathfrak{g}(d)$ is the algebra of the k -dimensional group $\mathcal{G}(d)$ and $\mathfrak{t}(d)$ is the algebra of translation group $\mathcal{T}(d)$.

For a left-invariant metric defined by a block structured matrix \mathbf{m}

$$\mathbf{m} = \left(\begin{array}{c|c} \mathbf{m}^{\mathcal{G}} & 0_{k \times d} \\ \hline 0_{d \times k} & \mathbf{m}^{\mathcal{T}} \end{array} \right) ,$$

the Riemannian exponential results in

$$\text{Exp}_I^{(\mathcal{G} \times \mathcal{T})} (U(\bar{\rho}, \bar{\tau})) = \left(\begin{array}{c|c|c} \text{Exp}_I^{\mathcal{G}}(U^{\mathcal{G}}(\bar{\rho})) & 0_{n \times d} & 0_n \\ \hline 0_{d \times n} & I_d & \bar{\tau} \\ \hline 0_n^T & 0_d^T & 1 \end{array} \right) ,$$

which is the direct product of geodesics from \mathcal{G} and $\mathcal{T}(d)$.

5.1.11 Projective group (a.k.a. Homography group), $\mathcal{PGL}(d)$

Projective transformation from this group are, in general, not proper spatial transformations. Technically, this group does not generate bijections for points of \mathbb{R}^d , but it is well defined for lines through the origin in \mathbb{R}^{d+1} . However, it is included in this list because it is a very useful tool in many computer vision applications.

A matrix representation of the group elements is attained considering the action on a *projective space*. In a projective space every point of \mathbb{R}^d is identified with a real line passing through the origin of a $(d+1)$ dimensional space. The mappings ψ and ψ^{-1} allow the identification of $y \in \mathbb{R}^d$ with its corresponding line in \mathbb{R}^{d+1} and viceversa

$$\begin{aligned} \psi(y) &= [y^T, 1]^T \\ \psi^{-1}(\bar{z}) &= \frac{(z^1, \dots, z^d)^T}{z^{d+1}} . \end{aligned}$$

Projective transformations act linearly on \mathbb{R}^{d+1} . If the denominator in the last expression vanishes, the corresponding \mathbb{R}^d points are not well defined and these points are known as *ideals* [Hartley and Zisserman 2004]. Due to the equivalence relation $\psi^{-1}(\alpha z) = \psi^{-1}(z)$ for any non-zero scalar α , linear transformations related by a scalar factor are identified. Therefore, the transformation group $\mathcal{PGL}(d+1)$ is defined as the quotient $\mathcal{GL}(d+1)/(\mathbb{R}_{\neq 0})$. Every transformation in a neighborhood of the identity can be identified with a matrix of $\mathbf{GL}(d+1)$

with determinant equal to 1 and therefore the group becomes, at least locally, isomorphic to $\mathcal{SL}(d+1)$ group.

Matrix representations of the group elements are the $(d+1) \times (d+1)$ matrices with determinant equal to 1, and matrix representations of the algebra vectors are $(d+1) \times (d+1)$ matrices with trace equal to 0, resulting in a $(d+1)^2 - 1$ dimensional space.

The action on \mathbb{R}^d points, when it is defined, can be written as

$$\hbar(\bar{p}) \star y = \psi^{-1}(H(\bar{p})\psi(y)),$$

where \bar{p} are the parameters of a matrix with determinant equal to 1.

Finally, under the left-invariant metric generated by $\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \text{trace}(U_1^T U_2)$, the Riemannian exponential function can be computed via Eq. (5.1.2).

5.1.12 Möbius Transformation group, $\mathcal{MO}(d)$

Similarly to the $\mathcal{PGL}(d)$ group, some elements of this group does not generate homeomorphisms of the Euclidean space. But it has nice properties in analysis and in computer graphics area [Arnold and Rogness 2008].

This group includes transformations of the Euclidean space that locally preserve angles. Möbius transformations are obtained by composition of translations, rotations, dilations and inversions with respect to the unit sphere ($\mathbf{i}: y \mapsto y/\|y\|^2$). Note that, an even number of inversions must be used in order to preserve the orientation. A parametrization and its action are given by $\hbar(s, \bar{\theta}, \bar{t}, \bar{w}) \star y = sR_{\bar{\theta}}\mathbf{i}(\mathbf{i}(y) + \bar{w}) + \bar{t}$.

In order to obtain a matrix representation of the transformation element, a mapping to an hyperbolic space can be done: let $\bar{z} \in \mathbb{R}^{d+1,1}$ with a norm defined by $\|\bar{z}\|_{hyp}^2 = \sum_i^{d+1} (z^i)^2 - (z^{d+2})^2$, and let $\psi: \mathbb{R}^d \rightarrow \mathbb{R}^{d+1,1}$ be the mapping given by

$$\begin{aligned} \psi(y) &= \left(y^T, \frac{\|y\|^2 - 1}{2}, -\frac{\|y\|^2 + 1}{2} \right)^T \\ \psi^{-1}(\bar{z}) &= \frac{(z^1, \dots, z^d)^T}{-z^{d+1} - z^{d+2}}. \end{aligned}$$

Note that $\|\psi(y)\|_{hyp}^2 = 0$ and $\psi^{-1}(\alpha\bar{z}) = \psi^{-1}(\bar{z})$ for any non-zero scalar α . With the mapping ψ , conformal transformations in \mathbb{R}^d consists of isometries in $\mathbb{R}^{d+1,1}$ modulo $\{\pm \mathcal{ID}(d+2)\}$, which is the projective orthogonal group $\mathcal{PO}(d+1, 1)$ [Gallot et al. 2004]. This group is locally isomorphic to the indefinite special orthogonal group $\mathcal{SO}(d+1, 1)$. Specifically, matrix representations of the elementary transformations that make up Möbius group are H_D , H_R , H_T and H_i corresponding to dilations, rotations, translations and the inversion, respectively:

$$\begin{aligned} H_D(s) &= \left(\begin{array}{c|cc} I_d & 0_d & 0_d \\ \hline 0_d^T & \frac{1}{2}(\frac{1}{s} + s) & \frac{1}{2}(\frac{1}{s} - s) \\ 0_d^T & \frac{1}{2}(\frac{1}{s} - s) & \frac{1}{2}(\frac{1}{s} + s) \end{array} \right) & H_R(\bar{\theta}) &= \left(\begin{array}{c|cc} R_{\bar{\theta}} & 0_d & 0_d \\ \hline 0_d^T & 1 & 0 \\ 0_d^T & 0 & 1 \end{array} \right) \\ H_T(\bar{t}) &= \left(\begin{array}{c|cc} I_d & -\bar{t} & -\bar{t} \\ \hline \bar{t}^T & 1 - \frac{\|\bar{t}\|^2}{2} & -\frac{\|\bar{t}\|^2}{2} \\ -\bar{t}^T & \frac{\|\bar{t}\|^2}{2} & 1 + \frac{\|\bar{t}\|^2}{2} \end{array} \right) & H_i &= \left(\begin{array}{c|cc} I_d & 0_d & 0_d \\ \hline 0_d^T & -1 & 0 \\ 0_d^T & 0 & 1 \end{array} \right) \end{aligned}$$

where $s \in \mathbb{R}^+$, $\bar{\theta}$ is a parametrization of a rotation matrix and $\bar{t} \in \mathbb{R}^d$.

In terms of the previous elementary transformations, a matrix representation of $\hbar(s, \bar{\theta}, \bar{t}, \bar{w})$ is given by

$$H(s, \bar{\theta}, \bar{t}, \bar{w}) = H_T(\bar{t}) H_D(s) H_R(\bar{\theta}) H_i H_T(\bar{w}) H_i.$$

As in the case of the projective transformations, the action of a Möbius transformation on a coordinate point of \mathbb{R}^d is $\tilde{\mu}(s, \bar{\theta}, \bar{t}, \bar{w}) \star y = \psi^{-1}(H\psi(y))$ when it is well defined and no vanishing denominator appears.

The following set of $(d+2) \times (d+2)$ matrices can be used as representation of the bases of the algebra:

$$\left(\begin{array}{c|cc} 0_{d \times d} & 0_d & 0_d \\ \hline 0_d^T & 0 & -1 \\ 0_d^T & -1 & 0 \end{array} \right) \cup \left\{ \left(\begin{array}{c|cc} \Sigma(e_i) & 0_d & 0_d \\ \hline 0_d^T & 0 & 0 \\ 0_d^T & 0 & 0 \end{array} \right) \right\}_{i=1}^{d(d-1)/2} \cup \dots$$

$$\dots \left\{ \left(\begin{array}{c|cc} 0_{d \times d} & -e_j & -e_j \\ \hline e_j^T & 0 & 0 \\ -e_j^T & 0 & 0 \end{array} \right) \right\}_{j=1}^d \cup \left\{ \left(\begin{array}{c|cc} 0_{d \times d} & e_j & -e_j \\ \hline -e_j^T & 0 & 0 \\ -e_j^T & 0 & 0 \end{array} \right) \right\}_{j=1}^d.$$

It can be proved that with the above representation, $\mathcal{MO}(d)$ group is a totally geodesic subgroup of $\mathcal{GL}^+(d+2)$ group under the left-invariant metric given by $\langle\langle U_1, U_2 \rangle\rangle_I = \text{trace}(U_1^T U_2)$. Therefore, for this metric, the Riemannian exponential function can be computed explicitly using Eq. (5.1.2).

Commonly used transformation groups without known closed-form solution for $\text{Exp}(\cdot)$

The group of affine transformations, $\mathcal{GA}(d)$ is very used in medical imaging analysis, for example in the image registration stage. However, up to our knowledge, there is no known closed-form solution for $\text{Exp}(\cdot)$ function under any left-invariant Riemannian metric (except for the trivial case $d = 1$). However, matrix representations of the algebra elements are included in the present list for completeness.

Orientation preserving General Affine group, $\mathcal{GA}^+(d)$

A matrix representation of the group elements is

$$H(\bar{p}, \bar{t}) = \left(\begin{array}{c|c} M_p & \bar{t} \\ \hline 0_d^T & 1 \end{array} \right),$$

a matrix representation of the algebra vectors is obtained by the $(d+1) \times (d+1)$ matrices with the last row equals to zero.

No known solutions for the Riemannian exponential function is known for any left- or right-invariant Riemannian metric. Moreover, the exponential group is no surjective for this group. If geodesics and distances are required, the problem has to be solved by numerical integration procedures of any of the descriptions of the differential system given in Chapter 4. Practical aspects of these numerical approaches will be treated in the following section.

Special Affine group, $\mathcal{SA}(d)$

It is the subset of transformations in $\mathcal{GA}^+(d)$ group that also preserve the volume of the objects. A matrix representation of the algebra are the $(d+1) \times (d+1)$ matrices with the last row equal to zero and trace equals to zero. Geodesics and distances also have to be computed by numerical integration for any invariant Riemannian metric.

5.2 Numerical solutions for the left-invariant Riemannian exponential function

In the previous Section, several closed-form expressions to calculate Riemannian exponential functions and the corresponding geodesics were gathered. However, most of these expressions are only valid for specific inertia operators. Moreover, for some groups, such as $\mathcal{GA}(d)$, there is no known expression to compute the Riemannian exponential under any metric. Although it is very common to elude in practice, in numerous cases there is not other solution than to solve geodesics by numerical integration.

There is a large number of numerical integration methods for ODEs and IVPs [Iserles et al. 2000] [Butcher 2000, 2008] [McLachlan and Quispel 2006] [Hairer et al. 2011]. Each numerical procedure with different levels of accuracy and complexity. Some of them are specifically designed to preserve the properties of dynamical systems such as the conserved quantities, the underlying group structure, the symplectic characteristics of Hamiltonians or the variational properties of the Lagrangian approaches.

In the Chapter 4 different ODEs describing the geodesic evolution were formulated. Although there are all equivalents, some of them are preferable in different scenarios. For example, the description in terms of chart coordinates given in Section 4.3.1 can result in a simple algebraic IVP which can be easily solved for some few particular groups and under certain inertia operators. Moreover, it is an approach valid even for non-matrix groups of finite dimension such as quotients groups. However the choice of appropriate charts is a difficult task. Solutions based on the Euler–Poincaré equation (Sections 4.3.3 and 4.3.4) are good choices for geometric integrators since the energy can be easily monitored and controlled by a quadratic expression (Section 4.4.1). The Optimal Control approach (Section 4.3.5) results in a system of differential equations with a symplectic structure derived from its Hamiltonian, Eq. (4.3.8). This structure of the coupled IVP system (4.3.11) can be exploited by specific integrators. The use of the symmetric formulation (Section 4.3.6) and the reduced to first order equation given in Section 4.5, are appropriate approaches to construct geometric integrators which inherently preserve the underlying group structure.

This thesis is not aimed to explore the vast families of numerical integrators and only explicit Euler and Taylor methods are explored as illustrative cases: two explicit Euler methods are presented to solve the differential equation (\dot{Q}) and the reconstruction step (recons); a fixed step-size Taylor series to solve the coupled IVP system formed by Eq. (4.3.7) and Eq. (recons); and an adaptive Taylor method where the time step-size is adapted in order to obtain solutions within a controlled tolerance.

5.2.1 Explicit additive Euler method

Let us recall the reconstruction equation for a left-invariant geodesic

$$\begin{aligned} \dot{Q}_{(t)} &= Q_{(t)} U_{(t)} \\ \text{with initial conditions } Q_{(0)} &= Q_0. \end{aligned} \tag{recons}$$

The explicit additive Euler method with a fixed step-size $\Delta = t_{j+1} - t_j$ is given by the iteration

$$Q_{(t_{j+1})} = Q_{(t_j + \Delta)} \approx Q_{j+1} = Q_j + Q_j \Delta U_{(t_j)} \tag{5.2.1}$$

where $U_{(t)}$ can be computed by different ways: in algebra coordinates by Eq. (4.3.5) following the approach given in Section 4.3.3; in its matrix representation by Eq. (4.3.7) given in Section 4.3.4; in terms of the co-state variable P using Eq.(4.3.11a); by the symmetric representation of the

geodesic as in Eq. (4.3.12a); or making use of the conserved quantity S as $U_{(t)} = \boldsymbol{\sigma} \left(Q_{(t)}^T S Q_{(t)}^{-T} \right)$ by following the reduction to an a first order equation proposed in Section 4.5. In the last case S is calculated at the beginning of the integration in terms of the initial conditions, by means of Eq. (4.4.2)

$$S = Q_0^{-T} \boldsymbol{\chi} (Q_0^{-1} V_0) Q_0^T.$$

Summarizing, the geodesic problem is completely described by specifying the initial geodesic point Q_0 and the initial velocity V_0 as $n \times n$ matrices. Moreover, the group where to solve must be specified by providing the basis of the algebra $\mathbf{B} = (\overline{B}_1, \dots, \overline{B}_k)$ and the metric properties must be detailed by a $k \times k$ symmetric and positive definite matrix \mathbf{m} representing the inertia operator at the identity. With these ingredients, compute the $n^2 \times n^2$ matrix $(\mathbf{B} \mathbf{m}^{-1} \mathbf{B}^T)$ representing the linear operator $\boldsymbol{\sigma}$ and the matrix $(\mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^{\dagger})$ representing its inverse operator $\boldsymbol{\chi}$. The left-invariant Riemannian geodesic can be computed by iterating with a (sufficiently small) step-size Δ the Eq. (5.2.1). Finally, the left-invariant Riemannian exponential function is obtained from the geodesic solution up to time $t = 1$, $\text{Exp}_{Q_0}(V_0) = Q_{(t=1)}$.

5.2.2 Explicit Lie–Euler method

In the previous additive method, generally the update $(Q + \Delta QU)$ might not belong to \mathbf{G} due to the curvature of the group set. The previous forward Euler integration method involving additive updates is well behaved in a vector space but, for an integration on a group structure, multiplicative updates are better to preserve the group structure.

If, in the evolution equation for $Q_{(t)}$, it is assumed that $U_{(t)}$ remains constant on each update step, then $\dot{Q} = QU$ can be solved at $t_{j+1} = t_j + \Delta$ in exact form by

$$Q_{(t_{j+1})} \approx Q_{j+1} = Q_j e^{\Delta U_{(t_j)}}. \quad (5.2.2)$$

Note the link with the group exponential function and the fact that if \dot{U} vanishes, then geodesics results in translated one-parameter subgroups.

With the scheme given by Eq. (5.2.2), each step is updated by group operations, and therefore Q_{j+1} inherently remains within the group set. This scheme can be used to integrate Eq. (dotQ) as well the IVP system (4.3.12).

5.2.3 Taylor series

The Taylor series integration method [Butcher 2008] [Ang and Park 2008] can be used to integrate the system defined by Eq. (4.3.7) for $U_{(t)}$ and the reconstruction step (recons) for $Q_{(t)}$. The Taylor series method approximates the time-dependent variables by their R -th order Taylor polynomial. Taylor series integrators can be formulated as interval methods guarantying the accuracy of the solutions [Jackson and Nedialkov 2002] [Moore et al. 2009]. In some approaches variable order formulations were performed to decrease the number of function derivatives evaluations [Barrio et al. 2005]. Other advantage is that the method allows the computation of dense output, obtaining the time evolution of the output. Assuming an ODE system $\dot{y}_{(t)} = \mathbf{f}(y_{(t)})$ with initial condition $y_{(t_0)} = y_0$, the approximate value of $y_j \approx y_{(t_j)}$ is computed as²

$$y_{(t_{j+1})} \approx y_{j+1} = y_j + \Delta_j \dot{y}_{(t_j)} + \frac{\Delta_j^2}{2!} \ddot{y}_{(t_j)} + \dots + \frac{\Delta_j^R}{R!} y_{(t_j)}^{(R)}$$

² $\dot{y}_{(s)}, \ddot{y}_{(s)}, \dots, y_{(s)}^{(j)}$ denote first, second, ..., j -th time derivative evaluated at instant s .

where $\Delta_j = t_{j+1} - t_j$. The method can be formulated as a variable step-size algorithm, where the time-interval Δ_j is appropriately chosen at each iteration to control the accuracy. For a large enough order, a large step-size can be used which could result in a reduction of the number of function evaluations.

A drawback of the Taylor series integrator is that it is not easy to implement as a general purpose algorithm. Some efforts were done in the past to automatically compute the high order derivatives of $\mathbf{f}(y(t))$ [Lara et al. 1999] [Jorba and Zou 2005]. Instead, in this work it is presented an explicit derivation of the expression for the m -th time derivatives of the involved variables.

Recalling Eq. (4.3.7)

$$\dot{U}_{(t)} = \boldsymbol{\sigma}[U_{(t)}^T, \boldsymbol{\chi}U_{(t)}],$$

where the transpose U^T will be written as the linear operation $\bar{\boldsymbol{\tau}}\bar{U} = \bar{U}^T$ for notational convenience (the linear operator $\boldsymbol{\tau}$ is represented by the commutation matrix K_{nn} [Abadir and Magnus 2005]). Using that the matrix commutation is a bilinear operator, $d_t[x, y] = [\dot{x}, y] + [x, \dot{y}]$, m -th time derivatives $U_{(0)}^{(m)}$ can be recursively computed as

$$\begin{aligned} U_{(t=0)} &= U_0 \\ \dot{U}_{(0)} &= \boldsymbol{\sigma}[\boldsymbol{\tau}U_{(0)}, \boldsymbol{\chi}U_{(0)}] \\ \ddot{U}_{(0)} &= \boldsymbol{\sigma}[\boldsymbol{\tau}\dot{U}_{(0)}, \boldsymbol{\chi}U_{(0)}] + \boldsymbol{\sigma}[\boldsymbol{\tau}U_{(0)}, \boldsymbol{\chi}\dot{U}_{(0)}] \\ \dddot{U}_{(0)} &= \boldsymbol{\sigma}[\boldsymbol{\tau}\ddot{U}_{(0)}, \boldsymbol{\chi}U_{(0)}] + 2\boldsymbol{\sigma}[\boldsymbol{\tau}\dot{U}_{(0)}, \boldsymbol{\chi}\dot{U}_{(0)}] + \boldsymbol{\sigma}[\boldsymbol{\tau}U_{(0)}, \boldsymbol{\chi}\ddot{U}_{(0)}] \\ &\vdots \\ U_{(0)}^{(m+1)} &= \sum_{j=0}^m \binom{m}{j} \boldsymbol{\sigma}[\boldsymbol{\tau}U_{(0)}^{(j)}, \boldsymbol{\chi}U_{(0)}^{(m-j)}]. \end{aligned} \quad (5.2.3)$$

Equation (5.2.3) resembles the general Leibniz rule. Then, via Taylor series it is possible to compute the dynamic of the variable $U_{(t)}$ given by Eq. (4.3.7)

$$U_{(t)} = \sum_{m=0}^{\infty} \frac{t^m}{m!} U_{(0)}^{(m)}, \quad (5.2.4)$$

valid up to a time smaller than the convergence radius of the previous series.

Similarly to the U evolution, a recursive procedure allows to compute the m -th time derivative of $Q_{(t)}$, at $t = 0$. From Eq. (recons)

$$\begin{aligned} Q_{(t=0)} &= Q_0 \\ \dot{Q}_{(0)} &= \dot{Q}_{(0)}U_{(0)} \\ \ddot{Q}_{(0)} &= \dot{Q}_{(0)}U_{(0)} + Q_{(0)}\dot{U}_{(0)} \\ \dddot{Q}_{(0)} &= \ddot{Q}_{(0)}U_{(0)} + 2\dot{Q}_{(0)}\dot{U}_{(0)} + Q_{(0)}\ddot{U}_{(0)} \\ &\vdots \\ Q_{(0)}^{(m+1)} &= \sum_{j=0}^m \binom{m}{j} Q_{(0)}^{(j)} U_{(0)}^{(m-j)} \end{aligned} \quad (5.2.5)$$

where $U_{(0)}^{(m-j)}$ is computed from Eq. (5.2.3). Finally, $Q_{(t)}$ can be computed via its Taylor series:

$$Q_{(t)} = \sum_{m=0}^{\infty} \frac{t^m}{m!} Q_{(0)}^{(m)}. \quad (5.2.6)$$

In order to use Eqs. (5.2.4) and (5.2.6) an estimation of their convergence radius is required. Let us start by given an upper bound of the Frobenius norm $\|U_{(0)}^{(m)}\|_F$ in terms of $\|U_{(0)}\|_F$. For any matrices U and $W \in \mathbb{M}_n$, the following upper bound is obtained

$$\|\sigma[\tau U, \chi W]\|_F \leq c \|\sigma\|_2 \|\tau\|_2 \|\chi\|_2 \|U\|_F \|W\|_F = C \|U\|_F \|W\|_F ,$$

where c is a constant bounding the norm of the commutator $\|[U, W]\|_F \leq c \|U\|_F \|W\|_F$, and $C = c \|\sigma\|_2 \|\tau\|_2 \|\chi\|_2$. Without any assumptions on matrices U and W , the minimal acceptable value for c is $\sqrt{2}$, see [Böttcher and Wenzel 2005, 2008] [Vong and Jin 2008]. The others involved constants $\|\sigma\|_2$, $\|\tau\|_2$ and $\|\chi\|_2$ are the spectral norm of the linear operators σ , τ and χ , respectively. They can be computed by the maximum singular value of their corresponding matrices

$$\begin{aligned} \|\sigma\|_2 &= \max \{\text{svd}(\mathbf{B}\mathbf{m}^{-1}\mathbf{B}^T)\} \\ \|\chi\|_2 &= \max \{\text{svd}(\mathbf{B}^{\dagger T}\mathbf{m}\mathbf{B}^\dagger)\} \\ \|\tau\|_2 &= \max \{\text{svd}(K_{nn})\} = 1 . \end{aligned}$$

In addition, it can be seen that the product $\|\sigma\|_2 \|\chi\|_2$ is no smaller than 1.

Bounding the terms of the recursive procedure (5.2.3) it is obtained

$$\|U_{(0)}^{(m)}\|_F \leq m! C^m \|U_0\|_F^{m+1} \quad (5.2.7)$$

which can be easily proved by induction. Replacing in Eq. (5.2.4) results:

$$\begin{aligned} \|U_{(t)}\|_F &\leq \sum_{m=0}^{\infty} \frac{t^m}{m!} \|U_{(0)}^{(m)}\|_F \\ &\leq \sum_{m=0}^{\infty} t^m C^m \|U_0\|_F^{m+1} . \end{aligned} \quad (5.2.8)$$

By using the root test, we conclude that the Taylor series given in Eq. (5.2.4) converges for any

$$t < \text{Cr}_{\text{bound}} = \frac{1}{C \|U_0\|_F} \quad (5.2.9)$$

Similarly to Eq. (5.2.7), upper bounds for $\|Q_{(0)}^{(m)}\|_F$ can be obtained

$$\begin{aligned} \|Q_{(0)}^{(m)}\|_F &\leq \left(\prod_{j=0}^{m-1} (1 + j C) \right) \|Q_0\|_F \|U_0\|_F^m \\ &< m! C^m \|Q_0\|_F \|U_0\|_F^m \end{aligned} \quad (5.2.10)$$

where the last strict inequality holds for $C > 1$. Comparing with Eq. (5.2.8) it is concluded that series (5.2.6) fulfils the same convergence bound than series (5.2.4) given in Eq. (5.2.9).

To exemplify the radius of convergence of the Taylor series, let us take the closed-form solution for $\mathcal{ST}(1)$ group (Eq. (5.1.1) in Section 5.1.4) and consider that given an initial velocity $U(\sigma, \tau)$ (for translational velocity $\tau \neq 0$), the evolution of the geodesic $\text{Exp}_I(t U(\sigma, \tau))$ remains analytic while denominator $(b^2 e^{2w} + 1)$ does not vanish. Replacing the values of w and b it is obtained that these equations have a pole at the complex time value

$$t_{\text{pole}} = \frac{\log \left(-\frac{\tau^2}{(\sqrt{\sigma^2 + \tau^2} - \sigma)^2} \right)}{2\sqrt{\sigma^2 + \tau^2}} = \frac{\log(|\tau|) - \log(\|U\|_F - \sigma) + i\frac{\pi}{2}}{\|U\|_F} , \quad (5.2.11)$$

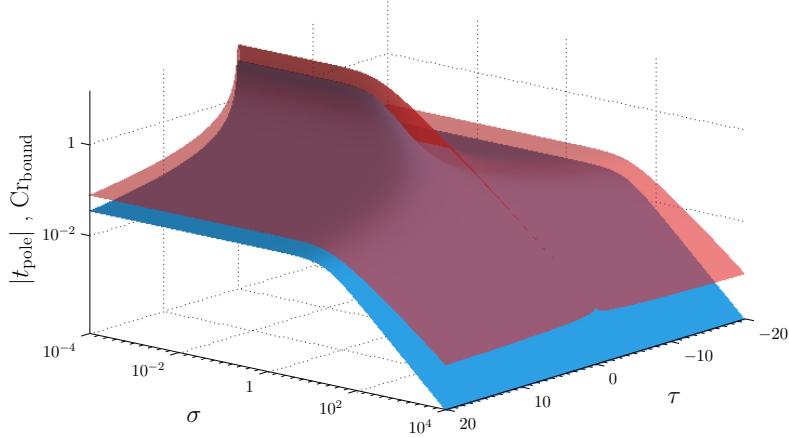


Figure 5.2.1. Red surface: $|t_{\text{pole}}|$ from Eq. (5.2.11) in terms of parameters σ and τ of the initial velocity $U(\sigma, \tau)$. Cyan surface: estimated lower bound Cr_{bound} of the convergence radius following Eq. (5.2.9).

for $\mathbf{i} = \sqrt{-1}$ and $\|U\|_F = \sqrt{\sigma^2 + \tau^2}$. Therefore the convergence of Taylor series (5.2.4) and (5.2.6) is guaranteed to times smaller than $Cr < |t_{\text{pole}}|$.

Figure 5.2.1 shows the value of t_{pole} from Eq. (5.2.11) and Cr_{bound} from Eq. (5.2.9) for different parameters σ and τ . It can be shown that Cr_{bound} (cyan surface) is almost one order of magnitude lower than t_{pole} (red surface) showing empirically that Cr_{bound} is a lower bound for the convergence radius of the Taylor series (5.2.4) and (5.2.6).

5.2.4 Fixed step-size Taylor method

In the iterative methods given in Sections 5.2.1 and 5.2.2, it were consider that either QU or U remains constant during each update step. The knowledge of the evolution for U allows to propose higher order methods that improve the accuracy at each integration step. In the iterative schemes based on Taylor series, a truncated Taylor expansion for U and Q is used at each update

$$\begin{aligned} U_{(t_{j+1})} \approx U_{j+1} &= U_j + \sum_{m=1}^R \frac{\Delta^m}{m!} U_j^{(m)} \\ Q_{(t_{j+1})} \approx Q_{j+1} &= Q_j + \sum_{m=1}^R \frac{\Delta^m}{m!} Q_j^{(m)}, \end{aligned}$$

where R is the degree of the truncated Taylor polynomial and the derivatives $U^{(m)}$ and $Q^{(m)}$ are computed using the iterative procedures given in Eqs. (5.2.3) and (5.2.5).

Due to numerical inaccuracies, the obtained U_{j+1} can departs from the algebra space. A re-projection can be performed by $\bar{U}_{j+1} = \mathbf{B}\mathbf{B}^\dagger \bar{U}_{j+1}$. Due to the truncation at the order R of the series, the computed Q_{j+1} departs from the group set. In this case, re-projection of Q_{j+1} usually involves non-linear functions or costly numerical procedures. For example, the projection (in the Frobenius norm sense) of a matrix onto $\mathbf{SO}(d)$ group set can be performed via the SVD decomposition [Gramkow 2001]. Also an optimization process have been proposed to project a matrix onto $\mathbf{SL}(d)$ [Lee et al. 2007].

5.2.5 Adaptive Taylor method

A further improvement can be obtained considering a variable step-size Taylor integrator where the integration error is controlled in terms of the step-size. A classical way to choose a guaranteed

step-size is to define bounds on the truncated coefficients of the Taylor expansion. Upper bounds for $\|U_0^{(m)}\|_F$ and $\|Q_0^{(m)}\|_F$ were given before in Section 5.2.3. These bounds are generally overestimated [Böttcher and Wenzel 2005] [Cheng et al. 2010] and step-sizes based on they result very small and are not used in the practice.

Alternatively, we propose to control the step-size at each iteration by monitoring the conserved quantities E and S obtained in Section 4.4. At each iteration, a geodesic segment is computed with a step-size that preserves E and S below prescribed thresholds. The step-size Δ is repeatedly reduced by a factor $0 < \alpha < 1$ until deviations of E and S are below prescribed thresholds thr_E and thr_S , respectively. At each iteration j , the reduction process starts from the value

$$\Delta^* = \sqrt[R]{\frac{R! \cdot \text{tol}}{\max \{\|U^{(R)}\|_{\max}, \|Q^{(R)}\|_{\max}\}}} \quad (5.2.12)$$

which satisfies that max-norms³ of the last terms in the truncated Taylor expansions are below a given tolerance specified in tol .

Once the reduction process is finished, the calculated segment is used to update the element along the requested geodesic. The update can be additive, as Taylor methods suggest, or multiplicative (by composing small geodesic segments), using Eq. (4.2.11). In the multiplicative update every geodesic segment is computed starting from the identity with the advantage that the matrices involved in the computation of S remain with a low condition number. Therefore, multiplicative update procedure behaves numerically better than additive updates.

The proposed algorithm is given in Alg. 1. Slight modifications allow the computation of dense outputs when the full geodesic segment is required. Additionally, the obtained U_Δ and Q_Δ after the computation of the U - and Q -series could be re-projected to the algebra and group set respectively as was noted in the fixed step Taylor method. Note that although the Q_j updates are done within the group structure, the computation of Q_Δ is performed additively and deviations from the group set can appear. In practice, re-projections increase the computation time and, based on our experience, it is enough to reduce the control parameters to improve the results.

5.2.6 Computing the Riemannian logarithm function

The Riemannian exponential function defines a geodesic in terms of its departing element and its initial velocity. However, in several applications it is desired to specify a geodesic segments in terms of its departing and arrival element (see Section 4.2.6). For this purpose the computation of the Riemannian logarithm is a key tool. The Riemannian logarithm is defined as an inverse function. This problem can be considered as a *two-point boundary value problem* (TPBVP) [Keller 1976] [Press et al. 2007] instead of an IVP. There are different approaches to solve this kind of problems such as *shooting* or *relaxation* methods.

The shooting method starts with an initial guess of the initial velocity, then integrates the geodesic up to time 1 and quantifies the discrepancy or error between the actual arrival point and the desired target. The computed discrepancy is used to smartly update the estimated initial velocity and the process continues until the discrepancy vanishes. The process can also be interpreted as a root-finding problem.

Given A and B the initial and final elements from a group \mathbf{G} to be joined by a geodesic segment, $\text{Log}_A(B)$ is formulated as the inverse problem of finding the initial velocity V such that $\text{Exp}_A(V) = B$. The following optimization problem is proposed to solve the inverse problem:

$$V^* = \underset{V \in T_A \mathbf{G}}{\text{argmin}} \|\text{Exp}_A(V) - B\|_F^2,$$

³ $\|M\|_{\max} = \max_i(|\bar{M}^i|)$

Algorithm 1 Adaptive Taylor method for compute left-invariant Riemannian exponential (multiplicative updates version).

```

function EXP(  $Q_0$  ,  $V_0$  ,  $\mathbf{B}$  ,  $\mathbf{m}$  )
    set  $R > 1$                                       $\triangleright$  order of the Taylor polynomial
    set tol
    set thrE and thrS                    $\triangleright$  for the estimation the initial step-size
    set  $0 < \alpha < 1$                           $\triangleright$  allowed deviations for  $E$  and  $S$  in each iteration
    compute  $\mathbf{B}\mathbf{m}^{-1}\mathbf{B}^T$             $\triangleright$  reduction factor
    compute  $\mathbf{B}^{\dagger T}\mathbf{m}\mathbf{B}^{\dagger}$ 
     $U_0 = Q_0^{-1}V_0$ 
    if  $[\chi U_0, U_0^T] = 0$  then
         $Q_1 = Q_0 \exp(U_0)$                        $\triangleright$  according to Section 4.3.4
    else
         $U_j = U_0$  ,  $Q_j = Q_0$  ,  $t = 0$ 
        while  $t < 1$  do
             $\{Q_{\Delta}, U_{\Delta}, \Delta\} = \text{GEODESICSEGMENT}(I, U_j, 1 - t)$ 
             $Q_j = Q_j Q_{\Delta}$                        $\triangleright$  multiplicate update step
             $U_j = U_{\Delta}$ 
             $t = t + \Delta$ 
        end while
         $Q_1 = Q_j$ 
    end if
    return  $Q_1$                                  $\triangleright$  return state of geodesic at ( $t = 1$ )  $\equiv \text{Exp}_{Q_0}(V_0)$ 
end function

procedure GEODESICSEGMENT(  $Q$  ,  $U$  ,  $s$  )
     $\triangleright$  Compute a controlled geodesic segment from  $Q$  up to time no greater than  $s$ 

 $E_0 = \text{trace}(\chi(U)^T U)$                    $\triangleright$  Eq. (4.4.1)
 $S_0 = Q^{-T} \chi(U) Q^T$                        $\triangleright$  Eq. (4.4.2). In particular,  $S_0 = U$  if  $Q$  is the identity
for  $m = 1$  to  $R$  do
    compute  $U^{(m)}$  using  $U, \dot{U}, \dots, U^{(m-1)}$            $\triangleright$  Eq. (5.2.3)
    compute  $Q^{(m)}$  using  $Q, \dot{Q}, \dots, Q^{(m-1)}$  and  $U, \dot{U}, \dots, U^{(m-1)}$        $\triangleright$  Eq. (5.2.5)
end for
with  $\|U^{(R)}\|_{\max}$  and  $\|Q^{(R)}\|_{\max}$ , compute  $\Delta^*$            $\triangleright$  Eq. (5.2.12), using tol
 $\Delta = \min(s, \Delta^*)$ 
loop
     $U_{\Delta} = U + \sum_{m=1}^R \frac{\Delta^m}{m!} U^{(m)}$ 
     $Q_{\Delta} = Q + \sum_{m=1}^R \frac{\Delta^m}{m!} Q^{(m)}$ 
     $E_{\Delta} = \text{trace}(\chi(U_{\Delta})^T U_{\Delta})$ 
     $S_{\Delta} = Q_{\Delta}^{-T} \chi(U_{\Delta}) Q_{\Delta}^T$ 
    if  $(|E_{\Delta} - E_0| < \text{thr}_E)$  and  $(\|S_{\Delta} - S_0\|_{\max} < \text{thr}_S)$  then
        return  $\{Q_{\Delta}, U_{\Delta}, \Delta\}$ 
    end if
     $\Delta = \alpha \Delta$                                  $\triangleright$  reduce step-size
end loop
end procedure

```

where the *Frobenius matrix norm* $\|\cdot\|_F$ was selected for simplicity.

Due to the metric invariance and using Eq. (4.2.11), the problem can be equivalently formulated as

$$V^* = A \left(\underset{U \in \mathfrak{g}}{\operatorname{argmin}} \| \operatorname{Exp}_I(U) - A^{-1}B \|_F^2 \right). \quad (5.2.13)$$

For most of the cases enumerated in Section 5.1 the $\operatorname{Exp}_I(\cdot)$ function can be algebraically inverted and there exist closed-form expressions for the $\operatorname{Log}_I(\cdot)$ function. Nevertheless, the algebraic expressions corresponding to groups such as $\mathcal{GL}^+(d)$, $\mathcal{SL}^+(d)$ cannot be algebraically inverted [Zacur et al. 2013]. For these cases, and those problems where $\operatorname{Exp}(\cdot)$ have to be computed numerically, a numerical procedure must be employed to compute $\operatorname{Log}(\cdot)$.

A gradient descent strategy is proposed in order to solve (5.2.13). As was sketched in Section 4.6, the corresponding descent direction can be computed using $\partial_X \|X - A^{-1}B\|_F^2 \Big|_{(X=Q)} = 2\overline{(Q - A^{-1}B)}^T$ and the derivatives of the discrepancy functional with respect to the components μ^i of U are

$$\partial_{\mu^i} \| \operatorname{Exp}_I(U) - A^{-1}B \|_F^2 = 2 \operatorname{trace} \left((\operatorname{Exp}_I(U) - A^{-1}B)^T \operatorname{Exp}_I(U) R_{(1)}^{B_i} \right),$$

where $R_{(1)}^{B_i}$ is the solution of Eq. 4.6.4 at $t = 1$ with $W = B_i$ and initial condition $R_{(0)} = 0$.

Alternatively, a finite differences scheme on $\| \operatorname{Exp}_I(U) - A^{-1}B \|_F^2$ can also be used to compute the descent direction [Press et al. 2007] [de Levie 2009]. Note that the derivatives have to be computed along perturbations of the components of U in the algebra \mathfrak{g}

$$\partial_{\mu^i} \| \operatorname{Exp}_I(U) - A^{-1}B \|_F^2 \approx \frac{\| \operatorname{Exp}_I(U + \varepsilon B_i) - A^{-1}B \|_F^2 - \| \operatorname{Exp}_I(U - \varepsilon B_i) - A^{-1}B \|_F^2}{2\varepsilon}$$

for an appropriate small ε . Once a descent direction is computed, a simple line-search procedure is performed to update U .

A major limitation of the descent strategy is that for some groups the stationary point reached after convergence may be not a global optimum, *i.e.* $\operatorname{Exp}_I(U) \neq A^{-1}B$, or even being a global optimum it may be not the smallest initial velocity which satisfies $U = \operatorname{Log}_I(A^{-1}B)$ [Zacur et al. 2013] (see also Section 6.3.1).

Whenever a right-invariant metric is used, the Riemannian logarithm can be computed from its left-invariant version as

$$\operatorname{Log}_X^{right}(Y) = -\operatorname{Log}_I^{left}(XY^{-1})X.$$

5.3 Results

The performance of the presented numerical methods to solve the Riemannian exponential function under a left-invariant metric was quantified in terms of accuracy with respect to a closed-form solution, the computation time, and the departure of the obtained solution from the group set.

5.3.1 Accuracy

Accuracy measure was quantified as the maximum of the absolute difference $\|Q_{\text{numeric}} - Q_{\text{exact}}\|_{\max}$, with Q_{exact} computed using the corresponding closed-form expression given in Section 5.1. Left-invariant $\operatorname{Exp}_I(\cdot)$ functions were computed for several random initial velocities.

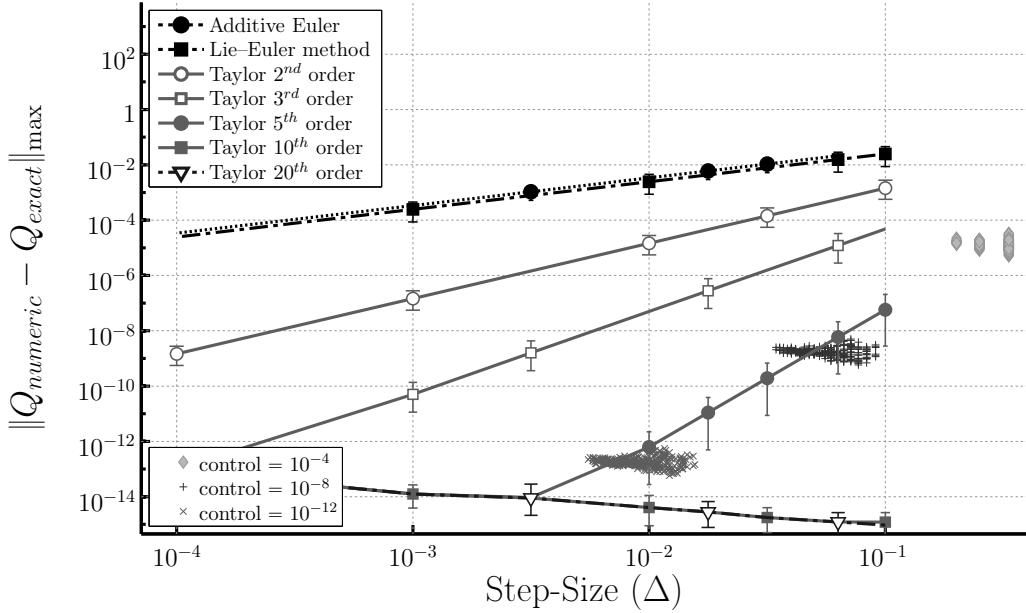


Figure 5.3.1. Accuracy measure for different integration methods on $\mathcal{GL}(3)$ group and initial velocities with norm $\|U_0\|_I = 1$. Lines represent average value among 1000 random initial velocities. Error bars depict 5-th/95-th percentiles. Scatter plots show results from Adaptive Taylor method under different values of the control parameters. Abscissa coordinates of scatter plots are the inverse of the number of integration steps required up to $t = 1$.

For each experiment, one thousand initial velocities were Exponentiated all having equal norm $\|U_0\|_I$ according to Eq. (4.2.10), *i.e.* they were drawn from a uniform distribution on a k -dimensional sphere in the algebra. Figure 5.3.1 shows results for $\mathcal{GL}(3)$ group and initial velocities with norm $\|U_0\|_I = 1$. The error bars depict the 5-th and 95-th percentiles of the accuracy measure and lines represent their average values. It can be seen that the numerical procedures are consistent as errors decrease with the step-size, at least within the floating point precision. Additive Euler and Lie–Euler methods present the worst accuracy. Better accuracies were obtained with fixed step-size Taylor methods, where the larger the order the smaller the error within the floating point precision. The increase of the error with respect to reductions of the step-size observed for 5-th, 10-th and 20-th order Taylor methods is due to accumulation of floating point round-off errors.

Results for Adaptive Taylor method with order $R = 5$ are shown as scatter plots using three different values of the error control parameters ($\text{tol} = \text{thr}_E = \text{thr}_S = \{10^{-4}, 10^{-8}, 10^{-12}\} = \text{control}$). Abscissa coordinates of the scatter plots represent, a “mean” step-size as the inverse of the number of integration steps needed up to $t = 1$.

Similar results were obtained for other transformation groups. For example, Figure 5.3.2 shows results for $\mathcal{SE}(3)$ group, but at this time $\|U_0\|_I = 10$, also for one thousand initial velocities. It can be observed that, for these larger magnitude of initial velocities, using fixed step-size Taylor method, the 10-th order is not enough to compute solutions within machine precision with large step-sizes. Then, either a smaller step-size or a larger Taylor order R is needed to get a similar accuracy.

Although the control parameters of the algorithm tol , thr_E and thr_S are designed to control the step-size with respect to the deviation for the conserved quantity E and the deviation for S , respectively, the accuracy measure remains at the same order of magnitude of these parameters. Then, as it was expected, the control parameters of the algorithm indirectly control the accuracy

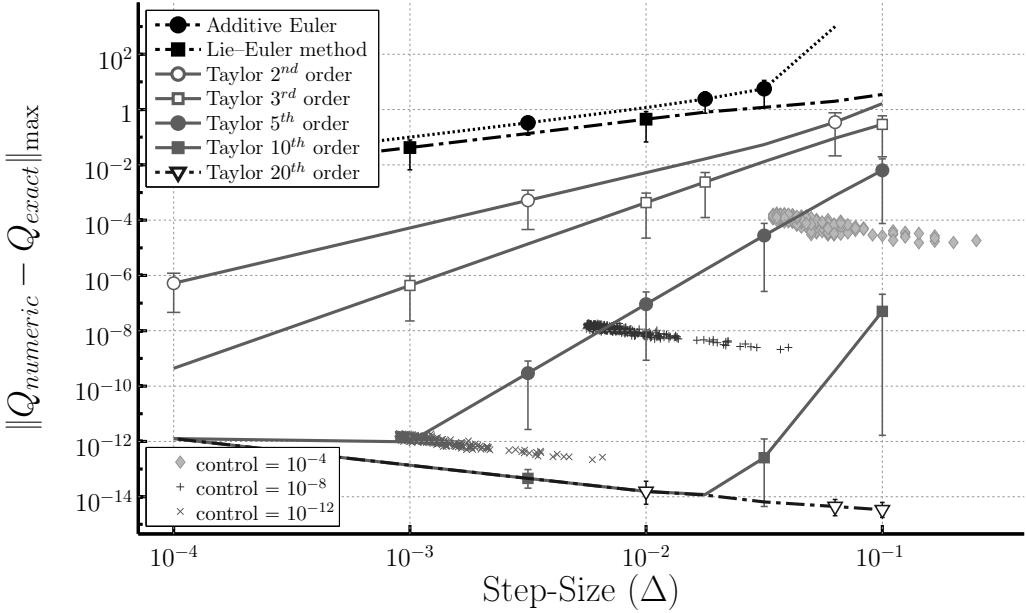


Figure 5.3.2. Accuracy measure for different integration methods on $\mathcal{SE}(3)$ group and initial velocities with norm $\|U_0\|_I = 10$. See caption of Figure 5.3.1 for more details.

of the results. In addition, for many initial velocities instances a larger “mean” step-size was estimated by the Adaptive Taylor method, compared to the fixed step Taylor of the same order, to get a similar final accuracy. Thus, reductions in the total number of integration steps were attained.

On one hand, the value of the step-size Δ in all numerical methods should be small enough to guarantee an accurate result. On the other hand, larger step-sizes require a smaller number of integration steps. Therefore, a smart strategy is to select the largest step-size Δ which provides a given accuracy performance. This is the working principle of the proposed Adaptive Taylor method.

5.3.2 Timing performance of Adaptive Taylor method

Timing performance was only evaluated for Adaptive Taylor method, given in Section 5.2.5, measuring the computation time in seconds with respect to the order R . The experiments were performed on a Intel i7-2600 processor with a clock frequency of 3.4 GHz using a MATLAB implementation of the Alg. 1. Algorithm parameters were set to $\text{tol} = \text{thre} = \text{thr}_S = 10^{-12}$ and the reduction factor $\alpha = 0.75$. Two thousand initial velocities were used in each experiment. Random initial velocities were uniformly drawn from the algebra, achieving three different values of the norm $\|U_0\|_I = \{0.1, 1, 10\}$, (see Eq. (4.2.10)).

Results on $\mathcal{ST}(1)$ group are shown in Figure 5.3.3 where, for each initial velocity norm, lines show average times and error bars depict 5-th/95-th percentiles. In general, the computation time of the Adaptive Taylor method decreases for low values of the order of the Taylor approximation. However, whenever a too small order is set, the computation time pronouncedly increases because very small values of step-size Δ are needed to fulfil the demanded error controls. These small values of Δ yield to many integration steps increasing the total computation time. It can be noted that for initial velocities with small enough magnitudes and large Taylor orders, only one integration step is needed to compute the Riemannian exponential. Accordingly no dispersion in the computation time was obtained, even for initial velocities having different directions. A

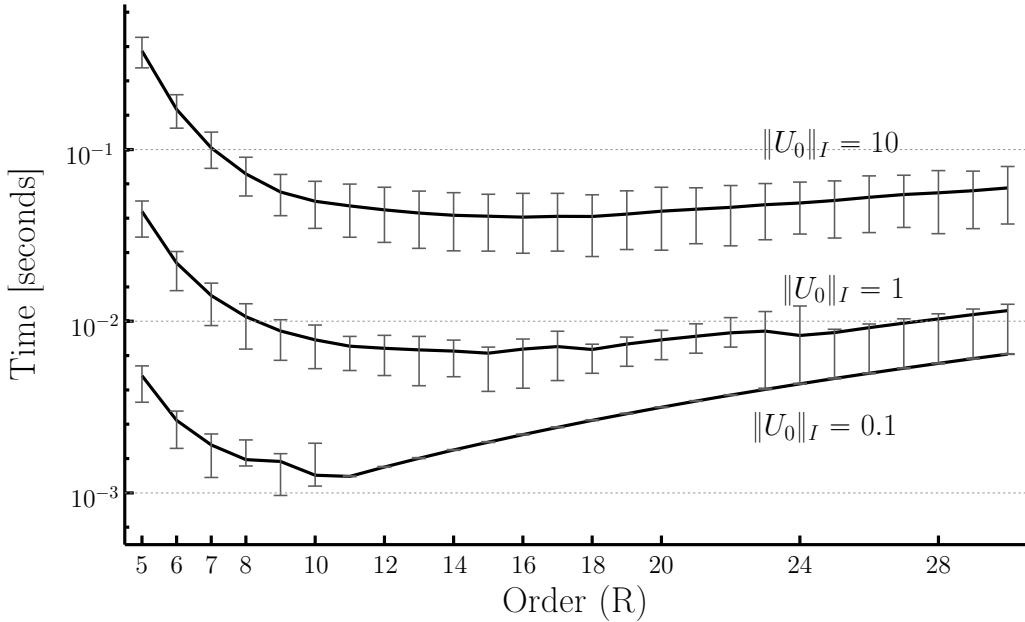


Figure 5.3.3. Computation time required to calculate the Riemannian exponential with the Adaptive Taylor method in terms of the Taylor order R . Results on $\mathcal{ST}(1)$ group and for different magnitudes of random initial velocities. Lines: average time (in seconds). Error bars: 5-th/95-th percentiles.

similar behavior was observed for several tested transformation groups. For example, timing performance for $\mathcal{SE}(5)$ group is shown in Figure 5.3.4.

According to our experience, a suitable value of the Taylor order R to get minimum computation time is more sensitive to the magnitude of the initial velocities $\|U_0\|_I$ than to the group itself or the parameters of the algorithm such as thr_E or thr_S . The larger the magnitude of the initial velocities, the larger value of the Taylor order to minimize the computation time. As a rule of thumb and according to our experience, a value of $R = \lceil 15\sqrt[6]{\|U_0\|_I} \rceil$ provides a good choice in terms of timing performance.

5.3.3 Departures from the group set

In order to evaluate if the solutions provided by the numerical algorithms belong to the group set, an experiment was performed on $\mathcal{SA}(3)$ group. The elements of this group are constrained to have determinant equal to 1. Again, thousand random initial velocities with norm $\|U_0\|_I = 1$ on the algebra were used in this experiment. The performance measure was the deviation $|\det(Q_{\text{numeric}}) - 1|$. Results are shown in Figure 5.3.5. As it was expected the best performance was obtained with Lie–Euler method because for this method each update is performed within the group structure.

5.4 Application examples

A few illustrative examples are given below to illustrate some applications of invariant Riemannian geodesics.

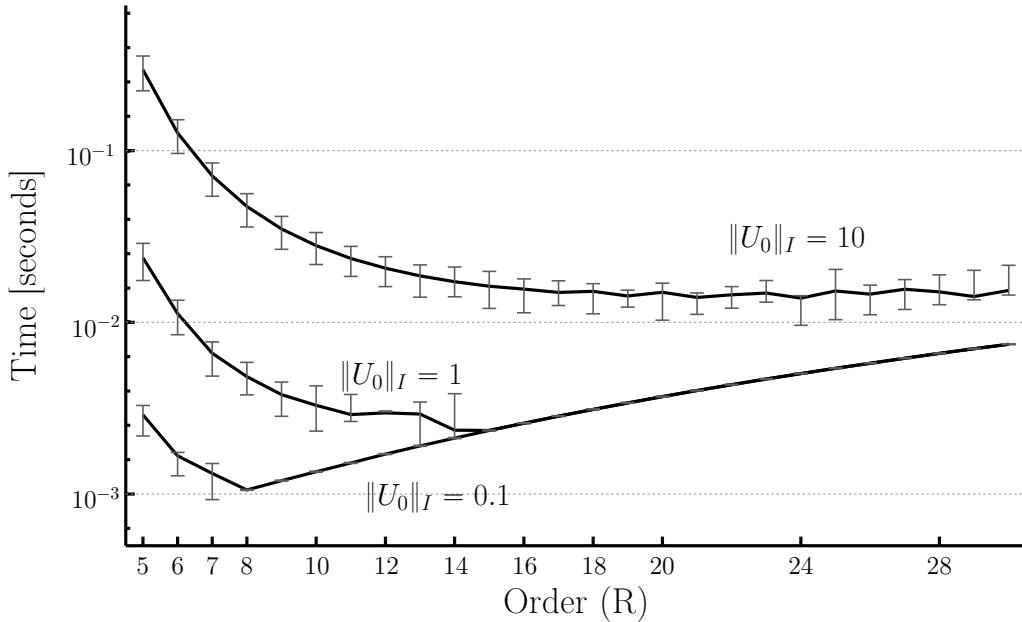


Figure 5.3.4. Timing performance of Adaptive Taylor integration method on $\mathcal{SE}(5)$ group.
See also caption of Figure 5.3.3.

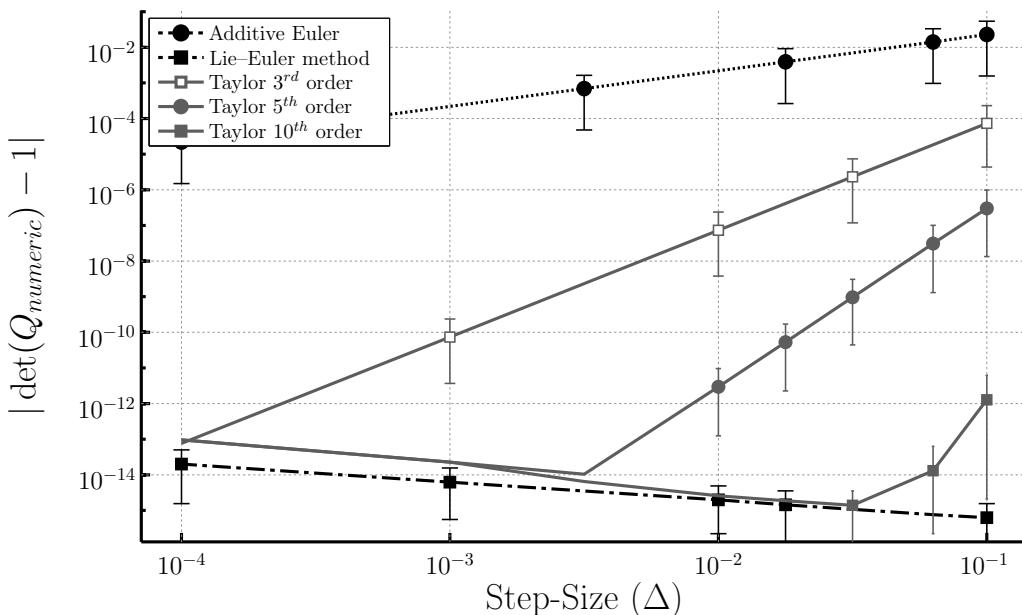


Figure 5.3.5. Deviations from the unity of the determinant for different integration methods at different step-sizes on $\mathcal{SA}(3)$ group. One thousand random initial velocities with norm $\|U_0\|_I = 1$ were used.

5.4.1 Interpolating spatial transformations

In this example, invariant Riemannian geodesics are used to interpolate between two transformations or between two transformed realizations of an object. The methodology provides a time-parameterized continuous path between two transformations which can be used, for example, in computer graphics for key-framing animation [Shoemake 1985] [Alexa 2002] [Begelfor and Werman 2005] [Li and Hao 2006].

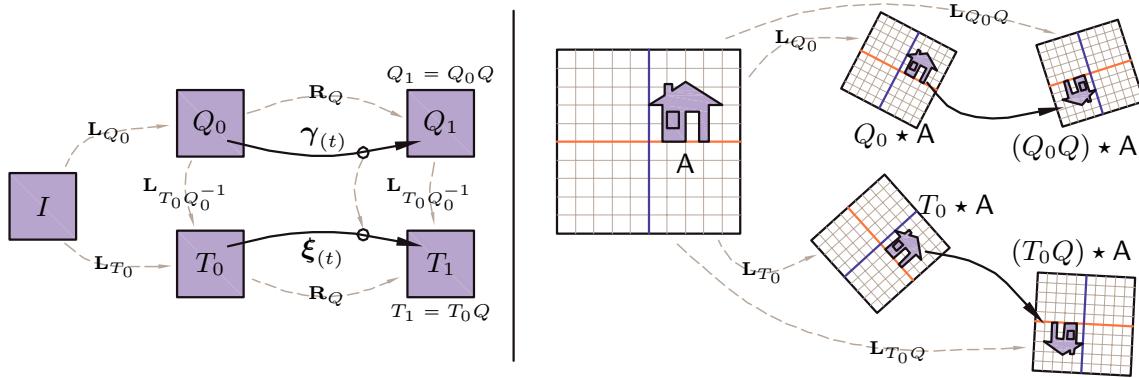


Figure 5.4.1. Illustration of the framework to interpolate spatial transformations.

The interpolation framework is depicted in Figure 5.4.1. In the left-panel it is illustrated the problem of finding the geodesic γ that interpolates two spatial transformations Q_0 and $Q_1 = \mathbf{R}_Q(Q_0) = Q_0Q$. The geodesic interpolating path is given by $\gamma(t) = \text{Exp}_{Q_0}(t \text{ Log}_{Q_0}(Q_1))$. Due to the left-invariance, the geodesic between any pair of left-translated instances $T_0 = \mathbf{L}_{T_0Q_0^{-1}}(Q_0)$ and $T_1 = \mathbf{L}_{T_0Q_0^{-1}}(Q_1)$ will be $\xi(t) = \mathbf{L}_{T_0Q_0^{-1}}(\gamma(t))$. In this framework the following distances are also preserved: $\text{distance}(Q_0, Q_1) = \text{distance}(I, Q) = \text{distance}(T_0, T_1)$. Right-panel shows how these transformations act on an object A defined in the Euclidean space \mathbb{R}^2 . To emphasize that the spatial transformations act over the whole ambient space a grid is also shown in the figure representing the ambient space. Note that the interpolation between transformations acts directly to the ambient space and is independent of the objects defined on it and the intermediate positions of the objects do not depend on them, because they are guided by the intermediate positions of the transformations solely. However, different objects may define different inertia operators and therefore different Riemannian metrics to take their geometries into account. This approach is similar to the formalism used to analyze the kinematics of a physical object, where rotations act on the space but the space is endowed by the inertia tensor of the physical object computed from its mass distribution. In addition, using the presented framework, a distance between the objects $Q_0 * A$ and $Q_1 * A$ may be inherited from the distance between the transformation mapping them.

A first experiment was designed to illustrate the results of interpolate spatial transformations when using different groups. Two instances to be interpolated were arbitrarily chosen: Q_0 as the identity transformation; and Q_1 as a clockwise rotation of 135 degrees followed by a translation of 4 units in the horizontal axis and 2 units along the vertical axis. Both transformations are elements from $\mathcal{SE}(2)$ group. In all interpolations the metric was selected as the one given by $\langle U, U \rangle_I = \text{trace}(U^T U)$, where U is the matrix representation of the velocity located at the identity.

The geodesic interpolation on $\mathcal{SE}(2)$ is shown in Figure 5.4.2(a). In this case the interpolating path corresponds to a linear interpolation of the transformation parameters t_x, t_y, θ matching both instances (see the attached panel with the time-course of the parameters describing the transformation: θ denotes the rotation angle; t_x and t_y denote translations; and s and \det

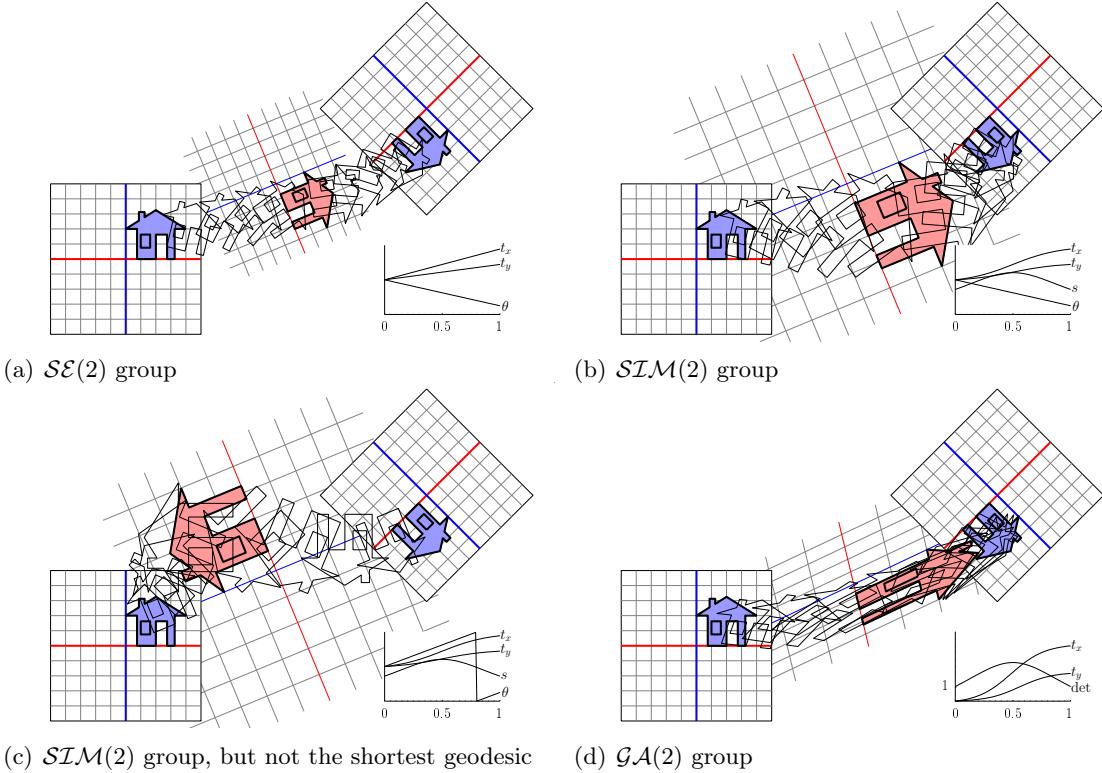


Figure 5.4.2. Geodesic interpolation between two instances of $\mathcal{SE}(2)$ on different groups. The time course of some parameters of the transformations are also shown in arbitrary units.

denote size and volume change, respectively; vertical axes of these panels are in arbitrary units). Figures 5.4.2(b)-(c) show two possible geodesics on $\mathcal{SLM}(2)$. Both curves are solutions of Eq. (5.2.13) both being global minima. This illustrates the fact that several geodesics can be solutions of the inverse problem for the Riemannian exponential function. The shortest curve, shown in Figure 5.4.2(b), defines the distance between the transformations. It is interesting to see that, although in both, the initial and final transformations, the scale factor equals to 1, the geodesics (b)-(c) change the object scale in order to minimize the curve lengths (this effect will also be discussed in Section 5.4.4). A similar behavior can be seen in Figure 5.4.2(d) for $\mathcal{GA}(2)$ group where determinant of the transformations changes along time. Note that the interpolating paths, and also the distances between both transformations, depend on the underlying group. The resulting distances are: 5.57 for $\mathcal{SE}(2)$; 4.83 (case (b)) and 6.57 (case (c)) for $\mathcal{SLM}(2)$; 4.66 for $\mathcal{SA}(2)$ (not shown in the figure); 4.42 for $\mathcal{GA}(2)$. It can be said that for self-contained subgroups, the larger the dimension of the subgroup, the smaller the distance between a pair of transformations.

Figure 5.4.3 shows similar interpolating paths as before but now obtained when using the following transformation groups: Projective group $\mathcal{PGL}(2)$; Möbius group $\mathcal{MO}(2)$; and two examples of centered transformations from $\mathcal{SO}(2) \times \mathcal{T}(2)$ group. In the cases of centered transformations, the interpolation paths depend on the choice of the rotation center (indicated by “+” markers in figure). This transformation model should be only used when a well defined center can be determined, otherwise the result shows a strong dependence on the arbitrary choice of the center.

Transformations can also be interpolated without using the left-invariant geodesic framework

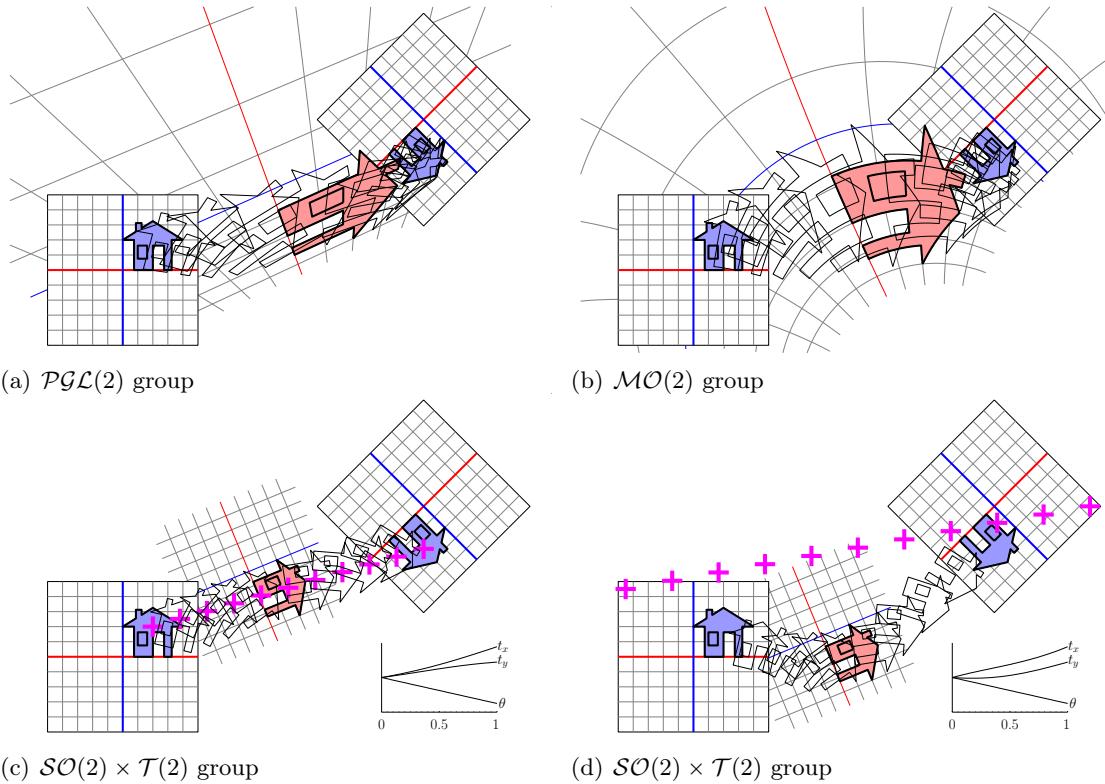


Figure 5.4.3. Interpolation curves using $\mathcal{P}\mathcal{G}\mathcal{L}(2)$, $\mathcal{M}\mathcal{O}(2)$ and centered transformation ($\mathcal{S}\mathcal{O}(2) \times \mathcal{T}(2)$) groups. In the cases (c) and (d) “+” markers indicate the center of rotation.

shown in Figure 5.4.1. For example, Figure 5.4.4(a) shows an interpolating geodesic on $\mathcal{SE}(2)$ group but this time using a right-invariant metric. Instead of using Riemannian geodesics, trajectories defined by one-parameter subgroup (or its translations) might also be used for interpolation [Alexa 2002] [Hawkins and Grimm 2005], although it does not always exist (see Section 4.2.5). Figure 5.4.4(b), shows the interpolating curve computed by using the group exponential function $\gamma_{(t)} = Q_0 \exp(t Q_0^{-1} Q_1)$ where no metric is specified. This curve is computationally simpler to compute [Moler and Van Loan 2003] [Al-Mohy 2010] than a Riemannian geodesic. The elements along the curve belong to the smallest group containing both transformations, in this example $\mathcal{SE}(2)$ group.

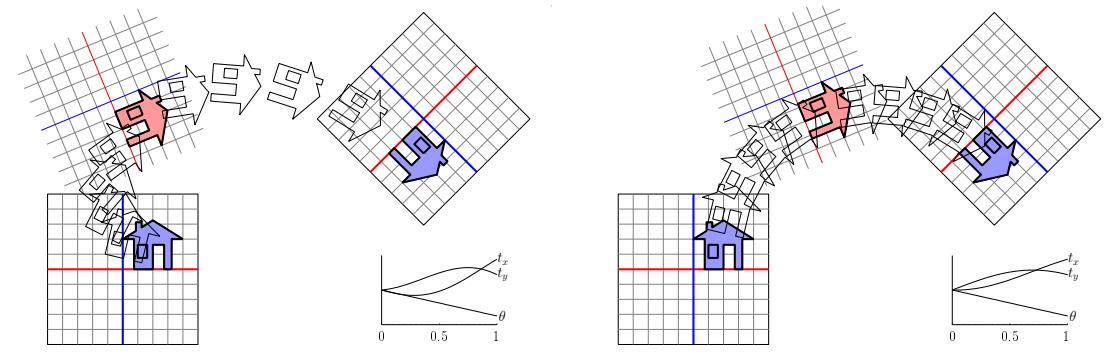


Figure 5.4.4. Alternative interpolation curves to the framework illustrated in Figure 5.4.1.

A second experiment was designed to illustrate the geodesic interpolation between two transformations from $\mathcal{GL}^+(2)$ which cannot be joined by a one-parameter curve. The instances to be interpolated are $Q_0 = I$ and Q_1 defined as the composition of a 180 degree rotation followed by a scale factor of 2 along the horizontal axis. Figure 5.4.5 shows the action of a left-invariant Riemannian geodesic on $\mathcal{GL}(2)$ group interpolating Q_0 and Q_1 . In this particular example there are two possible geodesics with minimal length joining Q_0 and Q_1 . In addition to the path shown in Figure 5.4.5 the other geodesic starts with a rotation in the opposite direction.

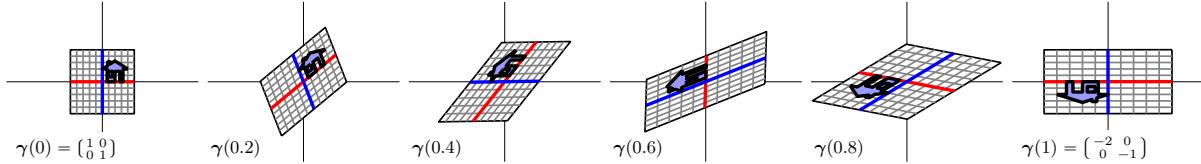


Figure 5.4.5. Action of the left-invariant geodesic on $\mathcal{GL}(2)$ interpolating Q_0 and Q_1 which cannot be joined by a translated one-parameter subgroup curve.

5.4.2 Intrinsic sample mean of spatial transformations

In the recent years, there is a considerable growing interest in performing statistical analysis on different manifolds. Many of this interest came from the areas of computer vision, image analysis and medical imaging. For a recent review gathering several applications see [Lui 2012].

Older approaches simply treat these spaces as embedded in a Euclidean space and ignoring their geometrical aspects. In that setting, the mean element is computed as the arithmetic average of the embedded coordinates and then it is projected back onto the closest Euclidean point of the embedded manifold.

Another school of thought defines a coordinate system on the embedded surface and describe the samples in these intrinsic coordinates instead of with the Euclidean embedded coordinates. Thus, statistical analyses are performed with classical tools but on the intrinsic coordinates. That approach is commonly solved by the ISometric MAPping (ISOMAP) and Local Linear Embedding (LLE) techniques [Tenenbaum et al. 2000] [Roweis and Saul 2000] which typically require a training step. However, the intrinsic coordinate system on the embedded surface usually inherit the metric structure from the embedding Euclidean space ignoring the metric structure of the original space.

Here we will investigate about *intrinsic means* which is an extension of the average concept defined for elements from a vector space. The elements have to be considered as belonging to a manifold endowed with a metric structure. A general description of the intrinsic sample mean for Riemannian manifolds is given in [Pennec 1999]. The specific case of the manifold of orthogonal matrices $\mathbf{SO}(d)$ has been studied by [Moakher 2002] [Hartley et al. 2013] An algorithm for computing the intrinsic means on the space of positive definite matrices has been presented in [Fletcher and Joshi 2007]. References of the subject dealing with other geometric spaces and manifolds can be found in [Buss and Fillmore 2001] [Chikuse 2003] [Vaillant et al. 2004] [Arsigny et al. 2007] [Subbarao 2008] [Mardia and Jupp 2009] [Vandereycken 2010] [Sommer et al. 2010] [Huckemann et al. 2010] [Pennec and Arsigny 2012] [Bhattacharya and Bhattacharya 2012] [Pennec 2013] [Jeuris et al. 2012] [Freifeld 2013]. For a more rigorous treatment, including statistical consistency of the intrinsic sample mean can be found in [Bhattacharya and Patrangenaru 2003, 2005].

In Section 1.3.1 it was sketched that the average \hat{m} of a sample of elements $\{\bar{x}_i\}_i^N$ from a

Euclidean space \mathbb{E} satisfies the variational condition

$$\hat{m} = \frac{1}{N} \sum_{i=1}^N \bar{x}_i = \operatorname{argmin}_{\bar{y} \in \mathbb{E}} \frac{1}{N} \sum_{i=1}^N \|\bar{x}_i - \bar{y}\|^2,$$

where $\|\cdot\|$ is the Euclidean norm in \mathbb{E} . Fréchet [Fréchet 1948] generalized the concept of means to general metric spaces by defining for a set of elements $\{\chi_i\}_{i=1}^N$ from a metric space \mathcal{M} , with a distance function $\operatorname{distance}(\cdot, \cdot) : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$,

$$\langle \{\chi_i\} \rangle = \operatorname{argmin}_{y \in \mathcal{M}} \sum_{i=1}^N \operatorname{distance}(\chi_i, y)^2. \quad (5.4.1)$$

The element $\langle \{\chi_i\} \rangle$ (from now on denoted as $\langle \chi \rangle$ for short), solution of the problem (5.4.1), is called *intrinsic sample mean*⁴. As the concept of intrinsic sample mean was extended from the average of samples from a vector space structure, whenever samples are vectors the same notation will be used, *i.e.* $\langle \bar{x} \rangle = \frac{1}{N} \sum_i^N \bar{x}_i$ with $\bar{x}_i \in \mathbb{E}$. Concerns about the solutions of the problem (5.4.1) were considered by Karcher [Karcher 1977], who proposed to extend the above definition of the intrinsic sample mean to local minima as opposed to global minimal. From the practical viewpoint, it is common to assume that samples are “sufficiently concentrated” [Buss and Fillmore 2001] [Hartley et al. 2013] such that the intrinsic sample mean is unique. An important remark about the intrinsic sample mean is that, under some mild assumptions about the structure of \mathcal{M} and the underlying distribution, the intrinsic sample mean is a strongly consistent estimator of the intrinsic mean of the distribution from where the samples were drawn [Bhattacharya and Patrangenaru 2003]. Hereinafter, when there is no danger of confusion, the word “sample” can be omitted.

Particularizing for a Riemannian manifold, where $\operatorname{distance}(\chi, y) = \|\operatorname{Log}_y(\chi)\|_y$, problem (5.4.1) results

$$\langle \chi \rangle = \operatorname{argmin}_{y \in \mathcal{M}} \sum_{i=1}^N \left\| \operatorname{Log}_y(\chi_i) \right\|_y^2,$$

whose solution (or local solutions) $\langle \chi \rangle$ satisfies

$$\sum_i^N \operatorname{Log}_{\langle \chi \rangle}(\chi_i) = 0, \quad (5.4.2)$$

where all vectors $\operatorname{Log}_{\langle \chi \rangle}(\chi_i)$ belong to the same tangent space $T_{\langle \chi \rangle} \mathcal{M}$ and the usual addition can be performed⁵. The following gradient descent procedure on \mathcal{M} can be performed to compute the intrinsic mean [Pennec 2006a]:

$$y_{k+1} = \operatorname{Exp}_{y_k} \left(\varepsilon \sum_i^N \operatorname{Log}_{y_k}(\chi_i) \right), \quad (5.4.3)$$

⁴ The word “sample” makes reference to that $\langle \cdot \rangle$ is computed for a finite set of elements. An *intrinsic mean* can be similarly computed for a probability distribution $\operatorname{Pr}(\cdot)$ on \mathcal{M} as the minimum

$$\min_{y \in \mathcal{M}} \int_{\chi \in \mathcal{M}} \operatorname{distance}(\chi, y)^2 \operatorname{Pr}(\chi) d\chi$$

but in general a *probability measure* defining $d\chi$ have to be assigned to \mathcal{M} .

⁵ In the same way that the vector structure was relaxed from the average based definition of the mean to the distance based definition of the intrinsic mean, [Pennec and Arsigny 2012] proposes to relax the Riemannian and metric structure keeping only the structure of an affine connected space to define the *exponential barycenters* as those points fulfilling Eq. (5.4.2). In that setting, there is no curve lengths, no distances, no variances, only geometrical transport of velocities between tangent spaces (see Section 4.2.11).

where ε is a small enough step such that

$$\sum_i \text{distance}(\chi_i, y_{k+1}^\varepsilon)^2 < \sum_i \text{distance}(\chi_i, y_k)^2$$

which can be determined by a backtracking line-search procedure. It can be shown that the sequence y_k from Eq. (5.4.3) converges to $\langle \chi \rangle$ (for appropriate ε).

An alternative iteration process was proposed in [Ho et al. 2013]. The algorithm is based on the observation that, given elements $\{\bar{x}_i\}_i^N$ from a Euclidean space \mathbb{E} , the average \hat{m} can be computed recursively by

$$\hat{m}_k = \frac{(k-1)\hat{m}_{k-1} + \bar{x}_k}{k} = \hat{m}_{k-1} + \frac{1}{k}(\bar{x}_k - \hat{m}_{k-1}), \quad (5.4.4)$$

with $\hat{m}_1 = \bar{x}_1$. At each iteration \hat{m}_{k-1} is updated such that \hat{m}_k results equal to $\langle \{\bar{x}_i\}_{i=1}^k \rangle$, and after N iterations, \hat{m}_N results equal to $\langle \bar{x} \rangle$. Now, given a sample $\{\chi_i\}_i^N$ from a Riemannian manifold, the adaptation of the recursive process given in Eq. (5.4.4) is given by

$$m_k = \text{Exp}_{m_{k-1}} \left(\frac{1}{k} \text{Log}_{m_{k-1}} (\chi_k) \right). \quad (5.4.5)$$

Both recursions, Eq. (5.4.4) and (5.4.5), can be interpreted geometrically as moving an appropriate distance away from m_{k-1} towards χ_k along the “straightest path” joining them. In general, in the manifold setting, m_N depends on the ordering of the instances $\chi_1, \chi_2, \dots, \chi_N$. However, it was shown in [Ho et al. 2013] that under some mild conditions on the distribution generating χ_i , as the number of samples goes to infinite, the recursive estimator m_k converges to the intrinsic mean of the underlying distribution irrespectively of the order. The computation of the recursive estimator does not require an optimization stage and provides a good starting point y_0 for the descent procedure (5.4.3).

Two experiments were performed in order to illustrate the properties of the intrinsic mean of spatial transformations when using either left- or right-invariant Riemannian distances. The experiments were performed using simulated spatial transformations from $\mathcal{SIM}(2)$ group, but any other group could be used without loss of generality.

- **Left-invariant case:**

Let $\{Q_i\}_i^3$ be a set of three spatial transformations from $\mathcal{SIM}(2)$ group. One way to illustrate each Q_i is by means of its action on a reference object which will be called *template* (denoted as T) and is illustrated in red in Figure 5.4.6. Therefore, objects $Q_i \star T$ in the left-panel of Figure 5.4.6 illustrate the transformations Q_i .

First, let us consider the left-invariant Riemannian metric defined in the algebra by $\langle\langle U_1, U_2 \rangle\rangle = \text{trace}(U_1^T U_2)$. The mean of $\{Q_i\}_i$ with the given metric is defined by Eq. (5.4.1) and $\langle Q \rangle^L$ can be computed with the iterative procedure from Eq. (5.4.3). The resulting $\langle Q \rangle^L$ is illustrated in Figure 5.4.6 by its action on the template, $\langle Q \rangle^L \star T$ (blue object).

In several scenarios, the observed spatial transformations may be left-translated by a common transformation M , which is usually uninformative. One example is when observing transformation using a different device. In this scenario, the transformation instances are given by $P_i = MQ_i$. These new instances are illustrated, again by means of their action on T , in the right-panel of Figure 5.4.6. The intrinsic mean of these instances, $\langle P \rangle^L$, are computed and its action on T is illustrated in blue in the right-panel of Figure 5.4.6.

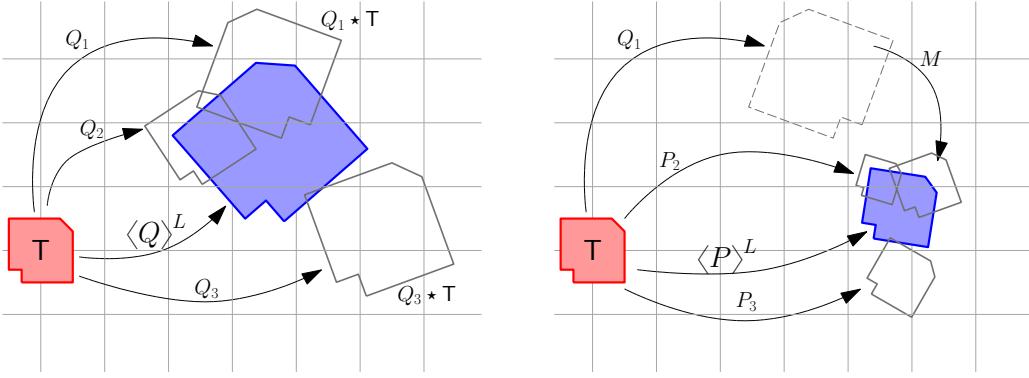


Figure 5.4.6. Illustration of the intrinsic mean of spatial transformations on $\mathcal{SLM}(2)$ with a left-invariant Riemannian metric. Comparing left and right panels, it can be checked the left-invariance of the intrinsic mean.

It can be easily checked the following relation between both intrinsic means:

$$\langle P \rangle^L = \langle MQ_i \rangle^L = M \langle Q \rangle^L.$$

That is, the intrinsic mean when using a left-invariant Riemannian metric does not depend on the reference system used for measuring the instances. This property is very important as it allows the assessment of statistic properties irrespectively of how the instances were measured (whenever the change of reference system is a member of the group under the metric is invariant).

- **Right-invariant case:**

A similar experiment can be designed but now considering a right-invariant Riemannian metric (again defined by $\langle\langle U_1, U_2 \rangle\rangle = \text{trace}(U_1^T U_2)$). Figure 5.4.7 illustrates the action of the three spatial transformations Q_i on the template T (shown in red color), as well as the intrinsic mean for the right-invariant distance $\langle Q \rangle^R$ (blue color).

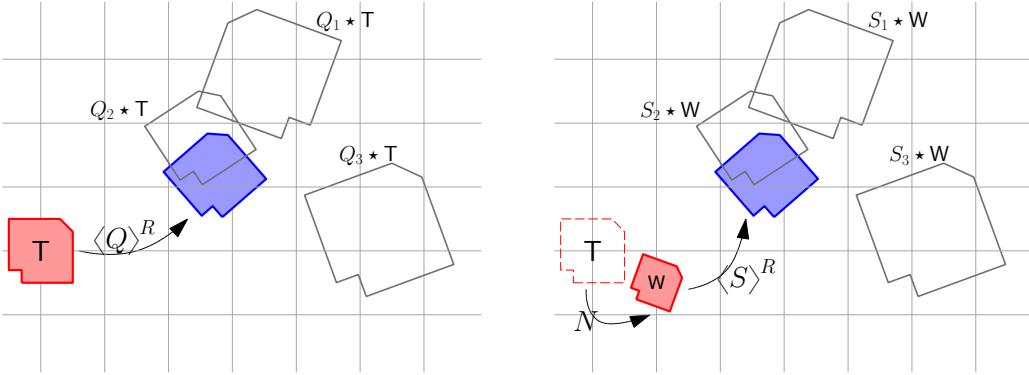


Figure 5.4.7. Illustration of the intrinsic mean of spatial transformations on $\mathcal{SLM}(2)$ with a right-invariant Riemannian metric. In this case it can be checked that the right-invariance of the metric provides a result independently of the selection of the template.

In some applications one is interesting in the assessment of statistical differences of a set of geometrical objects. Sometimes these objects are described by means of spatial transformations of a given template, such as in computational anatomy. In this context, the choice of the template should not affect the main conclusion of the study. The objects

initially described by the template T as the set $\{Q_i\}$, such that each object is $Q_i \star \mathsf{T}$, could be also described by using an alternative template $\mathsf{W} = N \star \mathsf{T}$, such that N is the spatial transformation relating both templates. Thus, in terms of W , the objects are described by $S_i \star \mathsf{W} = Q_i N^{-1} \star \mathsf{T}$ and the set of transformations $\{S_i\}_i = \{Q_i N^{-1}\}_i$ is an alternative description of the objects. The intrinsic mean of the set $\{S_i\}$ should be related to $\{Q_i\}$. This condition is achieved when using a right-invariant Riemannian metric, and it is easily seen that

$$\langle S \rangle^R = \langle Q_i N^{-1} \rangle^R = \langle Q \rangle^R N^{-1}.$$

The main conclusion from this experiment is that right-invariant distances are a good choice in applications where statistics of objects are computed by mean of spatial transformations descriptors.

5.4.3 Intrinsic filtering

In this application, a spatial filtering with a Gaussian kernel is performed on spatial transformations distributed on an open spatial domain $\Omega \subset \mathbb{R}^2$. The data to be filtered in this example is a field $J(\mathbf{x}) : \Omega \rightarrow \mathbf{GL}^+(2)$ of spatial derivatives of an invertible and differentiable mapping $\Phi : \Omega \rightarrow \Omega$. Figure 5.4.8(a) shows a deformation field Φ where the glyphs represent the local deformation defined by a 2×2 matrix with positive determinant

$$J(\mathbf{x}) = \begin{pmatrix} \partial_1 \phi^1|_{\mathbf{x}} & \partial_2 \phi^1|_{\mathbf{x}} \\ \partial_1 \phi^2|_{\mathbf{x}} & \partial_2 \phi^2|_{\mathbf{x}} \end{pmatrix} \in \mathbf{GL}^+(2).$$

Regarding practical details, the deformation mapping Φ was simulated from a continuous and differentiable piecewise cubic spline velocity field with vanishing boundary conditions. $\Phi(\mathbf{x})$ was obtained as the time integration up to time 1 of a particle on an initial point \mathbf{x} (see Section 2.5.4). Moreover, $J(\mathbf{x})$ can also be computed integrating up to time 1 the local deformation guided by the Jacobian of the velocity field, Section 2.5.4. As was argued in Chapter 2, this parametrization guarantees (with an appropriate integration scheme) that all Jacobian matrices $J(\mathbf{x})$ belong to $\mathbf{GL}^+(2)$.

Figure 5.4.8(b) shows a glyph representation of the matrix-valued data in a discrete grid which corresponds to the field $J(\mathbf{x})$ in *Lagrangian coordinates* [Holm et al. 2009, p.335]. The color of each cell represents the local volume change in logarithm units, $\log_{10}(\det(J(\mathbf{x})))$.

In order to simulate a noisy observation of the matrix-valued data $J(\mathbf{x})$, the deformation mapping Φ (Figure 5.4.8(a)) was composed with a random deformation with circular boundary conditions. The “noisy” Jacobian matrices from the perturbed deformation are shown in Figure 5.4.8(c).

The matrix-valued data of Figure 5.4.8(c) was filtered with a normalized 2D Gaussian kernel with *full width at half maximum* (FWHM) equals to $1/(2\sqrt{\log(4)})$ (corresponding to a Gaussian with standard deviation parameter σ equals to 1) and using circular boundary conditions. If the filtering is performed independently on each component of the Jacobian matrices the data elements are considered to belong to the vector space \mathbb{M}_2 . Figure 5.4.8(d) shows the result of the component-wise filtering. It can be seen that the filtering yielded two locations with negative values of the Jacobian determinant (shown as green pixels).

In order to preserve the group structure of the data, an intrinsic filtering was proposed in [Pennec et al. 2006]. Filtering of manifold-valued data can be interpreted as a *weighted intrinsic mean*. Given a data field $\bar{F} : \Omega \rightarrow \mathbb{E}$, where \mathbb{E} is a Euclidean space, and a normalized spatial kernel $K(r)$ the discrete convolution is $(K * \bar{F})(\mathbf{x}) = \sum_r K(r) \bar{F}(\mathbf{x} - r)$. The filtered field satisfies

$$\hat{F}(\mathbf{x}) = (K * \bar{F})(\mathbf{x}) = \operatorname{argmin}_{\bar{Y} \in \mathbb{E}} \sum_r K(r) \|\bar{F}(\mathbf{x} - r) - \bar{Y}\|^2.$$

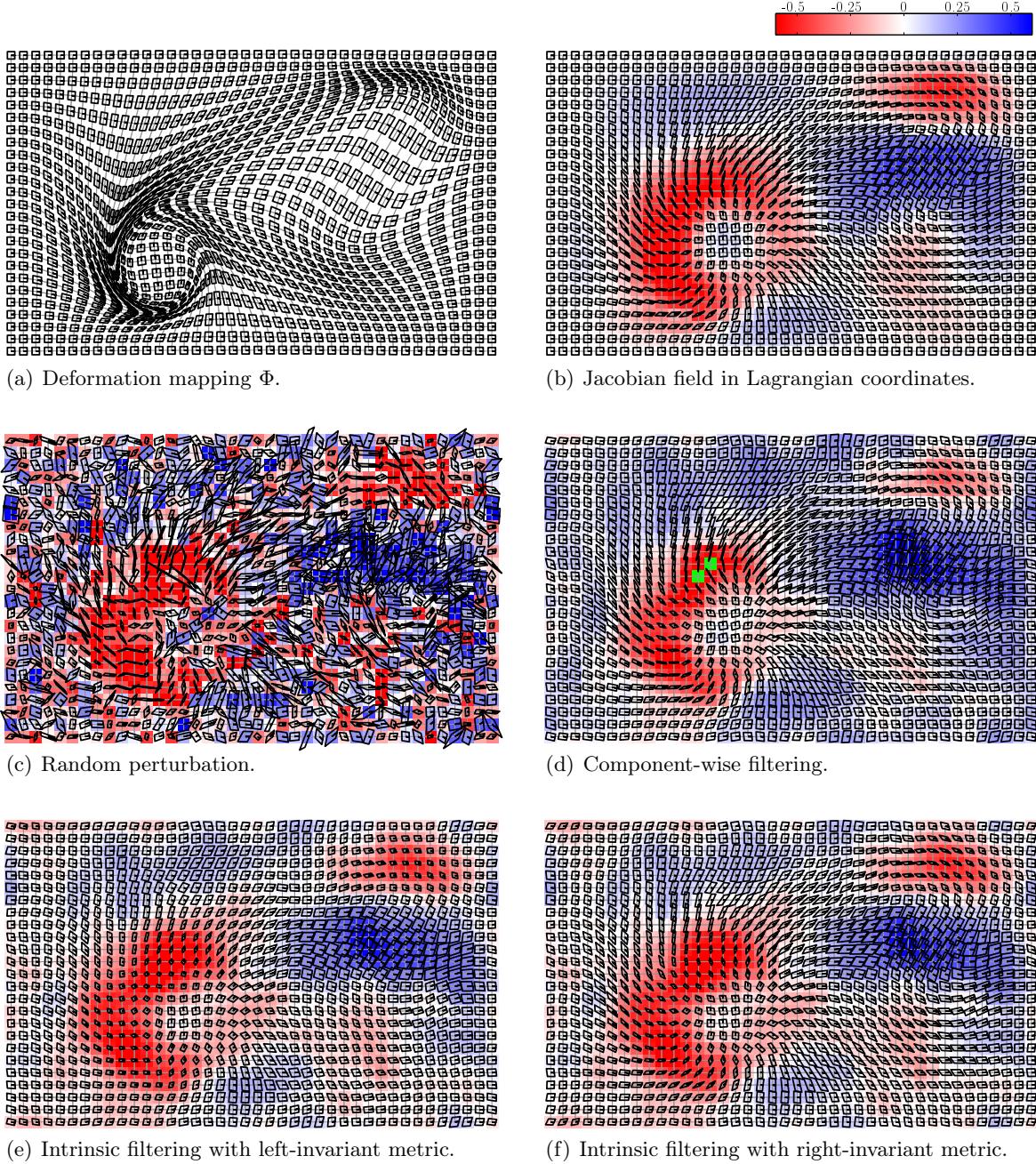


Figure 5.4.8. Intrinsic Riemannian filtering of a field of general linear transformations. In all images, glyphs represent the local deformation. In panels (b)-(f) glyphs are located in Lagrangian coordinates, while in panel (a) are located in deformed coordinates. Color represents local volume change in logarithm scale.

As it was done in Section 5.4.2, this interpretation of the convolution with a normalized kernel can be straightforwardly extended to Riemannian manifold-valued data by replacing the Euclidean distance by the Riemannian distance function. Note that two metrics are involved in this framework, the metric on the data, and the Euclidean metric from the spatial domain which defines the relative weights between neighboring elements via the kernel K .

For a manifold valued function $F: \Omega \rightarrow \mathcal{M}$, where \mathcal{M} is a Riemannian manifold, similarly to the computation of intrinsic mean in Section 5.4.2, a necessary condition for the intrinsic

filtering \hat{F} is

$$\sum_r K(r) \operatorname{Log}_{\hat{F}}(F(\mathbf{x} - r)) = 0.$$

The following gradient descent procedure on \mathcal{G} can be performed to compute the intrinsic filtering:

$$\hat{F}_{k+1} = \operatorname{Exp}_{\hat{F}_k} \left(\varepsilon \sum_r K(r) \operatorname{Log}_{\hat{F}_k}(F(\mathbf{x} - r)) \right),$$

where ε can be determined by backtracking such that

$$\sum_r K(r) \operatorname{distance}(F(\mathbf{x} - r), \hat{F}_{k+1}^\varepsilon)^2 < \sum_r K(r) \operatorname{distance}(F(\mathbf{x} - r), \hat{F}_k)^2.$$

While the convergence to a stationary point is guaranteed, necessary conditions ensuring uniqueness of minima are difficult to obtain, chiefly for compact groups (see, for example, [Buss and Fillmore 2001] [Hartley et al. 2013]). The proposed descent procedure converges to the intrinsic filtering for a sufficiently small dispersion of the instances.

Intrinsic filtering on $\mathbf{GL}^+(2)$ was applied to the Jacobian matrices of Figure 5.4.8(c) using invariant metrics defined at the identity by $\langle\!\langle U, U \rangle\!\rangle_I = \operatorname{trace}(U^T U)$. The results, are shown in Figure 5.4.8(e)-(f), for left- and right-invariant metrics respectively. Note that both results differ and the appropriate selection of the invariance depends on the application.

In this example, where the elements belong to $\mathbf{GL}^+(n) \equiv \mathbf{SL}(n) \times \mathbf{S}^+(1)$, it can be checked that the intrinsic filtering satisfies

$$\log(\det(K * F))|_{\mathbf{x}} = (K * \log(\det(F)))|_{\mathbf{x}}$$

for both left- and right-invariance cases. That is, the logarithms of the local volume change descriptor of the intrinsically filtered transformations coincide with the scalar filtering of the logarithms of the volume change descriptor from the original data.

5.4.4 Extension to any left-invariant metric of the closed-form expression for $\mathcal{ST}(1)$ group

This section makes use of the isometry between different metrics under the conjugacy action on geodesics explained in Section 4.7 to provide a closed-form expression for $\mathcal{ST}(1)$ group valid for any invariant metric.

A set of bases for the algebra of $\mathcal{ST}(1)$ is given by $\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\}$ (see Section 5.1.4). Let $\mathbf{m} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ be the representation of the inertia operator for which a closed-form solution for $\operatorname{Exp}(\cdot)$ was given in Eq. (5.1.1). Let Y be an element of $\mathcal{ST}(1)$ that is represented by the 2×2 matrix $Y = \begin{pmatrix} s & t \\ 0 & 1 \end{pmatrix}$ where $s \in \mathbb{R}^+$ is the scale factor and $t \in \mathbb{R}$ is the displacement parameter. Using Y in Eq. (4.7.1), the transformed and scaled inertia operator $\tilde{\mathbf{m}}$ takes the form

$$\tilde{\mathbf{m}} = \alpha \mathbf{B}^T (Y \otimes Y^{-T}) \mathbf{B}^{\dagger T} \mathbf{m} \mathbf{B}^{\dagger} (Y^T \otimes Y^{-1}) \mathbf{B} = \frac{\alpha}{s^2} \begin{pmatrix} s^2 + t^2 & t \\ t & 1 \end{pmatrix}. \quad (5.4.6)$$

It is easy to see that any possible inertia operator can be written as Eq. (5.4.6), where

$$\tilde{\mathbf{m}} = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \Rightarrow \begin{cases} s = \frac{\sqrt{\det(\tilde{\mathbf{m}})}}{B} \\ t = \frac{C}{B} \\ \alpha = \frac{\det(\tilde{\mathbf{m}})}{B} \end{cases} \quad (5.4.7)$$

Then, for any inertia operator represented by $\tilde{\mathbf{m}} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$, the element Y that satisfies Eq. (4.7.2), can be computed from Eq. (5.4.7) as

$$Y = \begin{pmatrix} \sqrt{\det(\tilde{\mathbf{m}})} / B & C / B \\ 0 & 1 \end{pmatrix}.$$

Finally, following Eq. (4.7.2), the general closed-form expression for the Riemannian exponential is

$$\text{Exp}_I^{\tilde{\mathbf{m}}}(U) = Y \text{Exp}_I(Y^{-1}UY)Y^{-1}$$

and $\text{Exp}_I(Y^{-1}UY)$ can be computed by using Eq. (5.1.1).

Figure 5.4.9(a) shows several geodesics, all starting from the identity with a given initial velocity $U = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$, but under different inertia operators $\tilde{\mathbf{m}}$. The corresponding inertia operator for each curve is given explicitly in the figure. White circles along each curve show equidistant points along the segment from $t = 0$ to 1 and the white square is the element at the mid point of the trajectory ($t = 0.5$). Geodesics under the inertia operator $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are known to be semicircles in the (t, s) plane with center on the $s = 0$ axis [Arsigny 2006] [Stahl 2007] (see red curves in Figure 5.4.9). Under a general $\tilde{\mathbf{m}}$ geodesics become elliptical trajectories (considering a straight line as a degenerated ellipse).

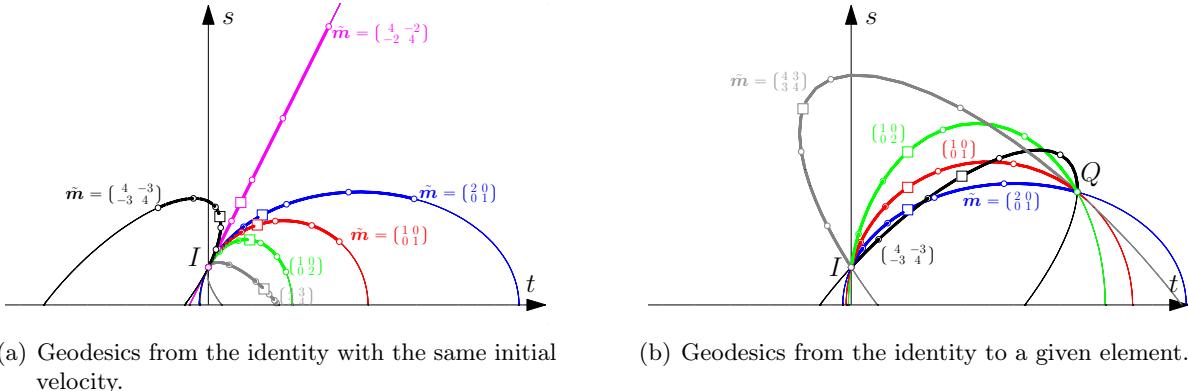


Figure 5.4.9. Geodesics for $\text{ST}(1)$ using different inertia operators $\tilde{\mathbf{m}}$.

Figure 5.4.9(b) shows several geodesics between the identity and the element $Q = \begin{pmatrix} 3 & 6 \\ 0 & 1 \end{pmatrix}$. The experiments performed on this simple group illustrate the strong dependence of the metric on the geodesic trajectories [Milnor 1976]. This behavior is expected to hold also on other spatial transformation groups. Accordingly, the choice of the metric is a crucial and application-dependent issue.

CHAPTER

6

Applications to brain morphometry

6.1 Statistical analysis of <i>pose</i> information from subcortical nuclei	146
6.1.1 Subject and group selection	147
6.1.2 Subcortical nuclei delineation	148
6.1.3 Spatial transformation model and mean shape for each subcortical object	148
6.1.4 Pose feature and intrinsic mean	150
6.1.5 Statistical analysis of pose features. A regression study	152
6.1.6 Hypothesis testing and statistical assessment	157
6.2 Tensor-Based Morphometry	159
6.2.1 Subject groups and template construction	160
6.2.2 Tensor-based morphometry (TBM) quantification	161
6.2.3 Statistical analysis	164
6.2.4 Assessment of statistical significance	166
6.3 Multivariate tensor-based morphometry	170
6.3.1 Invariance with respect to the template	171
6.3.2 Cramér test	177
6.3.3 Synthetic study	178
6.3.4 Results on ADNI dataset	181
6.4 Discussion	182

Previous chapters were focused on formulating the mathematical background and methodologies for image registration and analysis of spatial transformations. This chapter will provide applications of these methodologies to the neuroanatomy field. The main goal will be to identify and quantify anatomical group differences between Alzheimer's disease (AD) patients and control subjects. The statistical tools for the analysis will be defined specifically depending on the geometrical feature to be analyzed for each study: either pose information of subcortical nuclei characterized by spatial transformations, or deformation mappings characterized by volume changes scalar maps or Jacobian matrix fields.

Nowadays a wide variety of techniques are available for the analysis of brain anatomy. Many of these techniques have been applied in brain morphometry studies to characterize brain atrophy patterns present in AD. Morphometry studies can be classified into two main groups: studies at a population level; and single subject studies. In the former case, the focus of such studies has been either to identify the average anatomical differences among individuals from two or more subject groups in cross-sectional studies [Studholme et al. 2004, 2001] [Cardenas et al. 2007a,b] [Hua et al. 2008] [Bossa et al. 2010b] or to identify anatomical changes over time during a longitudinal study [Cardenas et al. 2007b] [Hua et al. 2009] [Bossa et al. 2010a] [Lorenzi et al. 2011] [Lorenzi and Pennec 2013a]. Also, to describe the relationship between anatomy and other relevant measurements (such as clinical variables, genetic information [Hibar et al. 2011], cognitive scores [Hua et al. 2008], biomarkers [Misra et al. 2009]).

6.1 Statistical analysis of *pose* information from subcortical nuclei

The oldest approach for quantitative analysis of brain anatomy is the volumetry technique, which measures the volume of specific brain structures. Although historically the anatomical changes were discovered in necropsy studies [Braak and Braak 1991], the macroscopical consequences of neurodegenerative changes at a cellular level can be visualized and measured with *in vivo* structural imaging techniques, such as magnetic resonance imaging (MRI). Currently, volumetry is a powerful, robust and intuitive technique that has yielded a wealth of findings. It relies on manual or automatic delineation of *regions of interest* (ROI). Volume measures of particular brain structures such as entorhinal cortex, hippocampus, parahippocampal gyrus, and amygdala [Laakso et al. 1995] [Krasuski et al. 1998] [Jack Jr et al. 1999] [Du et al. 2001] [Du et al. 2003] [Pennanen et al. 2004] have been long used as a neuroimaging markers of dementia in cross-sectional studies. In addition, longitudinal studies usually make use of the volume change rate.

More specific and subtle shape information of particular regions or structures, such as hippocampus, has also been analyzed by means of statistical shape analysis. In the shape analysis framework, shape is often defined as all the geometrical information of an object which is invariant to *pose*, usually defined as the information about location, attitude and, very often, the object size. Therefore, shape and pose are separate features that provide complementary information about the full description of the object of interest. Several shape features have been used so far, such as landmark coordinates [Csernansky et al. 2000, 2004], thickness or radial atrophy maps [Thompson et al. 2007] [Querbes et al. 2009], and medial representations [Styner et al. 2003]. In all these shape analysis studies of a single ROI, the pose information is rejected during an alignment stage because pose mainly depends on irrelevant external factors such as patient's location and its relative attitude within the scanner.

However, the information of relative pose among different structures belonging to a complex system may be useful for diagnosis, prognosis and monitoring disease status. Neurodegeneration

implies brain tissue loss which yield structural readjustments, including local volume changes of particular structures, but also positional shifts and deformations. A methodology to build statistical pose models was introduced in [Bossa and Olmos 2006]. Later, a joint statistical analysis of relative pose and shape was performed in [Bossa and Olmos 2007]. In [Rao et al. 2008] the correlation of the anatomical information from subcortical nuclei was analyzed by describing the geometrical objects as a set of landmarks after a global alignment, that can be considered as a joint shape and pose descriptor of the multi-structure system. The shape and the relative pose of the subcortical nuclei multi-structure was also analyzed in a longitudinal pediatric study on autism [Styner et al. 2006a] and a recent discrimination analysis can be found in [Gorcowski et al. 2010] [Bossa et al. 2011].

The aim of the following study is to assess the usefulness of pose information from subcortical nuclei for quantifying AD-induced anatomical changes. In particular, a regression study of pose features versus cognitive scores is performed in order to assess the relationship between cognitive decline and subcortical nuclei gross anatomy. It is also of interest to evaluate pose changes induced by healthy aging process and compare it with a the pathological induced change. Thus, a regression versus age is also presented.

6.1.1 Subject and group selection

A subset of 554 elderly subjects from the ADNI study with a three-year follow up (see Section 1.4.3) was selected. All subjects underwent clinical/cognitive assessment, as well as measurements of certain AD biomarkers at the time of scan acquisition.

In the ADNI study, as part of each subject's cognitive evaluation, the Mini-Mental State Examination (MMSE) was performed providing a global measure of mental status based on evaluation of several cognitive domains: orientation, attention, calculation, registration, language and recall [Cockrell and Folstein 2002]. The maximum score for MMSE is 30 corresponding to normal cognitive status, and scores lower or equal than 24 are usually consistent with a dementia status. The Clinical Dementia Rating–Sum of Boxes (CDRSB) was also assessed as a measure of dementia severity by evaluating six cognitive domains: memory, orientation, judgment and problem solving, home and hobbies, personal care and community affairs [Morris 1993]. CDRSB score has a dynamic range from 0 to 18 where higher scores correspond to more severe dementia. Also, subjects were evaluated using Alzheimer's Disease Assessment Scale Cognitive Subscale (ADAS-cog) [Mohs 1994] [Peña-Casanova 1997] which is a brief cognitive test battery that assesses learning and memory, language production, language comprehension, constructional praxis, ideational praxis, and orientation. ADAS-cog score ranges from 0 for healthy status up to 70 for very severe dementia. Diagnosis of the subjects were made by medical experts according to NINCDS-ADRDA criteria [McKhann et al. 1984] which assigns to each subject a diagnostic label of *control subject*, *mild-cognitive impairment patient* (MCI) or *Alzheimer's disease patient*.

Subjects were divided into four groups according to their diagnostic label resulting in the following distribution:

- a group called $\{\text{NOR}\}_{n=1}^N$ with $N = 207$ healthy control subjects (and remaining as control subjects along a 3-year follow up);
- a group called $\{\text{AD}\}_{a=1}^A$ with 176 AD patients;
- group $\{\text{sMCI}\}_{s=1}^S$ (or *stable* MCI) with 89 patients, formed by subjects whose diagnostic label was MCI all over the 3-year follow-up);
- and a group called $\{\text{cMCI}\}_{c=1}^C$ (or *converter* MCI) with 82 patients, including patients who started with an MCI label and converted to AD during the 3-year follow-up.

These groups will be used to define a disease stage. It should be noted that clinical evidence of dementia was only available for patients within $\{\text{AD}\}$ group, and for patients belonging to $\{\text{sMCI}\}$ after 3-year follow-up. It is unknown the percentage of patients from $\{\text{sMCI}\}$ or from $\{\text{NOR}\}$ that will convert to AD in longer follow-up intervals. In spite of this limitation, the order $\{\text{NOR}\} \rightarrow \{\text{sMCI}\} \rightarrow \{\text{cMCI}\} \rightarrow \{\text{AD}\}$ is used to characterize a sequence of “potential evolution of the disease”. Table 6.1 provides a summary of demographic data and cognitive baseline scores of each subject group.

Table 6.1. Demographic data and cognitive scores of subject groups (from ADNI database). Age, MMSE, CDRSB and ADAS format: *average \pm standard deviation [min, max]*.

Group	Gender (M/F)	Age	baseline MMSE	baseline CDRSB	baseline ADAS
$\{\text{NOR}\}$	101/106	76 ± 5 [62,90]	29 ± 1 [26,30]	0.0 ± 0.1 [0,0.5]	9 ± 4 [1,21]
$\{\text{sMCI}\}$	66/23	75 ± 7 [55,88]	28 ± 2 [24,30]	1.3 ± 0.6 [0.5,3]	17 ± 6 [3,40]
$\{\text{cMCI}\}$	55/27	75 ± 7 [55,90]	27 ± 2 [23,30]	1.9 ± 1.0 [0.5,5]	18 ± 6 [4,40]
$\{\text{AD}\}$	89/87	75 ± 8 [55,91]	23 ± 2 [20,27]	4.3 ± 1.6 [1,9]	28 ± 7 [13,50]

6.1.2 Subcortical nuclei delineation

Baseline preprocessed T1-weighted MR brain images (details at Section 1.4.3) were analyzed with FIRST tool, from FSL package [Smith et al. 2004] [Jenkinson et al. 2012], providing automatic delineation of the following subcortical structures: caudate nucleus (Caud.); nucleus accumbens (Accu.); putamen (Puta.); globus pallidus (Pall.); hippocampus (Hipp.); amygdala (Amyg.); and thalamus (Thal.). FIRST is a model-based registration/segmentation tool that makes use of statistical shape/appearance models to segment the subcortical structures. Delineations of subcortical structures are surfaces parameterized as triangle meshes whose vertices are assumed to be landmarks with correspondence across subjects.

In order to remove global effects of location and orientation of subjects within the scanner, and their head size, original MR images were aligned to a reference image by means of a registration of subject images using similarity transformation model $\mathcal{SITM}(3)$ (see Section 2.5.2). For each subject, the spatial transformation was also applied to the subcortical objects obtaining a set of landmarks coordinates in a normalized coordinate system.

For each subject, normalized landmarks coordinates from all subcortical structures are considered as a combined geometrical object. For example, the notation for the a -th instance from the group $\{\text{AD}\}$ is

$$\text{AD}_a \equiv (\text{LCaud}_a, \text{LAccu}_a, \text{LPuta}_a, \text{LPall}_a, \text{LHipp}_a, \text{LAmyg}_a, \text{LThal}_a, \\ \text{RCaud}_a, \text{RAccu}_a, \text{RPuta}_a, \text{RPall}_a, \text{RHipp}_a, \text{RAmyg}_a, \text{RThal}_a),$$

where prefixes L and R correspond to the left- and right-hemisphere, respectively. While subindex a runs over $\{\text{AD}\}$, subindex letters n, s and c run over $\{\text{NOR}\}$, $\{\text{sMCI}\}$ and $\{\text{cMCI}\}$, respectively. Subindex letter κ will be used for a generic instance from any group. Each subcortical structure is a geometrical object composed by landmarks in the normalized coordinates of the reference image.

6.1.3 Spatial transformation model and mean shape for each subcortical object

For a given structure and subject, for example the left caudate from the n -th instance within $\{\text{NOR}\}$ group, LCaud_n , its triangle mesh surface encloses an interior region. Let $\text{vol}(\text{LCaud}_n)$

be the volume of that interior region, and let $\text{center}(\text{LCaud}_n)$ be its barycenter. An efficient way to compute $\text{vol}(\cdot)$ is by means of the *divergence theorem* which involves the integration of a function with gradient equals to one across the surface triangles. The centroid $\text{center}(\cdot)$ can be computed by similar integrations [Millán et al. 2007] [Pozo et al. 2011].

Let us consider that pose is characterized by the group of centered transformations defined by $\mathcal{C} = \mathcal{S}^+(1) \times \mathcal{SO}(3) \times \mathcal{T}(3)$. Spatial transformations belonging to \mathcal{C} can be parameterized as $h(s, R_\theta, \bar{t})$, where $s \in \mathbb{R}^+$ is a scale parameter, R_θ is a rotation matrix, and $\bar{t} \in \mathbb{R}^3$ represents a translation. The action of $h(s, R_\theta, \bar{t})$ on a point y of the *object with center* (LCaud_n , $\text{center}(\text{LCaud}_n)$) is given by (see Section 2.5.2)

$$h(s, R_\theta, \bar{t}) \star y = sR_\theta(y - \text{center}(\text{LCaud}_n)) + \text{center}(\text{LCaud}_n) + \bar{t}$$

and simultaneously acting on $\text{center}(\text{LCaud}_n)$ by

$$h(s, R_\theta, \bar{t}) \star \text{center}(\text{LCaud}_n) = \text{center}(\text{LCaud}_n) + \bar{t}.$$

Then, given for example, the set of all left caudate nuclei $\{\text{LCaud}_\kappa\}$ (by considering together the left caudate objects from all groups), the *mean shape under \mathcal{C} group* is computed as the geometrical object $\langle \text{LCaud} \rangle$ whose distance to all instances is minimized¹. The mean shape $\langle \text{LCaud} \rangle$ is defined as the result of the optimization

$$\begin{aligned} \langle \text{LCaud} \rangle &= \underset{\mathbf{M}, \text{vol}(\mathbf{M})=1}{\operatorname{argmin}} \left(\sum_{\kappa=1} \min_{h_\kappa \in \mathcal{C}} \text{dis}(\mathbf{M}, h_\kappa \star \text{LCaud}_\kappa) \right) \\ &= \begin{cases} \underset{\mathbf{M}}{\operatorname{argmin}} & \sum_{\kappa=1} \text{dis}(\mathbf{M}, \mathcal{R}_c(\text{LCaud}_\kappa, \mathbf{M}) \star \text{LCaud}_\kappa) \\ \text{subject to} & \text{vol}(\mathbf{M}) = 1 \end{cases}, \end{aligned} \quad (6.1.1)$$

where the involved dissimilarity function $\text{dis}(\cdot, \cdot)$ is the sum of squared distances between corresponding landmarks, Eq. (3.2.1) in Section 3.2.1, and the registration operation $\mathcal{R}(\cdot, \cdot)$ was defined in Section 3.1. The solution of the registration process can be performed algebraically (see Section 3.4). The constraint $\text{vol}(\mathbf{M}) = 1$ is required to avoid the singular infimum at 0 resulting by “almost” collapsing all instances to a point by negligible scales (when the group under consideration does not produce scaled versions of the objects, this singularity is not an issue)². Algorithms to compute landmark coordinates of the mean shape can be found in [Goodall 1991] [Cootes et al. 1995] [Le 1998] [Dryden and Mardia 1998].

Figure 6.1.1 shows several left caudate instances LCaud_κ and the resulting mean shape $\langle \text{LCaud} \rangle$. Figure 6.1.2 shows the obtained mean shape of other subcortical structures: $\langle \text{LAccu} \rangle$; $\langle \text{LPuta} \rangle$; $\langle \text{LPall} \rangle$; $\langle \text{LHipp} \rangle$; $\langle \text{LAmyg} \rangle$; and $\langle \text{LTthal} \rangle$.



Figure 6.1.1. Left caudate instances LCaud_κ and their mean shape.

¹ By abusing of notation, the symbol $\langle \cdot \rangle$ is also used for geometrical objects, but they definition differs from the given in Section 5.4.2.

² The mean shape problem (6.1.1) has multiple solution, all being translated and rotated versions of the same object. Any of these solutions can be equally considered for the subsequent analysis.

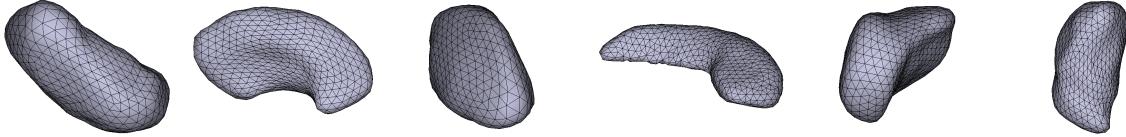


Figure 6.1.2. Mean shape of left subcortical structures. From left to right: $\langle \text{LAccu} \rangle$, $\langle \text{LPuta} \rangle$, $\langle \text{LPall} \rangle$, $\langle \text{LHipp} \rangle$, $\langle \text{LAmyg} \rangle$, and $\langle \text{LThal} \rangle$.

6.1.4 Pose feature and intrinsic mean

The volume magnitude of a region can be understood as the feature indicating how many times of a prototypical volume fits in the region. According to the *Erlangen program* modern description of the geometry [Klein 1893] [Yaglom et al. 1988], the volume is a magnitude which is preserved under the *diffeomorphism group generated by divergence free flows*. In a similar fashion, let us define the *pose feature* of a subcortical structure as the feature which is preserved under transformations on \mathcal{C} group. Then, *pose feature* is computed as the transformation from \mathcal{C} group that must be applied on the mean shape (considered as prototypical shape) in order to best fit the subcortical structure. Pose feature gives the relative location, size and attitude between the subcortical structure and the mean shape. For example, the pose feature of the κ -th left-caudate is

$$h_{\kappa}^{LCaud} = \mathcal{R}_{\mathcal{C}} (\langle \text{LCaud} \rangle, \text{LCaud}_{\kappa}) . \quad (6.1.2)$$

The action of the pose feature on the mean shape, $h_{\kappa}^{LCaud} \star \langle \text{LCaud} \rangle$, places the mean shape as close as possible to the structure LCaud_{κ} from the κ -th subject.

While all instances LCaud_{κ} are described in the coordinate system of the reference image, the mean shape $\langle \text{LCaud} \rangle$ is in an arbitrary coordinate system (it has **vol** equal to one and arbitrary location and attitude (see footnote 2 on page 149)). To place the mean shape $\langle \text{LCaud} \rangle$ in the coordinate system of the each image, the pose feature h_{κ}^{LCaud} is computed. Given the pose features of several structures $\{h_{\kappa}^{LCaud}\}$, the intrinsic mean pose is computed as in Section 5.4.2 by

$$\langle h^{LCaud} \rangle = \operatorname{argmin}_{q \in \mathcal{C}} \sum_{\kappa} \mathbf{distance}(q, h_{\kappa}^{LCaud})^2 , \quad (6.1.3)$$

where **distance**(\cdot, \cdot) is a distance on the group of spatial transformations \mathcal{C} . A left-invariant Riemannian metric will be considered following previous Chapters 4 and 5.

The group \mathcal{C} is the direct product of three well-known spatial transformation groups. Considering a block-structured inertia operator, where rotations, scales and translations do not interact, geodesics on \mathcal{C} are the direct product of geodesics in each subgroup (see Section 5.1.10). Therefore, the shortest geodesic connecting spatial transformations $q(r, R_{\theta}, \bar{d})$ and $h(s, R_{\theta}, \bar{t}) \in \mathcal{C}$ is obtained from the shortest geodesic between the isotropic scales r and s in $\mathcal{S}^+(1)$ (Section 5.1.2), the geodesic between rotations R_{θ} and R_{θ} in $\mathcal{SO}(3)$ (Section 5.1.5), and the geodesic between translations given by \bar{d} and \bar{t} in $\mathcal{T}(3)$ (Section 5.1.1).

Let us consider a metric in the rotation space given by the inner product in the algebra $\mathfrak{so}(3)$ of the form $\langle\langle U_1, U_2 \rangle\rangle = \operatorname{trace}(U_1^T U_2)$. Finally, the distance function results

$$\mathbf{distance}(q(r, R_{\theta}, \bar{d}), h(s, R_{\theta}, \bar{t})) = \sqrt{\rho_S 3 |\log(s/r)|^2 + \rho_R \|\log(R_{\theta}^{-1} R_{\theta})\|_F^2 + \rho_T \|\bar{d} - \bar{t}\|_F^2} , \quad (6.1.4)$$

where $\rho_S, \rho_R, \rho_T > 0$ are weights corresponding to rotation, scaling and translation components, respectively. For the current experiment these weights were set to 1. Note that the previous distance is not only left-invariant but also right-invariant. Therefore, the Riemannian exponential function coincides with the group exponential and can be easily computed by using the matrix exponential.

The algebra of the group \mathcal{C} is the direct sum of the algebras of its components, $\mathbf{c} = \mathfrak{s}(1) \oplus \mathfrak{so}(3) \oplus \mathfrak{t}(3)$. Therefore an element of this algebra can be described by the triplet $(\sigma, \Sigma(\bar{\theta}), \bar{t}) \in \mathbf{c}$, where

$$\Sigma(\bar{\theta}) = \begin{pmatrix} 0 & -\theta^z & \theta^y \\ \theta^z & 0 & -\theta^x \\ -\theta^y & \theta^x & 0 \end{pmatrix}.$$

The group exponential of the previous triplet results

$$\exp((\sigma, \Sigma(\bar{\theta}), \bar{t})) = h(e^\sigma, \exp(\Sigma(\bar{\theta})), \bar{t}) \in \mathcal{C}. \quad (6.1.5)$$

The Riemannian logarithm function results in

$$\text{Log}_{q(r, R_\theta, \bar{d})}(h(s, R_\theta, \bar{t})) = (r \log(s/r), R_\theta \log(R_\theta^{-1} R_\theta), \bar{t} - \bar{d}) \in T_q \mathcal{C}. \quad (6.1.6)$$

Instead of the triplets, a matrix representation in \mathbb{M}_7 can be used for the elements of \mathcal{C} as was done in Section 5.1.10. Then, representations of $h(s, R_\theta, \bar{t}) \in \mathcal{C}$ and $(\sigma, \Sigma(\bar{\theta}), \bar{t}) \in \mathbf{c}$ are given by

$$H(s, R_\theta, \bar{t}) = \left(\begin{array}{c|c|c} sR_\theta & | & 0_{3 \times 3} & | & 0_3 \\ \hline 0_{3 \times 3} & | & I_3 & | & \bar{t} \\ \hline 0_3^T & | & 0_3^T & | & 1 \end{array} \right) \quad \text{and} \quad (\sigma, \Sigma(\bar{\theta}), \bar{t}) \equiv \left(\begin{array}{c|c|c} \sigma I_3 + \Sigma(\bar{\theta}) & | & 0_{3 \times 3} & | & 0_3 \\ \hline 0_{3 \times 3} & | & 0_{3 \times 3} & | & \bar{t} \\ \hline 0_3^T & | & 0_3^T & | & 0 \end{array} \right).$$

With these representations, the distance on \mathcal{C} given in Eq. 6.1.4, can be computed by using the matrix logarithm function

$$\text{distance}(Q(r, R_\theta, \bar{d}), H(s, R_\theta, \bar{t})) = \|\log(Q^{-1}H)\|_F$$

Using Eqs. (6.1.5), (6.1.6) and the descent procedure explained in Section 5.4.2, the intrinsic mean given in Eq. (6.1.3) were computed for the 14 subcortical structures. Each intrinsic mean spatial transformation was applied to the corresponding mean shape allowing to place the mean shape in the reference image coordinate system.

In the definition of the pose feature for each subcortical structure given in Eq. (6.1.2), there are two hidden choices: the election of a reference image I to where all subjects are aligned to get normalized coordinates of the subcortical nuclei descriptors; and the election of a mean shape $\langle \text{LCaud} \rangle$ among all multiple solutions of the problem (6.1.1). If a different subject image I' had been chosen as reference, then all left caudate objects would have been of the form $\text{LCaud}'_\kappa = w \star \text{LCaud}_\kappa$, where $w \in \mathcal{C}$ is the spatial transformation relating both reference system defined by I and I' . Similarly, if a different mean shape $\langle \text{LCaud} \rangle'$ (also being a solution of the problem (6.1.1)) had been selected, it would have been of the form $\langle \text{LCaud} \rangle' = r \star \langle \text{LCaud} \rangle$, for some $r \in \mathcal{C}$. Using these new, also arbitrary, choices pose features would result

$$\begin{aligned} \tilde{h}_\kappa^{LCaud} &= \mathcal{R}_c(\langle \text{LCaud} \rangle', \text{LCaud}'_\kappa) \\ &= \mathcal{R}_c(r \star \langle \text{LCaud} \rangle, w \star \text{LCaud}_\kappa) \\ &= w \bullet \mathcal{R}_c(\langle \text{LCaud} \rangle, \text{LCaud}_\kappa) \bullet r^{-1} \\ &= w \bullet h_\kappa^{LCaud} \bullet r^{-1}. \end{aligned}$$

But, as the distance function on \mathcal{C} given in Eq. (6.1.4) is bi-invariant, the mean pose remains invariant to the arbitrary choice of w and r resulting in a mean pose of the form $\langle \tilde{h}^{LCaud} \rangle = w \bullet \langle h^{LCaud} \rangle \bullet r^{-1}$.

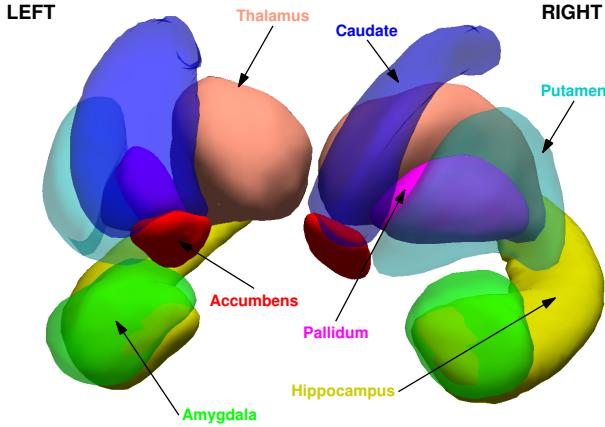


Figure 6.1.3. Mean shape of each subcortical nuclei at its mean pose.

The mean shape of the 14 subcortical structures, placed in the selected reference image system, is shown in Figure 6.1.3.

The intrinsic mean pose can also be computed for each subject group {NOR}, {sMCI}, {cMCI}, and {AD}, providing a summarized description of subcortical nuclei pose information at different stages of AD. Figure 6.1.4 shows the mean shapes at the mean poses for each subject group. It can be seen that both hippocampi shrink along disease stages, which is largely reported in the literature. Interesting, it is also seen that other subcortical nuclei suffer positional displacements. For example, caudate nuclei present positional shifts in the medial-lateral direction without relevant volume changes. These shifts are in agreement with ventricular growth, also reported in the literature. It should be kept in mind that this study is cross-sectional and the description of pose changes along disease is estimated from different subjects.

6.1.5 Statistical analysis of pose features. A regression study

Although visual and qualitative differences between subject groups can be seen in Figure 6.1.4, we are interested in quantifying the relationship between pose features and relevant regression variables. For each subcortical structure, we propose to perform a regression analysis to estimate the relationship between the pose feature and the cognitive score ADAS-cog.

As we are dealing with features belonging to a set without a vector space structure, appropriate statistical tools must be used. We begin by briefly reviewing a regression analysis on a vector space structure. Suppose we are interested in the relationship between non-random independent variables $\{t_\kappa\}$ and random dependent variables $\{x_\kappa\}$ under a generative model of the form

$$x_\kappa = \alpha + \beta t_\kappa + \varepsilon_\kappa, \quad (6.1.7)$$

where $\{\varepsilon_\kappa\}$ are unobservable random variables representing unknown hidden perturbations or measurement errors. In the linear regression language, the parameter α is called *intercept* and the parameter β is called *slope*. Assuming that ε is an *independent and identically distributed* Gaussian random variable, the maximum likelihood estimator of the parameters α and β is given by the well-known *least-squares* problem

$$(\alpha^*, \beta^*) = \underset{\alpha, \beta}{\operatorname{argmin}} \sum_{\kappa} |x_\kappa - \alpha - \beta t_\kappa|^2.$$

For the purposes of generalizing to a manifold case, following [Fletcher 2013], it is useful to think that in the idealized model (without the noise term) the variable $\hat{x} = \alpha + \beta t$ follows a geodesic path in \mathbb{R} , starting at α with velocity β and arc-length parameterized by the variable t .

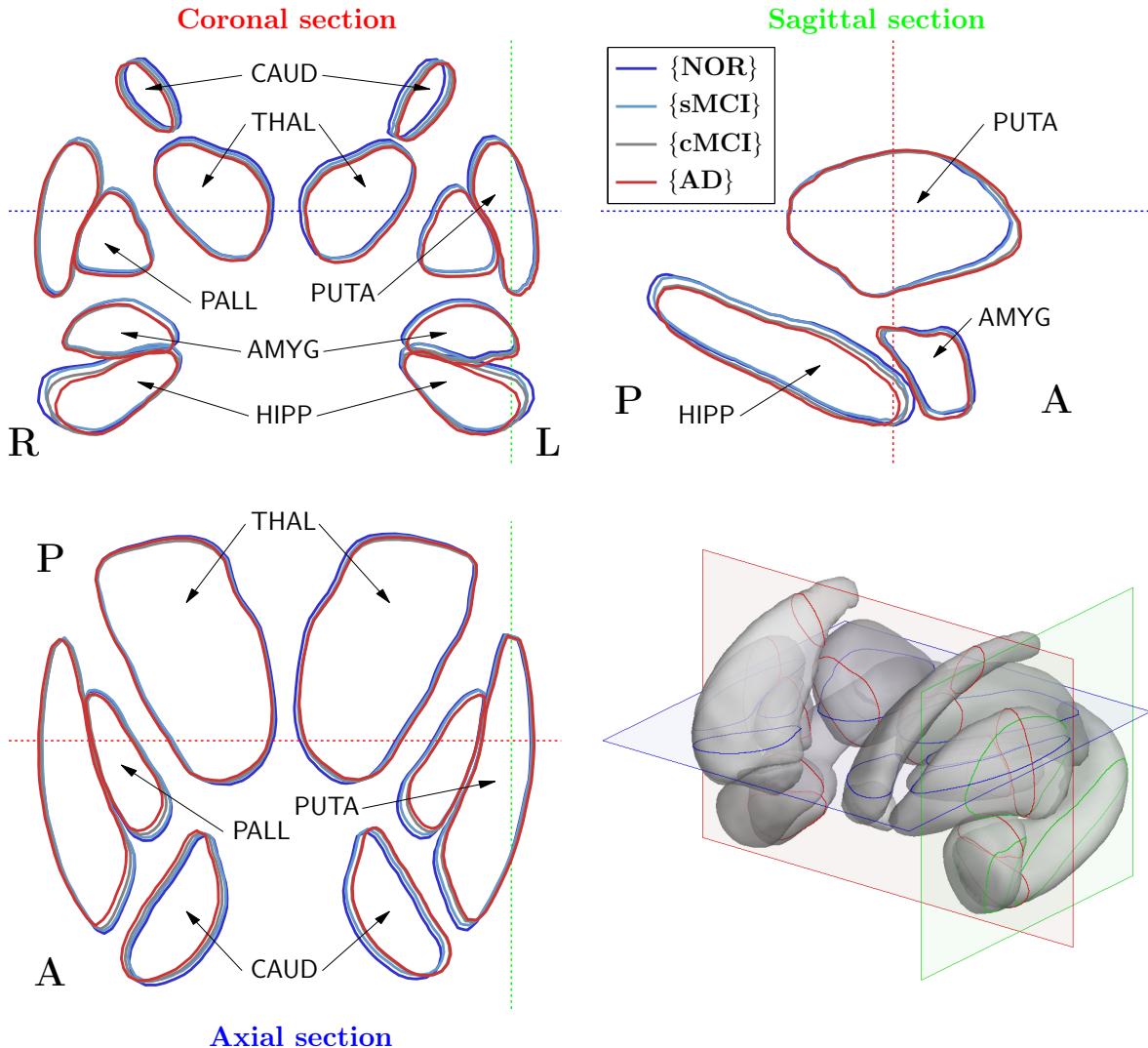


Figure 6.1.4. Illustration of the mean pose of subcortical nuclei for each subject group. The panels show slices of the structures in coronal, sagittal and axial views. R, L, P and A refer to right, left, posterior and anterior anatomical terms, respectively.

Now, let $\{\chi_\kappa\}$ be a set of instances from a Riemannian manifold \mathcal{M} , each of them with an associated scalar value $\{t_\kappa \in \mathbb{R}\}$. The goal is to find a geodesic $\gamma_{(t)} = \text{Exp}_a(t \mathcal{V})$ passing at $t = 0$ through the “intercept” $a \in \mathcal{M}$ with “slope” (velocity) $\mathcal{V} \in T_a \mathcal{M}$, such that the modelled variables $\hat{\chi}_\kappa = \text{Exp}_a(t_\kappa \mathcal{V})$ lie “as close as possible” to the instances χ_κ .

Generalizing the linear model (6.1.7), the instances $\{\chi_\kappa\}$ are modelled as a geodesic model³

$$\chi_\kappa = \text{Exp}_{\text{Exp}_a(t_\kappa \mathcal{V})}(\epsilon_\kappa),$$

where ϵ_κ is an unobservable living at the tangent space $T_{\hat{\chi}_\kappa} \mathcal{M}$. This geodesic regression model on \mathcal{M} is illustrated in Figure 6.1.5. Assuming that each ϵ_κ is distributed as an uniform Gaussian

³ In plain words the idea is to replace the addition operation in the linear model with translations along geodesics in \mathcal{M} . It is interesting to remark that, while the addition is a commutative operation, successive geodesic translations do not. An *a priori* election in the order of the operands must be taken into account in the definition of the model. Here we are assuming that the dependent variable, first evolve following the geodesic model, and it is afterwards perturbed by the noise. The non-commutativity complicates the generalization of a linear model in which confounding effects want to be included.

variable in the corresponding tangent space, the maximum likelihood estimation of the parameters a and \mathcal{V} is given by

$$(a^*, \mathcal{V}^*) = \underset{a \in \mathcal{M}, \mathcal{V} \in T_a \mathcal{M}}{\operatorname{argmin}} \sum_{\kappa} \operatorname{distance}(\operatorname{Exp}_a(t_{\kappa} \mathcal{V}), x_{\kappa})^2. \quad (6.1.8)$$

To formalize the previous interpretation, it is necessary to define a suitable procedure to translate Gaussian probability density functions (or any *pdf*) on a connected manifold between nearby tangent spaces. For the purposes and scope of this thesis, it is sufficient to define isotropic Gaussian distributions in terms of the metric of each tangent space. Therefore, the translation of the metric provides a generalization of *identically distributed* noise random variables along the tangent bundle of the manifold. For more elaborated procedures, see for example [Pennec 2006a] [Grenander 2008] [Freifeld 2013].

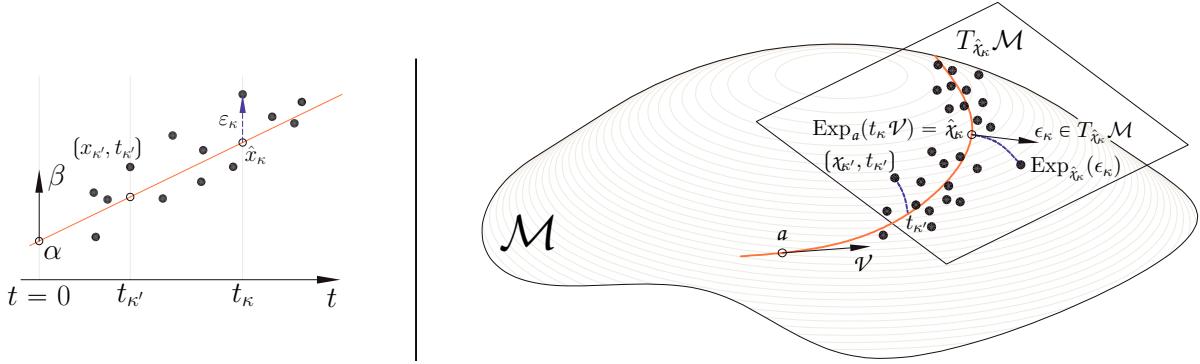


Figure 6.1.5. Left panel: illustration of a linear regression on a vector space. Right panel: Schematic illustration of a geodesic regression model on a manifold. The red curve is a geodesic model for the observed data (black dots). Each point x_{κ} has an associated value t_{κ} and the geodesic is parameterized in terms of this variable t . The geodesic passes through a (which is called *intercept*) with velocity \mathcal{V} (called *slope*) at $t = 0$. Realizations of the geodesic model \hat{x}_{κ} at $t = t_{\kappa}$ are compared with data by means of the *length of the residuals* $\|\epsilon_{\kappa}\|_{\hat{x}_{\kappa}}$.

With the previous formulation, the regression problem is defined intrinsically in the manifold. Let us particularize the regression analysis to a matrix group structure $\mathcal{M} \equiv \mathbf{G}$ with a left-invariant metric and consider that the velocity $V \in T_A \mathbf{G}$ is expressed in terms of the left-trivialized velocity $U \in \mathfrak{g}$ as $V = AU$. Therefore,

$$\begin{aligned} \operatorname{distance}(\operatorname{Exp}_A(t_{\kappa} V), X_{\kappa}) &= \operatorname{distance}(A \operatorname{Exp}_I(t_{\kappa} U), X_{\kappa}) \\ &= \operatorname{distance}(X_{\kappa}^{-1} A \operatorname{Exp}_I(t_{\kappa} U), I) \\ &= \|\operatorname{Log}_I(X_{\kappa}^{-1} A \operatorname{Exp}_I(t_{\kappa} U))\|_{\mathfrak{g}}. \end{aligned} \quad (6.1.9)$$

Replacing Eq. (6.1.9) in Eq. (6.1.8), the estimation of the optimal parameters A and V (or equivalently U) is given by

$$(A^*, U^*) = \underset{A \in \mathbf{G}, U \in \mathfrak{g}}{\operatorname{argmin}} \sum_{\kappa} \|\operatorname{Log}_I(X_{\kappa}^{-1} A \operatorname{Exp}_I(t_{\kappa} U))\|_{\mathfrak{g}}^2.$$

However, instead of performing the optimization for the variable A over the group space \mathbf{G} which is a non-linearly constrained space, it is convenient to express $A = \operatorname{Exp}_I(\sum_i B_i \alpha^i)$ and $U = \sum_i B_i \beta^i$ for some $\bar{\alpha}, \bar{\beta} \in \mathbb{R}^k$, coordinates of the algebra \mathfrak{g} with bases $\{B_i\}_{i=1}^k$.

In the case of pose features of subcortical nuclei, where the underlying considered group of spatial transformations is \mathcal{C} , the Riemannian exponential and logarithm functions $\operatorname{Exp}(\cdot)$ and

$\text{Log}(\cdot)$, result in the matrix exponential and logarithm functions $\exp(\cdot)$ and $\log(\cdot)$, respectively as was explained in Section 6.1.4. Therefore, the maximum likelihood estimation of the parameters is⁴

$$(\bar{\alpha}^*, \bar{\beta}^*) = \underset{\bar{\alpha}, \bar{\beta} \in \mathbb{R}^7}{\operatorname{argmin}} \sum_{\kappa} \left\| \log \left(X_{\kappa}^{-1} \exp(B_i \alpha^i) \exp(t_{\kappa} B_i \beta^i) \right) \right\|_F^2, \quad (6.1.10)$$

where the use of the Frobenius norm $\|\cdot\|_F$ follows from the distance defined in Eq. (6.1.4). Regarding existence and uniqueness issues of the geodesic problem, please refer to [Fletcher 2013].

In order to solve the problem (6.1.10) a descent procedure can be used. Derivatives of the each summand $\|\log(S_{\kappa})\|_F^2$, with $S_{\kappa} = X_{\kappa}^{-1} \exp(B_i \alpha^i) \exp(t_{\kappa} B_i \beta^i)$, can be computed using the following expressions⁵:

$$\begin{aligned} \mathcal{D}_{\bar{\alpha}} \|\log(S_{\kappa})\|_F^2 &= 2 \overline{\log(S_{\kappa})}^T \mathcal{D} \log(S_{\kappa}) \mathcal{D}_{\bar{\alpha}} S_{\kappa} \\ \mathcal{D}_{\bar{\alpha}} S_{\kappa} &= (\exp(t_{\kappa} B_i \beta^i)^T \otimes X_{\kappa}^{-1}) \mathcal{D} \exp(B_i \alpha^i) \mathbf{B} \\ \mathcal{D}_{\bar{\beta}} \|\log(S_{\kappa})\|_F^2 &= 2 \overline{\log(S_{\kappa})}^T \mathcal{D} \log(S_{\kappa}) \mathcal{D}_{\bar{\beta}} S_{\kappa} \\ \mathcal{D}_{\bar{\beta}} S_{\kappa} &= t_{\kappa} (I_7 \otimes X_{\kappa}^{-1} \exp(B_i \alpha^i)) \mathcal{D} \exp(t_{\kappa} B_i \beta^i) \mathbf{B}, \end{aligned}$$

where Fréchet derivatives $\mathcal{D}\exp(\cdot)$ and $\mathcal{D}\log(\cdot)$ were previously defined and used in Section 4.3.2.

Wise initial estimates for $\bar{\alpha}$ and $\bar{\beta}$ to start the optimization process can be computed. To do so, without loss of generality, the independent variable t (which also serves as parametrization of the geodesic) can be translated to $\tilde{t}_{\kappa} = t_{\kappa} - \langle t_{\kappa} \rangle$ such that $\langle \tilde{t}_{\kappa} \rangle = 0$. In this situation, a good estimation of the parameter A_0 is given by the intrinsic mean $\langle X_{\kappa} \rangle$ and the initial value $\bar{\alpha}_0$ is given by $\mathbf{B}^{\dagger} \log(A_0)$. An initial estimate for β can be computed by the linear regression of the residuals $A_0^{-1} X_{\kappa}$ at \tilde{t}_{κ} projected by the logarithm function on the algebra \mathbf{g} . That is, by solving

$$\bar{\beta}_0^* = \underset{\bar{\beta} \in \mathbb{R}^7}{\operatorname{argmin}} \sum_{\kappa} \left\| \bar{\beta} \tilde{t}_{\kappa} - \mathbf{B}^{\dagger} \overline{\log(A_0^{-1} X_{\kappa})} \right\|^2.$$

In the present study we propose to investigate the relationship between pose features $\{h_{\kappa}^{LCaud}\}$ and a cognitive status. Given the instances of a subcortical structure, geodesic regressions with respect to ADAS-cog score were independently computed for the 14 subcortical structures. Also, in order to compare with the effects if a normal aging process, geodesic regressions of pose were performed with respect to age of healthy subjects from {NOR} group.

The pose information is a 7 degrees of freedom feature (1 parameter for scale, 3 for relative attitude between the instance and the mean shape and 3 parameters defining the translation of the mean shape). Although the dimensionality of the feature is not so large, it is hard to visualize and give an intuitive interpretation to the intercept and velocity of the obtained geodesic. To aid its interpretation, it is helpful to illustrate the action of pose on the mean shape for each subcortical nuclei. Figure 6.1.6 shows subcortical nuclei slices obtained with the geodesic regression modeling of the pose with respect to ADAS-cog. This regression model allows to visualize the average subcortical nuclei pose, which is a compact and very coarse descriptor of the geometry of these structures, at different cognitive status.

An important finding is that most subcortical structures shift downwards (superior-inferior direction) and in the medial-lateral direction. These positional shifts are agreement with

⁴ In the following $B_i \alpha^i$ is used as a shorthand for $\sum_i B_i \alpha^i$.

⁵ Let A , B and C be $n \times n$ matrices, then $\overline{ABC} = (C^T B^T \otimes I_n) \overline{A}$ and therefore $\mathcal{D}_A(ABC) = C^T B^T \otimes I_n$, which is an $n^2 \times n^2$ matrix. Similarly, $\overline{ABC} = (C^T \otimes A) \overline{B}$ and therefore $\mathcal{D}_B(ABC) = C^T \otimes A$. Finally, $\overline{ABC} = (I_n \otimes AB) \overline{C}$ and then $\mathcal{D}_C(ABC) = I_n \otimes AB$.

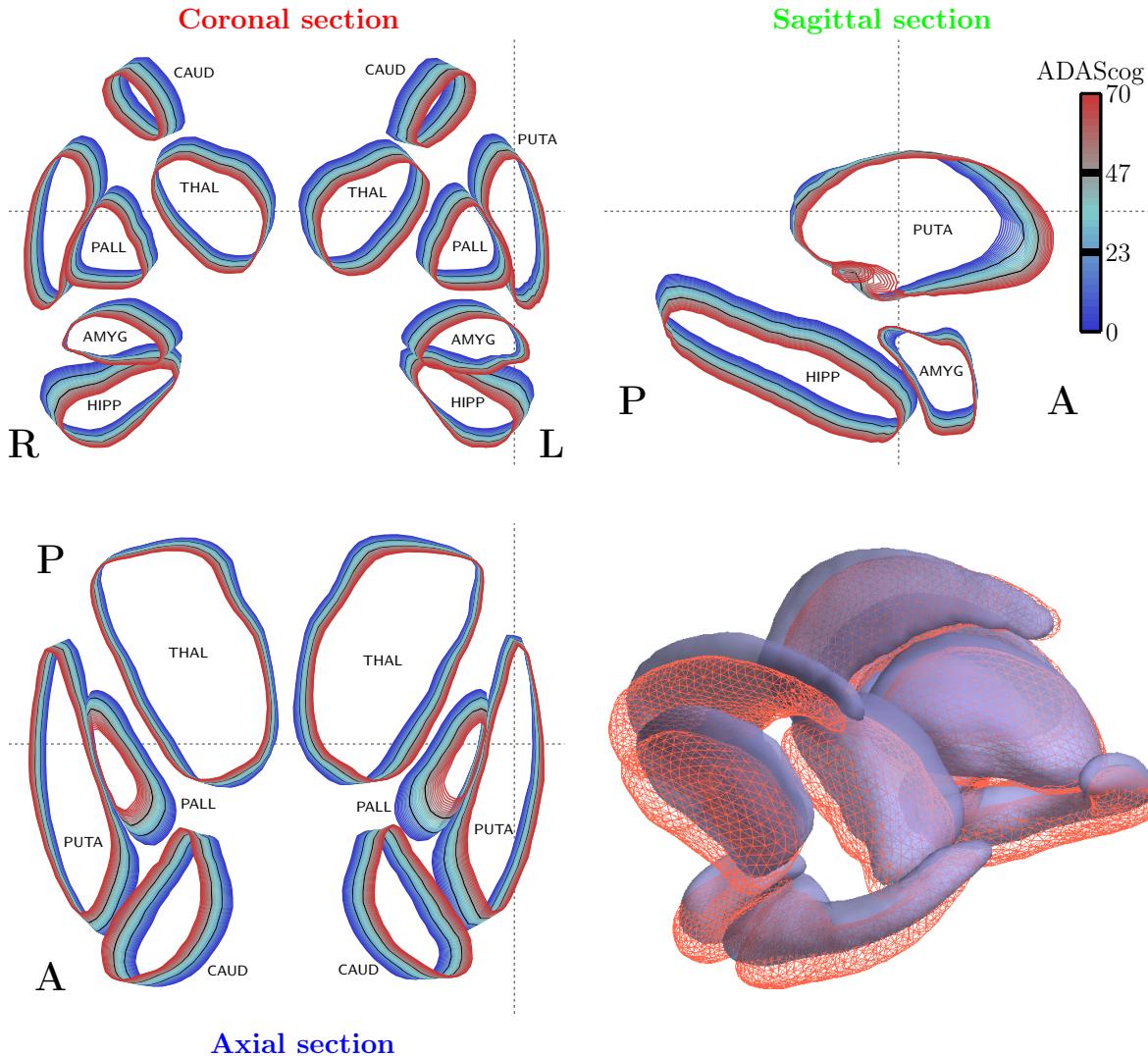


Figure 6.1.6. Illustration of the effects of the geodesic regression of pose features from all subjects with respect to ADAS-cog score. Panels show slices of modelled pose by the geodesic regression at different values of the ADAS-cog score. Blue colors represent low values of ADAS-cog and therefore a healthy status, while red colors are used to extreme values of ADAS-cog corresponding to severe dementia status. Although observed data range up to 55 points corresponding to medium AD status, regression model allows to extrapolate up to an ADAS-cog value of 70. Bottom-right panel shows 3D surface representations of subcortical nuclei at extreme values of the scale to facilitate the interpretation of the modelled anatomical and geometrical changes.

medial temporal atrophies and also with ventricular growth, typically reported in the literature. Both caudate nuclei give an example of subcortical structure with large shifts but without important volume changes. Others structures such as putamen, pallidum and hippocampus present positional shifts as well as volume reductions when evolving from low to high ADAS-cog scores.

It is well-known that age is a very important factor for AD. An interesting question is if the profile of anatomical changes of subcortical nuclei in a cognitive decline differ from the changes occurring in the process of normal aging. In order to get some relevant insight to answer this question, a regression model with respect to age of the pose feature of subcortical nuclei from

{NOR} group was performed. Results are illustrated in Figure 6.1.7. A visual comparison between Figures 6.1.6 and 6.1.7 reveals that while with respect to the more ADAS-cog scores, the more subcortical nuclei shift in the superior-inferior and medial-lateral direction, but with respect to age only medial-lateral shifts look relevant. Compare, for example, the shifts of caudate nuclei. They follow a downwards-lateral direction with respect to cognitive decline, but an upwards-lateral direction for normal aging study. Also, hippocampi suffer attitude changes, mainly pronounced at the anterior part (induced by the atrophy of the medial-temporal lobe) in the ADAS-cog regression, while attitude change are more pronounced in the posterior region in the normal aging study.

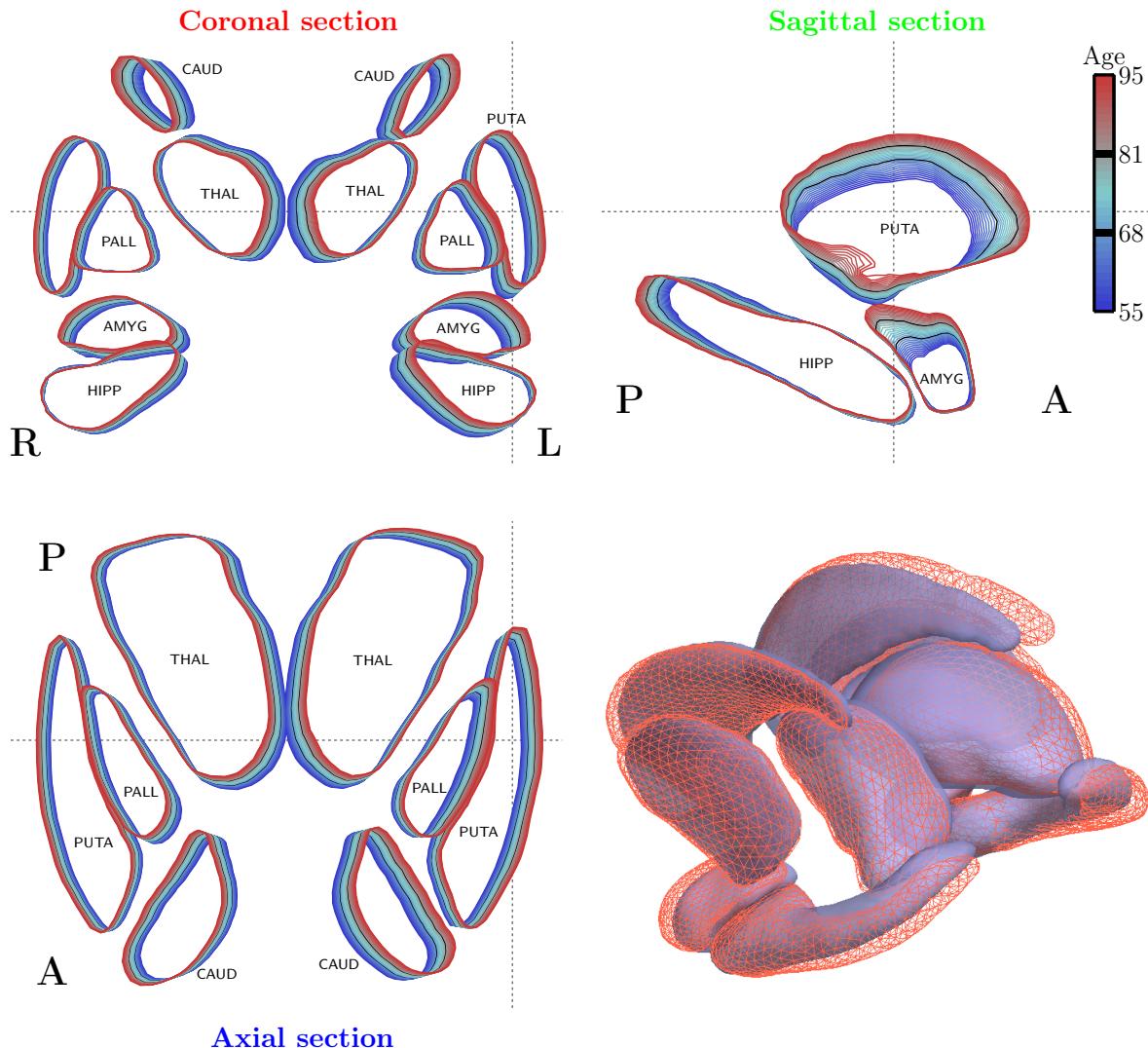


Figure 6.1.7. Illustration of results of geodesic regression model between subcortical nuclei pose and age for subjects from {NOR} group. See also caption of Figure 6.1.6 for more details.

6.1.6 Hypothesis testing and statistical assessment

A common practice to evaluate how well data fit a regression models is to measure the amount of variance explained by the model. While, in a linear model the variance is measured as the mean of the squared residuals $\|x_\kappa - \hat{x}_\kappa\|^2$, in a geodesic model it can be generalized as the mean value

of $\|\text{Log}_{\hat{x}_k}(\chi_k)\|_{\hat{x}_k}^2$, *i.e.* the distance between measurement and its model realization. The amount of explained variance by the regression model is quantified by the *coefficient of determination* R^2 which indicates how well the observations fit the model. The statistic R^2 is computed by comparing the variance of observations with and without considering the model

$$R^2 = 1 - \frac{\sum_k \|x_k - \hat{x}_k\|^2}{\sum_k \|x_k - \langle x \rangle\|^2}$$

which can be generalized to the geodesic regression case as

$$R^2 = 1 - \frac{\sum_k \|\text{Log}_{\hat{x}_k}(\chi_k)\|_{\hat{x}_k}^2}{\sum_k \|\text{Log}_{\langle x \rangle}(\chi_k)\|_{\langle x \rangle}^2},$$

where $\hat{x}_k = \text{Exp}_a(t_k \mathcal{V})$ is the model realization at $t = t_k$ and $\langle x \rangle$ is the intrinsic mean of the instances.

The coefficient of determination is always a value between 0 and 1. While a value of 1 indicates that observations perfectly fit the model, a value close to 0 indicates that the model is irrelevant for the measured data.

In order to assess if the model provides a significant description of the data, a *hypothesis test* can be conducted. Consider a null hypothesis \mathcal{H}_0 which express that the model is irrelevant for the data, in contrast to an alternative hypothesis assuming that the data actually present the behaviour proposed by the model. The *p*-value for a statistic is the probability of observing the current, or a more extreme, value of the statistic under \mathcal{H}_0 , *i.e.* the chance to have occurred if data were assumed to be irrespective from the model. The null hypothesis can be rejected in favor of the alternative hypothesis when the *p*-value turns out to be less than a certain significance level (typically selected as a value of 0.05).

Under \mathcal{H}_0 , there is no relationship between the instances χ_k and their associated values of the explanatory variable t_k . Therefore, any instance should be equally probable at any value of t and the instances χ_k are exchangeable among them [Aldous 1985]. A permutation test is a procedure to infer the distribution of a given statistic under the null hypothesis without making strong assumptions on the underlying distribution of the data [Good 2000] [Nichols and Holmes 2002] [Commenges 2003] [Manly 2006] [Moore et al. 2012]. It consists on simulating the null hypothesis by performing a large number of random permutations of the instances χ_k and computing for each permutation the statistic to be tested. Under some mild assumptions, such as exchangeability and homoscedasticity, and for a sufficient large number of random permutations, the outcome sequence of computed statistics empirically approximates the distribution of the statistic under \mathcal{H}_0 . Finally, in a regression study, a reliable estimation of the *p*-value of the fitting is given by the fraction of permutations that, by chance, have obtained a larger or equal value of R^2 than the “unpermuted” case.

On the regression problem of subcortical nuclei pose features with respect to ADAS-cog scores, the *p*-value of R^2 statistic were computed empirically by means of 20 000 random permutations. To compare the proposed methodology with a classical approach, a standard volumetric analysis was also performed by computing linear regressions on volumes of the different subcortical structures. Results are given in Table 6.2. Nucleus caudate and globus pallidum results significative for geodesics and not for linear regression, and with a larger coefficient of determination R^2 for the geodesic regression. It can be noted that as the *p*-values were computed by means of 20 000 random permutations, the resulting *p*-values are discretized at the resolution 1/20000 (5×10^{-5}) being also the smallest possible value.

Table 6.2. Statistical analysis of geodesic regressions between pose features and ADAS-cog scores. Also, results from linear regression on the volumes for each structure.

	L/R	Geodesic Regress. on pose features		Linear Regress. on volume	
		$\sqrt{R^2}$	p-value	$\sqrt{R^2}$	p-value
Accu.	L	0.105	1.00E-4	0.161	5.00E-5
	R	0.097	3.50E-4	0.186	5.00E-5
Amyg.	L	0.110	1.00E-4	0.124	8.50E-4
	R	0.095	7.50E-4	0.194	5.00E-5
Caud.	L	0.140	5.00E-5	0.040	2.73E-1
	R	0.135	5.00E-5	0.029	4.27E-1
Hipp.	L	0.109	1.00E-4	0.346	5.00E-5
	R	0.094	7.50E-4	0.340	5.00E-5
Pall.	L	0.131	5.00E-5	0.069	5.67E-2
	R	0.112	1.00E-4	0.091	1.26E-2
Puta.	L	0.120	5.00E-5	0.175	5.00E-5
	R	0.114	5.00E-5	0.093	1.06E-2
Thal.	L	0.119	5.00E-5	0.109	2.55E-3
	R	0.101	4.00E-4	0.128	4.00E-4

6.2 Tensor-Based Morphometry

The analysis presented in the previous section is a ROI-based (*region of interest*) analysis and the brain structures where to model and test the hypotheses were established *a priori*. In the case of neuroimaging studies, this class of analyses is typically limited to no more than a few tens of ROIs usually with previous wisdom and knowledge about the anatomical regions implicated in the disease. If the objective of the study is to identify and to localize those regions whose anatomies are affected by or related to the disease, different paradigms must be used. These paradigms have emerged to make inferences about regional responses in the brain without knowing where those responses were going to be expressed.

Tensor-based morphometry (TBM), *deformation-based morphometry* (DBM), and *voxel-based morphometry* (VBM) are techniques developed to quantify anatomical differences throughout the whole brain, offering a detailed perspective of characteristic geometrical patterns between groups of subjects, namely a control group and a group with a special characteristic of interest. DBM identifies regions with consistent structural positional shifts between groups of images, VBM was developed to identify consistent group differences in the grey-/white-matter density, and TBM identifies brain regions with a different atrophy/growth pattern between groups. The TBM technique is useful to detect structural differences among groups or over time, and to find relationships between anatomical patterns and different stages of a pathology in cross-sectional and longitudinal studies. It can also be used at individual level to assist in the diagnose task by finding and comparing anatomical patterns.

TBM relies on performing an image registration step between an anatomical *template* and the images under study. The resulting spatial transformations encode the geometrical differences between the anatomies, and at the same time, establish a dense correspondence between the template and the images. Using this correspondence, a local description of the deformations can

be represented in template coordinates, allowing the comparison of the anatomies from different subjects in a single standardized or stereotaxic space of reference. The most commonly used description of the local deformation is by the determinant of the gradient of the deformation mapping, which locally quantifies relative volume changes between the deformed and undeformed template. For each subject, an individual compression/expansion map is derived from the deformation mapping obtained by the registration. Afterwards, statistical analysis is performed at each location of the template in order to highlight those regions where atrophy/growth of tissues significantly differ between subject groups and then identifies regionally specific effects on, for example, diseases-related changes.

Although, it is theoretically possible to perform a statistical test at every point of the ambient space where the template is defined, it is common practice to carry out the quantification at the voxel-level resolution. Let T be an *anatomical template* (also called an *atlas*) which contains the anatomical features shared from all subject geometrical objects I_κ . When the template is registered towards I_κ , a spatial transformation $h_\kappa = \mathcal{R}(T, I_\kappa) : \mathbb{A} \rightarrow \mathbb{A}$ is obtained such that the deformed template $h_\kappa * T$ looks spatially similar to I_κ . Moreover, using a flexible enough spatial transformation model for h_κ , a dense correspondence between points in the template and the subject images is assumed. This correspondence is not restricted to only visually salient and significant points and features, but will be assumed that correspondence also identifies every point of the template $y \in \mathbb{A}$ with the point $h_\kappa * y = y_\kappa$ of the image I_κ . From each spatial transformation h_κ it is possible to compute a local description of the compressions/expansions by analyzing the deformation mapping. As all subject images were registered to the same template, the obtained compression/expansion maps share the anatomical coordinates defined by the template. At each voxel of the template, a group comparison test is irrespectively performed on the compression/expansion values and each result is laid out on a statistical map represented in the template reference system. The resulting *statistical map image* depicts, the spatial distribution of the statistic indicating a level of evidence of anatomical differences for each location.

From now on, all geometrical objects instances I_κ will be T1-weighted MR brain images from ADNI study (see Section 1.4.3). Also, image registration is performed with the *stationary velocity field* (SVF) registration technique presented in Section 3.6. Therefore, every spatial transformation h_κ will be parameterized by its stationary velocity field $\mathbf{v}_\kappa \mapsto h_\kappa = h^{\mathbf{v}_\kappa}$. Previous to the registration step, each image was preprocessed for intensity normalization with a criterion of intensity histogram matching with the template.

6.2.1 Subject groups and template construction

Two groups of preprocessed T1-weighted images from ADNI database were considered for this TBM study: a group $\{\mathbf{NOR}\}_{n=1}^N$ formed by $N = 231$ healthy and elderly subjects, without previous cognitive or neurological disorders; and a group $\{\mathbf{AD}\}_{a=1}^A$ formed by 200 sporadic AD patients.

A template from anatomical images has a wide impact in applications such as template-based segmentation [Lorenzo-Valdés et al. 2002] [Heckemann et al. 2006] [Xue et al. 2006] and construction of sparse shape models [Frangi et al. 2002] [Rueckert et al. 2003]. The estimation of the template is a key component in a TBM study. The template must represent common intensities patterns, geometrical content and anatomical features shared by all subjects. In some works such as [Thompson et al. 2000b] a single image instance was used as template. In order to avoid possible missregistrations due to large regularization forces, it is preferable that the template anatomy is not too far from the instances in terms of deformation. Additionally, the template should be unbiased to any instance or to any group. That is, the template could be

defined as the geometrical object that minimizes the amount of deformation required to match the template with all sample anatomies simultaneously. A minimal deformation template is an unbiased average template image created to represent common anatomical features from a set of subjects, typically with a mathematically-defined mean geometry for the population [Good et al. 2001] [Kochunov et al. 2002] [Joshi et al. 2004] [Studholme et al. 2004] [Kovačević et al. 2005] [Christensen et al. 2006] [Lorenzen et al. 2006] [Leporé et al. 2007].

For this TBM study the unbiased template was estimated using 40 randomly chosen subject images from $\{\text{NOR}\}$ and $\{\text{AD}\}$ groups. All 40 images, denoted by I_s with $s = 1, \dots, 40$, were affine-registered towards the International Consortium for Brain Mapping atlas (ICBM-53) [Mazziotta et al. 2001], by computing for each image I_s the affine spatial transformation $r_s = \mathcal{R}_{GA}(I_s, \text{ICMB-53})$. Then, an initial template T^0 was estimated by means of averaging the intensities of the registered images I_s , *i.e.* for all points y of the ambient space, $T_{(y)}^0 = \langle (r_s \star I_s)_{(y)} \rangle$. Next, an iterative process was performed to estimate the template, including four steps for each iteration:

1. perform a mixed affine/SVF registration of the current estimated template T^k and each subject image obtaining the deformation mapping $h_s^k = h^{v_s^k} = \mathcal{R}_{SVF}(T^k, r_s^k \star I_s)$, where $r_s^k = \mathcal{R}_{GA}(I_s, T^k)$;
2. compute the “barycentric” spatial transformation from the deformation mappings as $h^{\langle v_s^k \rangle}$ which under some conditions can be considered as an approximation of the intrinsic mean $\langle h_s^k \rangle$ [Arsigny 2006] [Pennec 2013];
3. for every point y , update the intensity of the template T^{k+1} by averaging intensities such that $T_{(y)}^{k+1} = \left\langle \left(h^{\langle v_s^k \rangle} \bullet \left(h^{v_s^k} \right)^{-1} \bullet r_s^k \star I_s \right)_{(y)} \right\rangle$.
4. and finally, in order to preserve as much as possible the stereotactic reference system defined by ICBM-53 atlas, align under affine transformation T^{k+1} to (ICBM-53), *i.e.* perform $T^{k+1} = \mathcal{R}_{GA}(T^{k+1}, \text{ICMB-53}) \star T^{k+1}$.

Figure 6.2.1 shows an illustration of the template construction process. At early stages of the process anatomical structures of the template are poorly defined due to the averaging of the intensity values from missaligned anatomical structures. As the process progresses, anatomical details and image characteristics such as the edges of the structures become sharper and better defined leading to a high detailed template.

The resulting template, after $k = 6$ iterations, is detailed in Figure 6.2.2. It is a $220 \times 220 \times 220$ volume with voxel size of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ where intracranial volume occupies approximately 1 800 000 voxels.

6.2.2 Tensor-based morphometry (TBM) quantification

Let $h = \mathcal{R}(T, I) : \mathbb{A} \rightarrow \mathbb{A}$ be the spatial transformation registering the template T towards the subject image⁶ I obtained by the SVF registration algorithm explained in Section 3.6. The geometrical and anatomical differences between both objects are encoded in the deformation mapping given by the spatial transformation h . The deformation mapping is from the SVF spatial transformation model and therefore it is by construction an invertible and differentiable

⁶ Actually, to avoid to quantify global differences in positioning, orientation and size between the images, a previous affine registration step is performed between the image and the template. Therefore, the deformation mapping results in $h = \mathcal{R}_{SVF}(T, \mathcal{R}_{GA}(I, T) \star I)$.

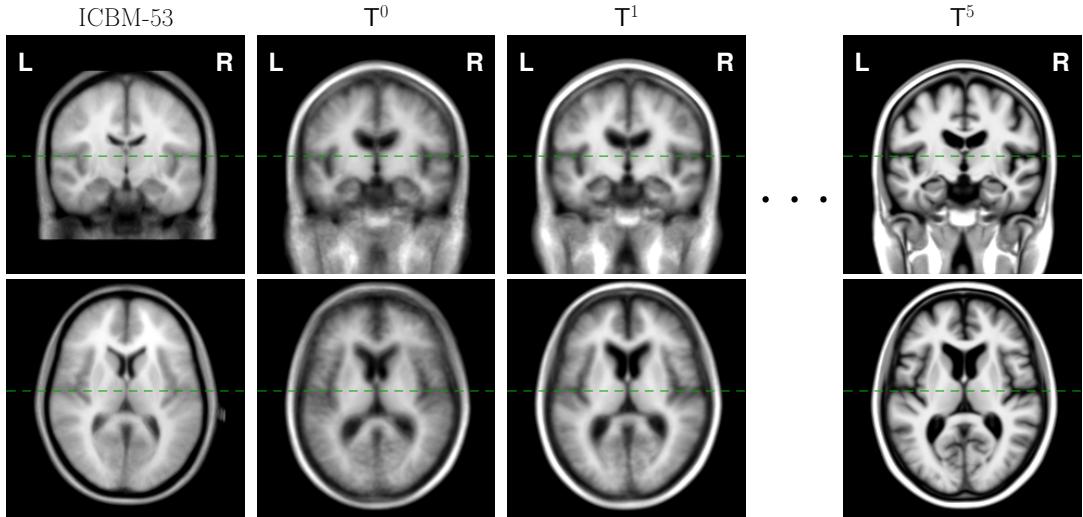
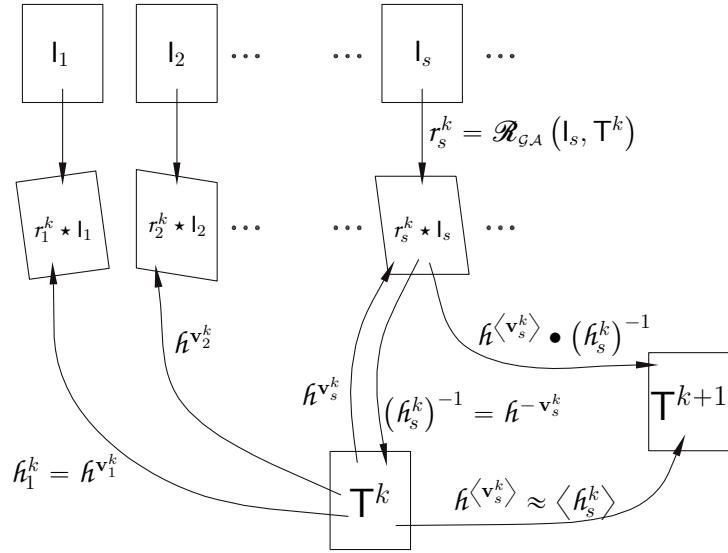


Figure 6.2.1. Construction of the unbiased template. Top panel: illustration of the procedure. Bottom panel: ICBM-53 atlas defining the stereotactic space, initial atlas T^0 , and several steps of the iterative procedure. Dashed lines show the location of the axial and coronal slices.

mapping from the ambient space to itself (see Section 2.5.4). Up to first order, this mapping acts on a differential neighborhood of a point $y \in \mathbb{A}$ as $h \star (y + dy) = h \star y + J(y)dy + O(dy^2)$, where J is a linear transformations field, such that $J(y): T_y \mathbb{A} \rightarrow T_{h \star y} \mathbb{A}$ for every $y \in \mathbb{A}$. In this thesis it is considered that the ambient space \mathbb{A} is a Euclidean space, such as \mathbb{R}^2 or \mathbb{R}^3 , and accordingly it can be considered that every tangent space $T_z \mathbb{A}$ are equivalent⁷ to \mathbb{R}^2 or \mathbb{R}^3 . Due to the invertibility and orientation preserving properties of the deformation mapping h , each linear transformation $J(y)$ belongs to the group $\mathcal{GL}^+(d)$.

Given a coordinate system for the ambient space $\mathbb{A} \equiv \mathbb{R}^d$ such that $y = (y^1, y^2, \dots, y^d)^T$ and $h \star y = (h_{(y)}^1, h_{(y)}^2, \dots, h_{(y)}^d)^T$, where h^c are the components of deformation mapping, each

⁷ In some works such as [Boucher et al. 2011] a strength measure of J is used to quantify and to regularize the deformations between embedded surfaces, but in those cases different tangent spaces are being considered and appropriate translations between them must be defined.

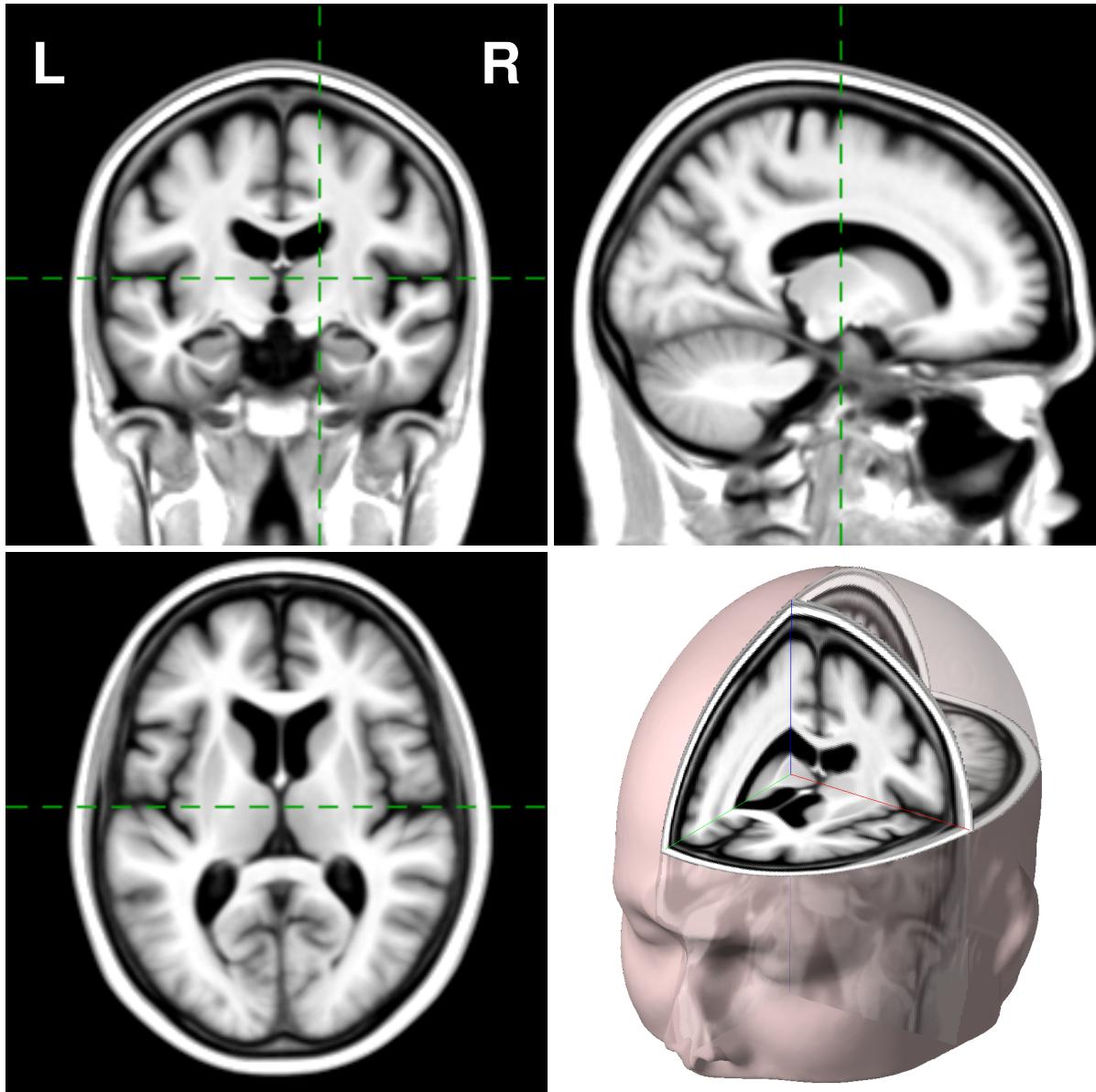


Figure 6.2.2. Obtained template, after $k = 6$ iterations.

linear transformation $\mathcal{J}(y)$ of $\mathcal{GL}^+(d)$ can be represented by a $d \times d$ matrix $J(y)$ belonging to the matrix group $\mathbf{GL}^+(d)$. The $d \times d$ matrix $J(y)$ is built by the spatial derivatives of the mapping h

$$J(y) = \mathcal{D}_z (h \star z)|_{z=y} = \begin{pmatrix} \partial_1 h^1|_y & \partial_2 h^1|_y & \cdots & \partial_d h^1|_y \\ \partial_1 h^2|_y & \partial_2 h^2|_y & \cdots & \partial_d h^2|_y \\ \vdots & & & \vdots \\ \partial_1 h^d|_y & \partial_2 h^d|_y & \cdots & \partial_d h^d|_y \end{pmatrix}$$

where $\mathcal{D}_z \cdot$ is the Jacobian operator with respect to spatial variations and $\partial_i h^j|_y$ is the derivative of the j -th component of $(h \star y)$ with respect to perturbations of y along the i -th spatial coordinate. When the spatial transformation is parameterized by a velocity field \mathbf{v} , such that the coordinates

of $h^{\mathbf{v}} \star \mathbf{y}$ is the solution at $t = 1$ of the IVP

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{v}(\mathbf{x}(t)) \\ \text{with initial condition } \mathbf{x}(0) &= \mathbf{y},\end{aligned}$$

then, the Jacobian matrix $J(\mathbf{y})$ can be obtained as the solution, at $t = 1$ of

$$\begin{aligned}\dot{J}(\mathbf{y}, t) &= (\mathcal{D}\mathbf{v})_{(\mathbf{x}(t))} J(\mathbf{y}, t) \\ \text{with initial condition } J(\mathbf{y}, 0) &= I_d.\end{aligned}$$

Statistical analysis is typically performed independently at each location \mathbf{y} on features extracted from the corresponding *Jacobian matrix*. The simplest, and still most widely used, feature in TBM studies is the determinant of the Jacobian matrix⁸ which describes the amount of local volume change between the template and each subject image indicating the local excess or deficit of the underlying tissues with respect to the template anatomy. Determinant values greater than one correspond to local expansions while values smaller than one correspond to local compressions. This feature has an intuitive interpretation providing a measure of the loss or excess of the underlying tissues with respect to the template anatomy [Davatzikos et al. 2001] [Studholme et al. 2006] [Hua et al. 2008] [Bossa et al. 2010b].

Figure 6.2.3 shows results of the registration of the template towards an instance from **{NOR}** group and an instance from **{AD}** group. Additionally, determinant of the Jacobian matrices (in logarithmic scale) are depicted in template coordinates. Zoomed panels show the obtained deformation mapping and also the compression/expansion map in subject coordinates. As the registration process obtains spatial correspondences between points of the template and each image, both compression/expansion maps can be compared in the template coordinates. It can be shown different atrophy/growth patterns of the tissues between both subjects. The ventricular system of the template must be compressed to match the control subject, but has to be expanded to match the AD patient. On the other hand, hippocampal regions must be expanded towards the control subject, but have to be compressed towards the AD patient.

6.2.3 Statistical analysis

A crucial part of experimental research is the statistical analysis of the results. *Two-sample Student's t-test* [Snedecor and Cochran 1991] [Woolson and Clarke 2011] is one of the simplest and most commonly used statistical methods which evaluates the difference between the means of two groups. Given two groups of scalar measurements or values $\{X\}$ and $\{Y\}$, the test relies on the t -statistic which is computed as⁹

$$t(\{X\}, \{Y\}) = \frac{\langle X \rangle - \langle Y \rangle}{\sqrt{\frac{\text{var}(X)}{|X|} + \frac{\text{var}(Y)}{|Y|}}}$$

where $\langle X \rangle$, $\text{var}(X)$ and $|X|$ are the classical mean, the unbiased variance and the cardinality of the sample set $\{X\}$, respectively.

Given the two groups of images $\{\mathbf{I}_n\}_{n=1}^{N=231}$ and $\{\mathbf{I}_a\}_{a=1}^{A=200}$ and the unbiased anatomical template \mathbf{T} , the deformation mappings $h_k = \mathcal{R}(\mathbf{T}, \mathbf{I}_k)$ and their

⁸ Some authors use the term Jacobian to refer to the whole matrix, while others use this terms to refer to the determinant of that matrix.

⁹ The expression corresponds to the statistic of the t -test for unequal sample sizes and unequal variances. This test is also known as the *Welch's t-test*.

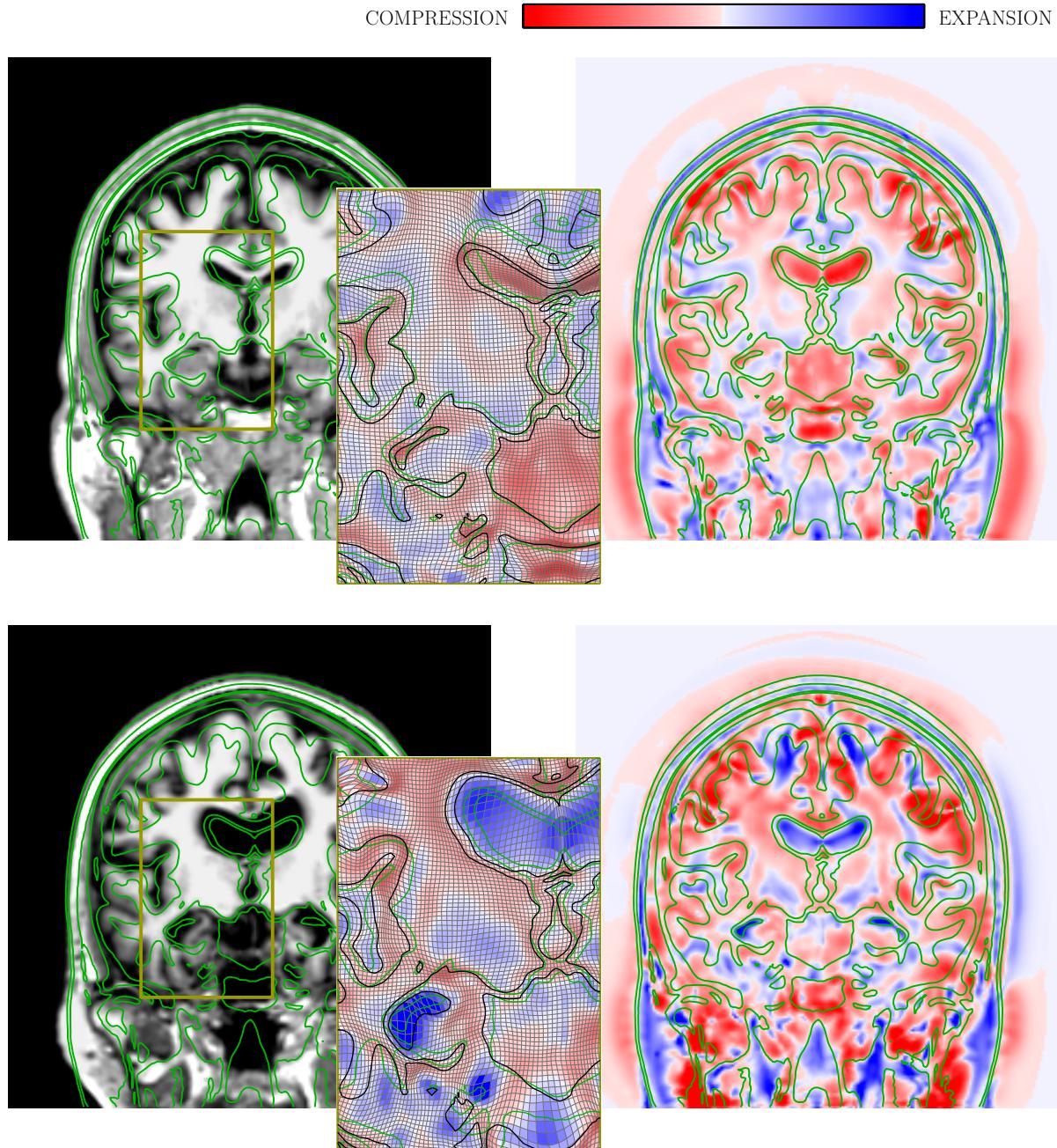


Figure 6.2.3. Top row: example of deformation mapping and corresponding compression/expansion map for a subject from **{NOR}** group. At left it is shown the subject image after affine registration to the template. Template contours are superimposed in green to help to visualize the alignment. At right it is shown the compression/expansion map (in logarithmic scale) in template coordinates. Compression/expansion map were computed from the determinant of Jacobian matrix of the deformation mapping. Zoomed panel shows the deformation mapping and compression/expansion map in subject coordinates. Bottom row: example of compression/expansion map for a subject from **{AD}** group. Both maps can be compared at template coordinates and it can be shown different atrophic patterns.

compression/expansion map $\det(J_\kappa)$ are computed for all images from both groups. As the deformation mappings κ are everywhere invertible, determinants of Jacobian matrices $J_\kappa(y)$ are inherently positive numbers. In order to adequate statistical analyses involving those quantities, it is a common practice to consider the logarithm of the determinants. After the logarithmic transform, determinant values 0.5 and 2 are both equally far from 1, which means that a local decrease in the volume of tissue to the half contributes the same than as a local expansion which double the volume.

Voxel-wise t -test statistical analysis is performed by comparing the sets $\{\log \det(J_n(y))\}_n^N$ from $\{\text{NOR}\}$ and $\{\log \det(J_a(y))\}_a^A$ from $\{\text{AD}\}$ group. Figure 6.2.5 shows the t -statistic map obtained from the $\{\text{NOR}\}$ - $\{\text{AD}\}$ group comparison. The t -value value computed for each voxel is laid out on the template. Blue to green regions correspond to regions where the tissue from instances of $\{\text{NOR}\}$ occupies, in mean, more volume than in the corresponding locations of the instances from $\{\text{AD}\}$. By contrast, red to yellow areas are associated with consistent tissue atrophies of the subjects from $\{\text{AD}\}$.

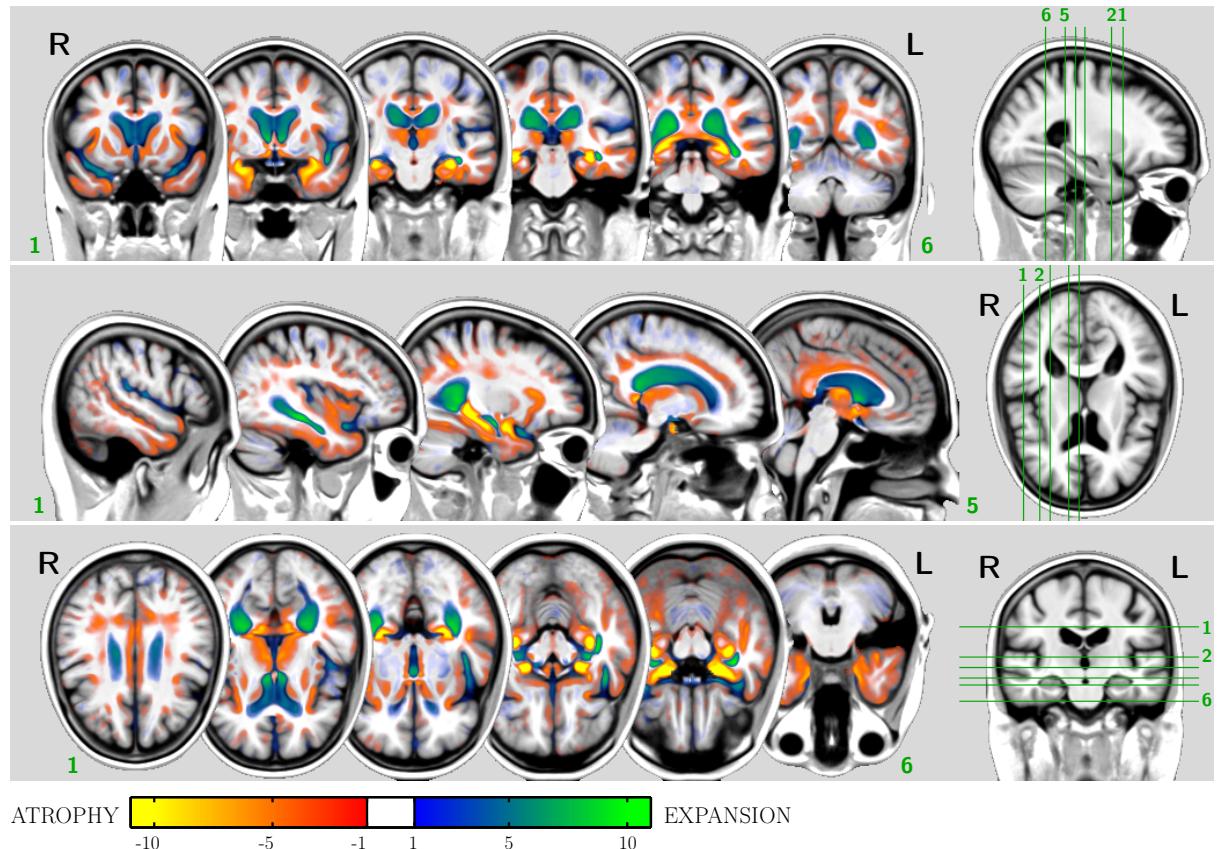


Figure 6.2.4. Voxel-wise t -value on logarithm of determinant of Jacobian matrices between $\{\text{NOR}\}$ and $\{\text{AD}\}$ groups.

6.2.4 Assessment of statistical significance

In a hypothesis testing experiment, the hypothesis is usually specified in terms of a test statistic, which is typically selected in such a way to stand out differences between the null and the alternative hypothesis. In the present study the chosen statistic is the Student's t -statistic which corresponds to a measurement about the difference in the means from both groups. However, as important as the statistic value is its p -value which gives a degree of confidence to the

rejection or acceptance of the null hypothesis. For the case of a t -statistic, its p -value can be assessed parametrically or non-parametrically. Assuming that samples from both groups follows a Gaussian distribution, the distribution of t -statistic is approximated as a Student's t distribution with a degrees of freedom dependent of the variances from both samples. The p -value is accordingly obtained by integrating the probability from the tails of the Student's t distribution. On the other hand and without making strong assumptions on the underlying distribution of the data, p -value can be empirically assessed in a random permutation test as it was done in Section 6.1.6.

In a *permutation test* for a group comparison experiment, the null hypothesis is simulated by randomly permuting the group labels among all instances. The distribution of the statistic under \mathcal{H}_0 is empirically estimated, by computing the test for all $(|X|+|Y|)!/|X||Y|!$ possible allocations of the group labels (being $|X|$ and $|Y|$ the sizes of the groups). The permutation test provides a formal and simple mechanism to quantify the “surprise” in terms of probability of obtain a given value of the statistic. The p -value of the test is the proportion of random permutations having a statistic value larger or equal¹⁰ than the value of the statistic computed on the original group partition (the unpermuted partition). This randomisation test provides a control on *type I errors*, *i.e.* it controls the probability of rejecting a true null hypothesis. However, the full randomisation experiment is impractical to be accomplished. For example, the number of possible permutations for two sample groups with 20 instances each is about 1.4×10^{11} which is almost impossible to be conducted (in the present case it would suppose about 6.9×10^{127} permutations!). Therefore, the p -value is typically estimated by a *Monte Carlo evaluation test* [Ridgway 2009] were a sufficiently large number of random permutations are performed. For the present study, 20 000 random permutations were conducted to estimate the p -value. Following [Edgington 1969] [Anderson and Robinson 2001] [Ridgway 2009], the estimated p -value from a random subset of N_p permutations may be assumed to be approximately normally distributed with mean p and variance $p(1-p)/N_p$. Thus, when conducting 20 000 random permutations, empirically estimated p -value of 0.10 or 0.05 can be assumed with a variance of $\sim 4.5 \times 10^{-6}$ or $\sim 2.4 \times 10^{-6}$, respectively .

In a TBM study, due to the absence of an *a-priori* anatomical hypothesis, entire compression/expansion maps are assessed in the finding of significant experimental effects involving a large number of hypothesis tests. When a single test is performed at a level of confidence α , which means that, by chance alone, α out of 100 tests will incorrectly reject \mathcal{H}_0 . Therefore, the overall probability of getting type I errors clearly increases when multiple statistical tests are performed and it becomes more likely that the groups being compared will appear to differ by chance, in at least one attribute. Therefore, the greater the number of simultaneous test are performed the more strict confidence levels are demanded to meaningfully reject the overall null hypothesis. If N independent simultaneous tests are performed, and \mathcal{H}_0 is assumed true in all of them, the chance of occurrence of at least one incorrect rejection increases up to $\bar{\alpha} = 1 - (1 - \alpha)^N$ (known as *Šidák correction*). For example, for a typical value $\alpha = 5\%$, there is a chance of about $\bar{\alpha} = 40\%$ of rejecting at least one null hypothesis when 10 test are assessed simultaneously. Moreover, this probability increases up to $\bar{\alpha} = 92.3\%$ for $N = 50$ simultaneous tests. Without the assumption that comparisons are independent, then $\bar{\alpha} \leq N\alpha$ is commonly used, which is known as *Bonferroni correction*. The *multiple comparisons problem* addresses these simultaneous statistical inferences and proposes different mechanisms to control the chance of false-positives results when several tests are performed. The multiple comparisons problem is an active area of research in the neuroimaging area, in particular within the functional neuroimaging community and in statistical parametric mapping tools.

¹⁰ In the case that the null hypothesis is rejected for small enough values of the statistic, the opposite proportion must be considered.

The previous mechanisms, Šidák and Bonferroni corrections, are considered well posed adjustments under modest numbers of comparisons but are considerably conservative for applications where hundred thousands or millions of tests are taken into account (take into account that about 1.8 millions of voxels are involved in the brain parenchyma of the template). Moreover, in the study of spatially distributed responses adjacent results commonly tend to be highly correlated. This situation is common in functional MRI studies, as well as in VBM and TBM studies which is the case we are dealing with. Two less conservative mechanisms have been proposed to make comparable the significance levels from single and from multiple simultaneous hypothesis testing for the case of large-scale multiple comparisons.

In the first, it is assumed that null hypothesis is true in any individual test as opposite to the alternative hypothesis that groups are different if any of the individual tests result positive. This mechanism belongs to the *family-wise error rate* (FWER) control mechanisms. The correction is based on Monte Carlo simulation performed by permuting the group labels. On each permutation, individual p -value at each voxel is computed (which can be assessed by permutations test or parametrically if enough certainty about the data is known). The minimum of these p -values across the whole test collection is computed and the distribution of these minimum values is empirically estimated throughout the random permutations. Finally, the corrected significance at a confidence level α is obtained as the α -quantile of the estimated distribution [Nichols and Holmes 2002] [Pantazis et al. 2004] [Styner et al. 2006b].

In some situations, the control of type I errors provided by FWER criterion is so strict that the large number of type II errors (false-negative errors) might be of more concern than number of type I errors. This motivated the second usually employed control mechanism called *false-discovery rate* (FDR) correction [Benjamini and Hochberg 1995] [Genovese et al. 2002] [Storey 2003] [Chumbley and Friston 2009]. The FDR criterion aims to provide a correction on the proportion of false-positives among the rejected null hypotheses, rather than the probability of falsely rejecting a single null hypothesis.

Both approaches, FWER and FDR, are equivalent if all null hypotheses are true, *i.e.* if there are no significant voxels anywhere. However, FDR procedure exerts a less strict control over false discoveries compared to FWER. The FDR mechanism seeks to reduce the probability of even one false discovery, as opposed to the expected proportion of false discoveries. FDR provides an adaptive criterion with higher statistical power than FWER. The FDR correction is computed by [Genovese et al. 2002] [Styner et al. 2007]: first select the desired FDR bound q (this is the maximum proportion of false-positives among the significant tests that you are willing to tolerate on average); sort the p -values from smallest to largest; let i be index of the largest sorted p -value such that $p_i \leq i q/N$, with N the number of performed tests; declare all individual tests with a p -value smaller or equal than p_i significant at level q .

Figure 6.2.5 shows the FDR corrected significance map from the TBM study between {NOR}- {AD} groups comparison. Individual non-corrected p -values were obtained in two ways: parametrically under the assumption that $\log \det(J)$ features are Gaussian distributed [Bossa et al. 2010b]; and non-parametrically by empirically estimating the distribution of the statistic under the null hypothesis by means of 20 000 random permutations of the group labels. Both estimations yielded similar results. Blue-green colors depict grown regions and red-yellow depict atrophic regions (these “signs” have been taken from the comparison of the mean of $\log \det(J)$ from both groups, see also Figure 6.2.4). The anatomical regions with significative differences are in agreement with anatomical structures affected by dementia. The following anatomical structures show significant atrophies: superior temporal sulcus; posterior part of the cingulate gyrus (precuneus region); temporo-occipital sulcus; hippocampus, mainly affecting subiculum and CA1 regions; entorhinal cortex and parahippocampal gyrus; amygdala; temporal pole. Also significant regions are detected in the temporal cortex and around the limbic system. Regarding expansions, high

significance levels are found at the ventricular system and at the insula. These expansions are indirect findings of the neurodegeneration which occur because the intracranial volume must be preserved. Although indirect, the expansions provides a pattern of the anatomical changes in the disease condition.

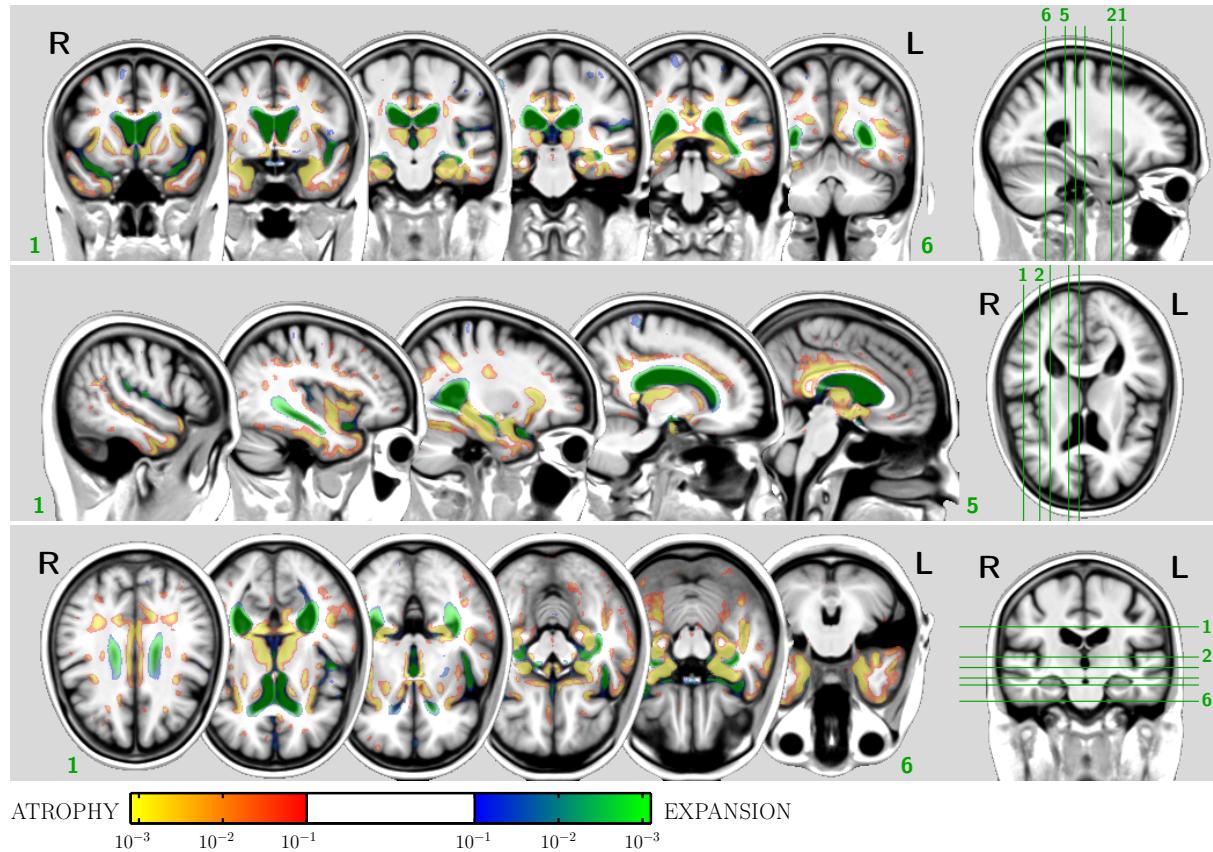


Figure 6.2.5. FDR corrected significance map of t -test statistic from the {NOR}-{AD} group comparison shown in Figure 6.2.4.

The difference between FDR and FWER control mechanisms is illustrated in Figure 6.2.6 for the TBM study between {NOR}-{AD} groups. Left panel shows that the anatomical regions with significant differences at the 5 % confidence level corrected with FDR (red shaded regions) are larger than the significant regions under FWER (cyan outlined regions) criterion at the same significance level. This behaviour was expected because FDR provides a less strict control over false positives than FWER. Right panel shows the volume (# of voxels) of significant regions with a variable confidence level, illustrating the difference between both correction criteria. It can be also shown, that parametric and non-parametric estimation of uncorrected p -values yields almost the same results. The parametric estimation assumes Gaussian distribution of the logarithms of determinants and is computed by integrating the tails of the Student's t distribution. On the other hand, the non-parametric estimation of uncorrected p -values was performed by permutation test. The non-parametric estimation have a discreteness from the finite number of random permutations and all p -values are multiples of $1/N_p$. In this experiment, N_p equals to 20 000 with a lowest achievable p -value of 5×10^{-5} . This discreteness of the non-parametric estimation of the uncorrected p -value avoids to resolve the distribution of the minimum p -value across all tests [Pantazis et al. 2005]. If the 5 % of the minimum p -value is desired, about 2.5 millions permutations test are required to be carried out.

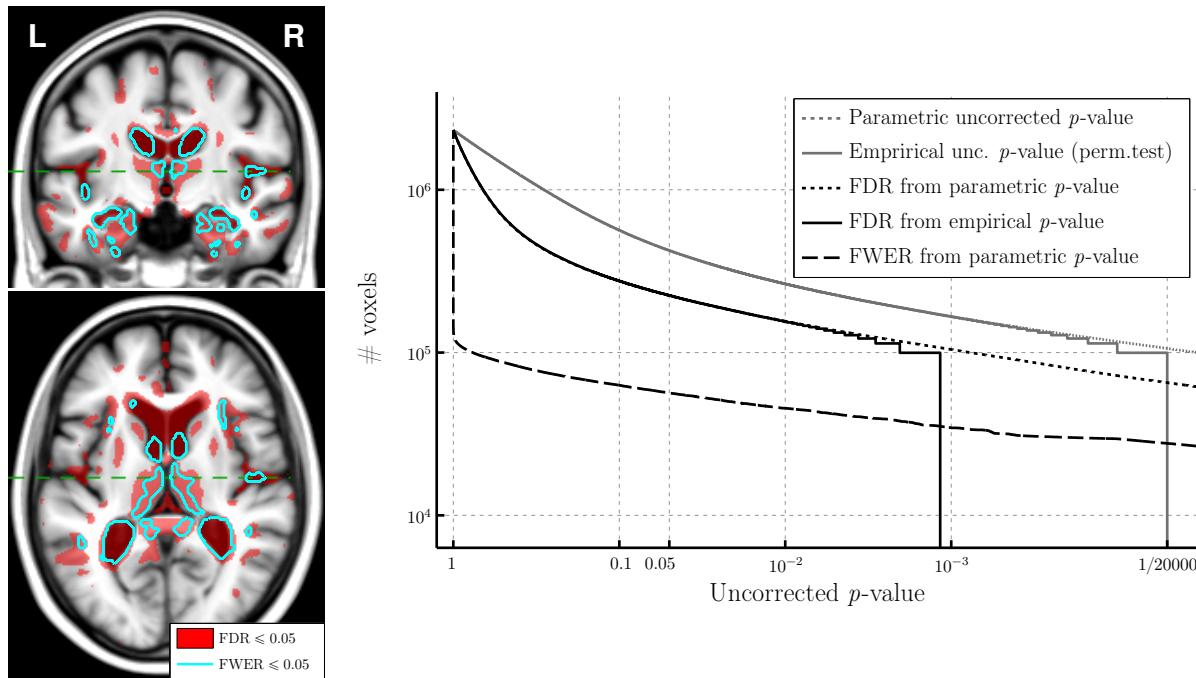


Figure 6.2.6. Differences between False Discovery Rate (FDR) correction mechanism and Family Wise Error Rate (FWER) correction. Left: coronal and axial views showing significant regions at 5 % confidence level for FDR and FWER control mechanisms. Right: Volume of the regions (number of voxels) versus significance level for uncorrected p -value, FDR and FWER corrected significance maps. Non-parametrically assessed p -values result discrete because the finite number of permutations N_p and the lowest achievable value is 5×10^{-5} .

6.3 Multivariate tensor-based morphometry

In the previous section the TBM methodology was performed by considering the determinant of the Jacobian matrix as the feature that encodes the geometrical information of the anatomical changes between the template and subjects. Because of familiarity with scalar data, both in terms of visualization and analysis, and due to the intuitive interpretation of the results, the determinant feature is the most commonly used quantity. However, the main limitation of the Jacobian determinant is that it is an unspecific descriptor of the geometrical change, providing a coarse description of the deformation, because it only quantifies local volume changes. In fact, there are many possible deformations without local volume change which may be relevant for a specific application. To overcome this limitation, several multivariate descriptors of the Jacobian matrix have been proposed. For a complete review of multivariate features from the Jacobian matrix see [Ridgway 2009, Chapter 4]. An example of a multivariate feature is the Cauchy–Green deformation tensor [Lepore et al. 2008, 2006]. This is a *symmetric positive definite* (SPD) tensor and it is represented by the matrix $C(y) = J(y)^T J(y)$. The analysis of *diffusion tensor images* (DTI) has fostered the development of formal statistical tools for the analysis of SPD tensors [Moakher 2005] [Batchelor et al. 2005] [Pennec et al. 2006] [Schwartzman 2006] [Fletcher and Joshi 2007] [Arsigny et al. 2007] [Boucher et al. 2011].

6.3.1 Invariance with respect to the template

In TBM, the inference is performed at each template coordinate using the local description of the deformation mapping. The results are laid out on a statistical map that reflects the spatial distribution of the voxel-wise inference. It is desired and expected to mainly obtain the same result for any template choice allowing to make general statements about the anatomical location of the findings. In the following it will be discussed that the template invariance requirement can be achieved by using a symmetric and transitive registration procedure [Thompson and Toga 1997] [Christensen and Johnson 2003] [Škrinjar et al. 2008]. These properties ensure that an anatomical label can be assigned to each location of the images irrespectively of the template choice. To our knowledge very few registration procedures guarantee these requirements (see for example [Škrinjar et al. 2008]).

Let T and W be two possible templates and let $w = \mathcal{R}(W, T)$ be the differentiable and invertible deformation mapping relating W and T , *i.e.* $w \star W$ looks like T . Let $h_\kappa = \mathcal{R}(T, I_\kappa)$ be the deformation mapping to register T towards I_κ , and let $\tilde{h}_\kappa = \mathcal{R}(W, I_\kappa)$ be the deformation mapping corresponding to the registration of the template W towards the same image I_κ . Assuming that the registration process is transitive and symmetric, the following relations between the registration results hold

$$\begin{aligned} \mathcal{R}(W, T) &= \mathcal{R}(T, W)^{-1} \\ \tilde{h}_\kappa = \mathcal{R}(W, I_\kappa) &= \mathcal{R}(T, I_\kappa) \bullet \mathcal{R}(W, T) = h_\kappa \bullet w \end{aligned} \quad (6.3.1)$$

The transitivity property of the deformation mappings is schematically illustrated in Figure 6.3.1. The mapping h_κ relates a point y of the template T to its corresponding point y_κ in the image I_κ . If the template W is used instead, the mapping \tilde{h}_κ relates each point \tilde{y} of the template W to the same point in image domain. Due to the transitivity property and following Eq. (6.3.1), the mapping relating both templates will be $w = h_\kappa^{-1} \bullet \tilde{h}_\kappa$, and each point \tilde{y} is in correspondence with the point $y = w \star \tilde{y}$. Within this framework each point y of the template T can be given an anatomical label which coincides with the one in $\tilde{y} = w^{-1} \star y$ of the template W . The anatomical label, in coordinates from either T or W , can be propagated to the corresponding instance location y_κ through the mappings either h_κ or \tilde{h}_κ .

Under the previous framework and by making use of the chain rule¹¹ on Eq. (6.3.1), the Jacobian matrix of \tilde{h}_κ evaluated at a point \tilde{y} is given by $\tilde{J}_\kappa(\tilde{y}) = J_\kappa(w \star \tilde{y})P(\tilde{y})$, where $P(\tilde{y}) = \mathcal{D}_z(\mathcal{R}(W, T) \star z)|_{z=\tilde{y}}$. The location \tilde{y} in the template W corresponds to the same anatomical location $y = w \star \tilde{y}$ in the template T and it is expected that Jacobian matrices $\tilde{J}_\kappa(\tilde{y})$ and $J_\kappa(y) = J_\kappa(w \star \tilde{y})$ provide the same findings about that location. Therefore, under a change of template, the information provided by Jacobian matrices transform as $J \mapsto JP$ with P being the Jacobian matrix of the mapping w between both templates. In terms of the registration function $\mathcal{R}(\cdot, \cdot)$ the effect of changing the template is written as

$$\mathcal{D}_z(\mathcal{R}(W, I_\kappa) \star z)|_{(z=\tilde{y})} = \mathcal{D}_z(\mathcal{R}(T, I_\kappa) \star z)|_{(z=\mathcal{R}(W, T) \star \tilde{y})} \mathcal{D}_z(\mathcal{R}(W, T) \star z)|_{(z=\tilde{y})} .$$

In TBM, statistical analysis is performed voxel-wise at each template coordinate, either in y coordinates for the template T or equivalently in \tilde{y} coordinates for W (see Figure 6.3.1). As the choice of the template is arbitrary, subsequent statistical analysis should be independent of that choice to avoid biased results. When a statistic based on distances is used, the template

¹¹ It have to remark that the composition operation $(\cdot \bullet \cdot)$ is equivalent to the function composition operation $(\cdot \circ \cdot)$ (see Section 2.1) and therefore derivatives of the composition operation follows the usual chain rule.

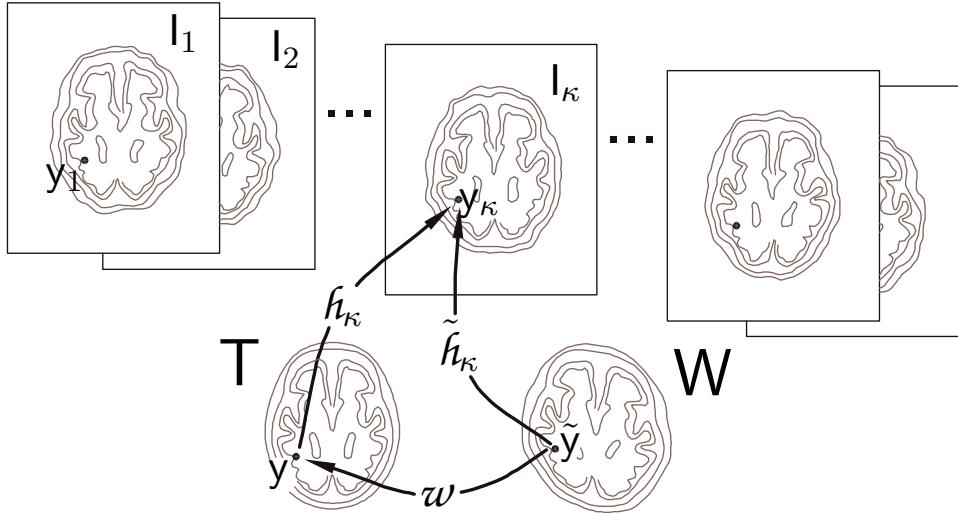


Figure 6.3.1. Illustration of template invariance in a TBM study. Using the transitivity property assumption of the mappings, each image location and its corresponding location for any template can be given with the same anatomical label.

invariance is achieved if the following sufficient condition on the distance function¹² $\mathbf{d}(\cdot, \cdot)$ fulfils

$$\begin{aligned}\mathbf{d}(\tilde{J}_\kappa(\tilde{y}), \tilde{J}_{\kappa'}(\tilde{y})) &= \mathbf{d}(J_\kappa(y)P(\tilde{y}), J_{\kappa'}(y)P(\tilde{y})) \\ &= \mathbf{d}(J_\kappa(y), J_{\kappa'}(y)),\end{aligned}$$

where $y = w * \tilde{y}$. Therefore template invariance will hold when the statistic is based on a right-invariant distance, *i.e.* a distance between instances κ and κ' fulfilling

$$\mathbf{d}(J_\kappa(y)P, J_{\kappa'}(y)P) = \mathbf{d}(J_\kappa(y), J_{\kappa'}(y)) \quad (6.3.2)$$

for any $J_\kappa(y), J_{\kappa'}(y), P \in \mathbf{GL}^+(d)$ which is the space where Jacobian matrices belong to.

Henceforth, for brevity of the notation, the dependence of the spatial variable will be omitted, and J will refer to the Jacobian matrix at the location under study.

A brief discussion is included below regarding the transitivity assumption. The transitivity property is not fully accomplished by most of the registration algorithms currently available, with few exceptions such as [Škrinjar et al. 2008].

If it had unique spatial transformations $h_\kappa \in \mathcal{H}_A$ fulfilling $h_\kappa * T = I_\kappa$ (with the equality holding strictly) and a unique spatial transformation w fulfilling $w * W = T$, then the transitivity property would be automatically achieved. However, in the practice this is not the typical case for several reasons. Firstly, because images usually have flat regions where small changes in deformation mappings results in the same deformed image. Secondly, image registration process is usually performed in a smaller set than \mathcal{H}_A and the equality relation is not fully achieved. And thirdly, acquired images are usually perturbed by noise and images are not strictly a deformed version of the template. Therefore, the non-transitivity implies that, among all homeomorphism of A , there exist more than one possible deformation mapping which can be understood as an uncertainty in the mapping $\mathcal{R}(T, I_\kappa)$. Also, it can be understood as that the exact spatial correspondence for all the points in the ambient space is unknown, and registration process only provides a reasonable estimation of the correspondences [Rohlfing 2012].

¹² In previous sections and chapters, distance functions were denoted by **distance**(\cdot, \cdot) but, in order to alleviate the notation, from now on they will be denoted by $\mathbf{d}(\cdot, \cdot)$.

In a general case the mappings h_κ and \tilde{h}_κ from the templates T and W respectively to each instance κ shown in Figure 6.3.1 can be written using an instance-dependent right-translation

$$\tilde{h}_\kappa = h_\kappa \bullet w_\kappa \quad (6.3.3)$$

and it is straightforward that

$$h_\kappa \bullet w_\kappa \bullet \tilde{h}_\kappa^{-1} = \text{id}_{\mathbb{A}} .$$

If a transitive registration algorithm had been used, all right-translations w_κ in Eq. (6.3.3) would be the same, $w_\kappa = \mathcal{R}(W, T) = w$ for all instances κ . The difference between the mapping $h_\kappa \bullet w \bullet \tilde{h}_\kappa^{-1}$ and the identity map $\text{id}_{\mathbb{A}}$ gives a measurement of the impact of the non-transitivity in a TBM study. When these differences are small enough relative to the intergroup deformation amount, then the non-transitivity could be considered a not a relevant issue.

Jacobian determinant

In Section 6.2.3 it was argued that most TBM studies perform voxel-wise statistics on the determinant of Jacobian matrices and more specifically on their logarithm. Here the logarithmic transform will be justified in accordance to the group structure suggested by the composition of spatial transformations and the invariance requested to the analysis.

The determinant of a composition of deformations is the product of their determinants. Therefore, determinant of Jacobian matrices belong to the group of positive numbers under multiplication which can be identified with $S^+(1)$ group. Following results from Chapter 5, an invariant distance on this Lie group is given by $\mathbf{d}_{S^+(1)}(s, r) = \log(s/r) = |\log(s) - \log(r)|$, where $s, r \in \mathbb{R}^+$. Distance $\mathbf{d}_{S^+(1)}$ induces the following distance on Jacobian matrices¹³

$$\mathbf{d}_{DET}(J_\kappa, J_{\kappa'}) = |\log(\det(J_\kappa)) - \log(\det(J_{\kappa'}))| .$$

It is easy to show that $\mathbf{d}_{DET}(\cdot, \cdot)$ fulfils Eq. (6.3.2) and therefore the template invariance requirement.

The determinant of a Jacobian matrix quantifies the local volume change induced by the deformation. However, note that $\mathbf{d}_{DET}(J_\kappa, J_{\kappa'}) = 0$ does not imply that $J_\kappa = J_{\kappa'}$. Then, \mathbf{d}_{DET} does not satisfy the *coincidence axiom* ($\mathbf{d}(x, y) = 0$ if and only if $x = y$), also known as the *identity of indiscernibles* principle. More precisely, $\mathbf{d}_{DET}(J, LJ) = \mathbf{d}_{DET}(J, JL) = 0$ for any matrix L with $\det(L) = 1$.

Deformation tensor

In the area of continuum mechanics, the Cauchy–Green deformation tensor $C = J^T J$ is a commonly used feature to measure the local deformation in a Lagrangian framework [Lubliner 2008, Chapter 8] [Ciarlet 1988] [Yavari et al. 2006] [Pennec 2006b]. The tensor C is a symmetric positive definite (SPD) matrix which measures changes of local lengths. Under a change of template, Jacobian matrices transform as $J \mapsto JP$ and therefore Cauchy–Green deformation tensors transform as $C \mapsto P^T CP$.

In [Moakher 2005] [Batchelor et al. 2005] a distance between SPD matrices have been proposed¹⁴ $\mathbf{d}_{SPD}(C_\kappa, C_{\kappa'}) = \|\log((C_\kappa)^{-1/2}(C_{\kappa'})(C_\kappa)^{-1/2})\|_F$ satisfying the following invariance:

¹³ Strictly, \mathbf{d}_{DET} is not a proper distance function but can be a *pseudo-distance* function. Anyway, from the practical point of view, the statistical analysis can be performed in the same way considering that the analysis is being performing with a proper *distance function* between *features* of the whole Jacobian matrices.

¹⁴ In the following, function $\log(\cdot)$ refers to the matrix logarithm, not to the group logarithm function as in Chapter 4.2.5. SPD matrices do not form a group under matrix multiplication.

$\mathbf{d}_{SPD}(C_\kappa, C_{\kappa'}) = \mathbf{d}_{SPD}(P^T C_\kappa P, P^T C_{\kappa'} P)$ for any non-singular linear transformation P . The distance \mathbf{d}_{SPD} induces a distance between Jacobian matrices

$$\mathbf{d}_{AFF}(J_\kappa, J_{\kappa'}) = \left\| \log \left((J_\kappa^T J_\kappa)^{-1/2} (J_{\kappa'}^T J_{\kappa'}) (J_\kappa^T J_\kappa)^{-1/2} \right) \right\|_F$$

satisfying the right-invariance property $\mathbf{d}_{AFF}(J_\kappa, J_{\kappa'}) = \mathbf{d}_{AFF}(J_\kappa P, J_{\kappa'} P)$ for any matrix P with positive determinant, and therefore fulfils the template invariance requirement given in Eq. (6.3.2).

As in the case of \mathbf{d}_{DET} , there may exists a pair of different Jacobian matrices with a null distance $\mathbf{d}_{AFF}(\cdot, \cdot)$ between them. Specifically, $\mathbf{d}_{AFF}(J, RJ) = 0$ for any rotation matrix R . Note that, for any Jacobian matrix J , the set of RJ is a proper subset of LJ , where $\det(L) = 1$. Therefore \mathbf{d}_{AFF} is sensitive to a larger set of deformations than \mathbf{d}_{DET} . Alternatively, it can be thought that C is a more complete descriptor than $\det(\cdot)$, and by using appropriate distances on it, better sensibility will be obtained.

Another distance between SPD matrices is commonly used in TBM studies [Lepore et al. 2008] [Boucher et al. 2011]. Under the Log-Euclidean framework [Arsigny et al. 2005, 2007], a distance on SPD matrices, $\mathbf{d}_{LE}(C_\kappa, C_{\kappa'}) = \|\log(C_\kappa) - \log(C_{\kappa'})\|_F$, was proposed. However, this distance is not invariant under general linear transformations although it is invariant only under transformations from the $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$ group (if $G = sR$ is a rotation followed by an isotropic positive scaling, then $\mathbf{d}_{LE}(C_\kappa, C_{\kappa'}) = \mathbf{d}_{LE}(G^T C_\kappa G, G^T C_{\kappa'} G)$).

A proper distance function on $\mathbf{GL}^+(d)$

The two previous distances \mathbf{d}_{DET} and \mathbf{d}_{AFF} fulfil the template invariance requirement. However they do not satisfy the property that $\mathbf{d}(J_\kappa, J_{\kappa'}) = 0$ if and only if $J_\kappa = J_{\kappa'}$ and accordingly some geometrical differences may not be measured. To overcome this drawback a Riemannian metric on the space of Jacobian matrices can be used. Riemannian metrics were defined in Section 4.2.6 and they can be used to measure the length of a curve in the space of Jacobian matrices. The shortest curve between two given Jacobian matrices defines the Riemannian distance function. This distance fulfils all the distance axioms, including the coincidence axiom. Therefore, Riemannian distances on $\mathbf{GL}^+(d)$ seem to be good candidates for a distance function $\mathbf{d}(\cdot, \cdot)$ to measure in full differences between Jacobian matrices. In particular, right-invariant distances, derived from right-invariant Riemannian metrics (see Section 4.2.8), will also fulfil the template invariance requirement.

The concept of Riemannian distance is closely related to the Riemannian exponential function $\text{Exp}(\cdot)$ in the sense that the length of a curve segment $\gamma_{(t)} = \text{Exp}_Q(tV)$ between $t = 0$ to $t = 1$ is $\|V\|_Q = \langle V, V \rangle_Q^{1/2}$. An expression for the left-invariant Exponential function $\text{Exp}_I(\cdot)$ on $\mathbf{GL}^+(d)$ was already given in Section 5.1.8 for the particular metric case $\langle U_1, U_2 \rangle = \text{trace}(U_1^T U_2)$:

$$\text{Exp}_Q(V) = Q \exp(V^T Q^{-T}) \exp(Q^{-1}V - V^T Q^{-T}).$$

In order to obtain an expression for the right-invariant case, the isometry between Exp^{left} and Exp^{right} given in Eq. (4.2.12) can be used, resulting in

$$\begin{aligned} \text{Exp}_Q^{right}(V) &= \left(\text{Exp}_{Q^{-1}}^{left}(-Q^{-1}VQ^{-1}) \right)^{-1} \\ &= \exp(VQ^{-1} - Q^{-T}V^T) \exp(Q^{-T}V^T) Q \end{aligned}$$

and, in particular for $Q = I$,

$$\text{Exp}_I^{right}(U) = \exp(U - U^T) \exp(U^T), \quad (6.3.4)$$

where, again, the metric is defined by the inner product at the identity $\langle\!\langle U_1, U_2 \rangle\!\rangle_I = \text{trace}(U_1^T U_2)$.

From the right-invariant Riemannian metric on $\mathbf{GL}^+(d)$ the following distance is induced on elements of the same connected component of the group: $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = \langle\!\langle V^*, V^* \rangle\!\rangle_{J_\kappa}^{1/2}$, where V^* is the smallest initial velocity satisfying $\text{Exp}_{J_\kappa}(V) = J_{\kappa'}$. The distance $\mathbf{d}_{RI}(\cdot, \cdot)$ inherits the right-invariance from the Riemannian metric, and therefore $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = \mathbf{d}_{RI}(J_\kappa P, J_{\kappa'} P)$ for any P in $\mathbf{GL}^+(d)$. Thus, the invariance under the template holds. In addition, \mathbf{d}_{RI} fulfils the coincidence axiom, $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = 0$ if and only if $J_\kappa = J_{\kappa'}$ and it is a proper distance function over the full description of Jacobian matrices.

The Riemannian distance can be computed from the Riemannian logarithm function as $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = \|\text{Log}_{J_\kappa}(J_{\kappa'})\|_{J_\kappa}$, which results in the following problem (see also Section 5.2.6):

$$\underset{V \in T_{J_\kappa} \mathbf{GL}^+(d)}{\text{minimize}} \quad \|\text{Exp}_{J_\kappa}(V) - J_{\kappa'}\|_F^2 \quad , \quad (\text{logGL1})$$

where the Frobenius norm of the Euclidean difference is chosen as the objective function for simplicity. Using the right-invariance property, problem (logGL1) can be reformulated as

$$\underset{U \in \mathbf{gl}(d)}{\text{minimize}} \quad \|\text{Exp}_I(U) - J_{\kappa'} J_\kappa^{-1}\|_F^2 \quad , \quad (\text{logGL2})$$

where $\mathbf{gl}(d) \equiv \mathbb{M}_d$ is the Lie algebra of $\mathbf{GL}^+(d)$. Due to the right-invariance of the metric, $\langle\!\langle U^*, U^* \rangle\!\rangle_I = \langle\!\langle V^*, V^* \rangle\!\rangle_{J_\kappa}$ and therefore, $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = \langle\!\langle U^*, U^* \rangle\!\rangle_I^{1/2}$, where U^* is the smallest solution of (logGL2).

In order to compute U^* , a descent procedure can be performed on (logGL2). The derivative of the objective function $E(U; Q) = \|\text{Exp}_I(U) - Q\|_F^2$, with $Q = J_{\kappa'} J_\kappa^{-1}$, with respect to U is

$$\mathcal{D}_U E(U; Q) = 2 \left(\overline{\text{Exp}_I(U) - Q} \right)^T \mathcal{D}_U \text{Exp}_I(U) . \quad (6.3.5)$$

Taking derivatives of Eq. (6.3.4) it is obtained (see footnote 5 on page 155)

$$\begin{aligned} \mathcal{D}_U \text{Exp}_I(U) &= (I_d \otimes \exp(U - U^T)) \mathcal{D}\exp(U^T) K_{dd} + \\ &\quad + (\exp(U) \otimes I_d) \mathcal{D}\exp(U - U^T) (I_{d^2} - K_{dd}) , \end{aligned} \quad (6.3.6)$$

where I_m is the $m \times m$ identity matrix, K_{dd} is the commutation matrix [Abadir and Magnus 2005] defined by $K_{mm} \overline{M} = M^T$ for an $m \times m$ matrix M , and $\mathcal{D}\exp(\cdot)$ is the Fréchet derivative of the matrix exponential function (see Section 4.3.2).

Once the optimal velocity $U^* \in \mathbf{gl}(d)$ is obtained by solving (logGL2), the length of the geodesic segment from I to Q is calculated by $\langle\!\langle U^*, U^* \rangle\!\rangle_I^{1/2} = \text{trace}(U^{*T} U^*)^{1/2} = \|U^*\|_F$. If its Riemannian exponentiation generates the shortest curve segment between I and $J_{\kappa'} J_\kappa^{-1}$, then $\mathbf{d}_{RI}(J_\kappa, J_{\kappa'}) = \mathbf{d}_{RI}(I, J_{\kappa'} J_\kappa^{-1}) = \|U^*\|_F$.

Details for computing \mathbf{d}_{RI}

Regarding the existence of a solution for the problem of compute the initial velocity of a geodesic starting at the identity I to the target matrix $Q = J_{\kappa'} J_\kappa^{-1} \in \mathbf{GL}^+(d)$, it can noted that the set $\mathbf{GL}^+(d)$ is connected. As the group $\mathbf{GL}^+(d)$ consists of only one connected component, then the existence of an initial velocity to generate a geodesic connecting J_κ and $J_{\kappa'}$ is guaranteed for any pair. Moreover, $\text{Exp}_I(\cdot)$ in Eq. (6.3.4) is surjective (see *Remarks* in Section 4.2.8) on $\mathbf{GL}^+(d)$ and therefore there always exists at least one zero-minimizer of the objective function $E(U; Q) = \|\text{Exp}_I(U) - Q\|_F^2$.

While the existence of a zero minimizer is guaranteed, a serious drawback of this formulation is the non-uniqueness of the velocities satisfying $\text{Exp}_I(U) = J_{\kappa'}J_{\kappa}^{-1}$. There may exists different initial velocities which generate geodesic segments between J_{κ} and $J_{\kappa'}$ (like in the rotations case where adding a rotation of 2π the same rotation is obtained). Therefore, the objective function $E(U; Q)$ is in general not convex in $\mathfrak{gl}(d)$ and its minimization procedure may have many attraction basins. Nevertheless, it can be shown that all local minima have zero energy and therefore they are also global minima. This can be proved by noticing that the function $\text{Exp}_I(\cdot)$ is a local diffeomorphism and therefore $\text{Exp}_I(\cdot)$ maps open subsets of $\mathbb{M}_d \equiv \mathfrak{gl}(d)$ around U to open subsets of $\mathbb{M}_d \supset \mathbf{GL}^+(d)$ around $\text{Exp}_I(U)$. The same applies for the function $\text{Exp}_I^{-1}(\cdot)$. Moreover, the Frobenius norm is a convex function in \mathbb{M}_d . Then, for any $U \in \mathfrak{gl}(d)$ either, $E(U; Q)$ is zero or there is an U' in a neighborhood of U with a lower value of the objective function.

It is remarkable that for some target matrices Q there may exists descending *valleys* in the objective function which extends up to infinity. For example, in the extreme case of $Q = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, there exist a descent valley in $\mathfrak{gl}(d)$ along the direction $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. These valleys asymptotically ends up in an element of $\mathfrak{gl}(d)$ whose Riemannian exponential is the null matrix. Although we could not find a simple proof, we conjecture that those *attraction basins towards infinity* have a zero measure in $\mathfrak{gl}(d)$. For the previous matrix Q , descent procedures starting from any symmetric matrix tend to the null matrix, but by slightly perturbing with a skew-symmetric matrix the procedure converges to a global minimum. For a continuous descending path $U(t)$ such that $\|U(t)\|_F \rightarrow +\infty$ for $t \rightarrow +\infty$, it can be proved that $\lim_{t \rightarrow +\infty} \text{Exp}_I(U(t))$, either it is a singular matrix or it does not exist. Then, to circumvent those *valleys towards infinite* it is convenient to provide to the descent algorithm with a control in the determinant of $\text{Exp}_I(U_k)$ (with U_k the current iteration of the algorithm) and perform a random perturbation on U_k if the algorithm is converging to a singular matrix.

An initial estimate for U is needed to start the optimization procedure. One possibility is to initialize U as the solution of the initial velocity in a “simpler” subgroup of $\mathbf{GL}^+(d)$ where the Riemannian logarithm can be easily computed. Let $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$ be the group of matrices of the form sR , with $s > 0$ and $R \in \mathbf{SO}(d)$. Its algebra, $(\mathfrak{s}(1) \oplus \mathfrak{so}(d))$, is the set of matrices of the form $\alpha I_d + \Sigma$ where $\alpha \in \mathbb{R}$ and Σ is a $d \times d$ skew-symmetric matrix. Under the metric $\langle U_1, U_2 \rangle_I = \text{trace}(U_1^T U_2)$, the Riemannian exponential over $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$ generates curves which are also geodesics in $\mathbf{GL}^+(d)$, i.e. $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$ is a totally geodesic subgroup of $\mathbf{GL}^+(d)$ under the given metric (see Section 4.2.9). For a matrix $H = \alpha I_d + \Sigma \in (\mathfrak{s}(1) \oplus \mathfrak{so}(d))$ the Riemannian exponential function given by Eq. (6.3.4) results in $\exp(H - H^T) \exp(H^T) = \exp(H)$ because, in this particular case, $(H - H^T)$ and H^T commute. Therefore, the matrix logarithm of a matrix $G \in (\mathbf{S}^+(1) \times \mathbf{SO}(d))$ gives the solution of the problem (logGL2) when it is restricted to the group $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$. The proposed initial estimate is the matrix logarithm of the closest matrix, in the Frobenius norm sense, belonging to $(\mathbf{S}^+(1) \times \mathbf{SO}(d))$ to the target matrix $Q = J_{\kappa'}J_{\kappa}^{-1}$

$$U_0 = \log \left(\underset{G \in (\mathbf{S}^+(1) \times \mathbf{SO}(d))}{\text{argmin}} \|G - Q\|_F^2 \right). \quad (6.3.7)$$

To compute this, let $Q = ZDX^T$ be the singular value decomposition of the target matrix, then the matrix $G^* = \text{trace}(D)/dZ \text{diag}(1, 1, \dots, 1, \det(ZX^T)) X^T$ is the solution of the minimization problem in Eq. (6.3.7) [Lee et al. 2007].

To show the performance of descent strategies the following experiment was performed: one thousand 3×3 random matrices were generated with positive determinant and with several predefined condition numbers; for each target matrix a solution U^* of the problem (logGL2)

was computed by using two descent strategies: along the negated gradient direction given by Eqs. (6.3.5) and (6.3.6); and along the Gauss–Newton direction [Nocedal and Wright 2006] [Chong and Zak 2013]. Both descent strategies were performed with an exact line-search, for example, the golden section search. The stopping criterion was set to a relative error in the Frobenius norm smaller than 10^{-10} . Figure 6.3.2 shows in the left panel the number of iterations needed to reach the stopping criterion for different condition numbers of the target matrix. Right panel in Figure 6.3.2 shows the relative error for 100 random target matrices with a condition number of 10 for both descent strategies. It is very clear that Gauss–Newton strategy requires a much smaller number of iterations than gradient descent.

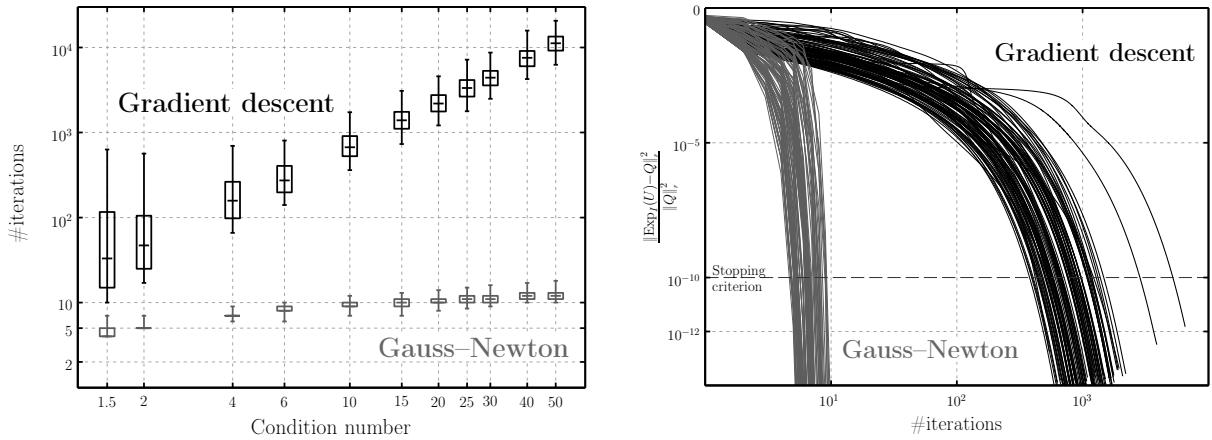


Figure 6.3.2. Performance of gradient descent and Gauss–Newton strategies using the golden section line-search. Left panel: box-plot showing 5, 25, 50, 75 and 95 percentiles of the number of iterations needed to reach the stopping criterion for target matrices with different condition numbers. Right panel: evolution of the relative error for 100 random target matrices with a condition number of 10.

6.3.2 Cramér test

Three different distance functions over Jacobian matrices (or, more precisely over features of Jacobian matrices) have been considered above: \mathbf{d}_{DET} ; \mathbf{d}_{AFF} ; and \mathbf{d}_{RI} . Most statistical analysis techniques can be defined in terms of a distance function between observations. In the current experiment, the two-sample Cramér test [Baringhaus and Franz 2004] [Székely and Rizzo 2004] [Whitcher et al. 2007] [Ridgway 2009] is used to assess differences between groups. Given two sets of Jacobian matrices $\{J_a\}_{a=1}^A$ and $\{J_b\}_{b=1}^B$, the statistic for the two-sample Cramér test is computed by

$$\sigma(\{J_a\}, \{J_b\}) = \frac{AB}{A+B} \left(\frac{1}{AB} \sum_{a=1}^A \sum_{b=1}^B D_{ab} - \frac{1}{2A^2} \sum_{a=1}^A \sum_{a'=1}^A D_{aa'} - \frac{1}{2B^2} \sum_{b=1}^B \sum_{b'=1}^B D_{b'b'} \right),$$

where $D_{\kappa\kappa'} = \mathbf{d}(J_\kappa, J_{\kappa'})$ is the inter-element distance and A and B is the cardinality of the groups $\{J_a\}$ and $\{J_b\}$, respectively.

In a TBM study, the statistic σ is calculated at each point of location y of the template and denoted by $\sigma_y = \sigma(\{J_a(y)\}, \{J_b(y)\})$. Also, at each location, statistical significance of σ_y can be assessed empirically by means of random permutations of group labels. The null hypothesis for this test is that instances from both groups follow the same distribution and it is rejected for large enough values of σ_y . Therefore, the p -value is the proportion of the permutations having a

σ_y value larger or equal than the value without relabeling. Note that the statistic for a random permutation can be obtained by permuting distances $D_{\kappa\kappa'}$, without the need to recalculate distances between Jacobian matrices, which is the most intensive computational task.

The Cramér test was selected because its statistic can be computed in terms of the inter-element distances solely. In addition, it can be directly used on both univariate and multivariate data. It may be of interest to extend the use of the presented Riemannian distance to another statistical tools such as the multivariate Hotelling's T^2 test. Unluckily, an optimization step will be required to estimate the intrinsic mean on $\mathbf{GL}^+(d)$ group (see Section 5.4.2). Moreover, up to our knowledge, there is no method to parametrically assess the significance of hypothesis test on Jacobian matrices and the use of random permutations would require to compute the intrinsic mean for each permutation leading in an even greater computational cost.

6.3.3 Synthetic study

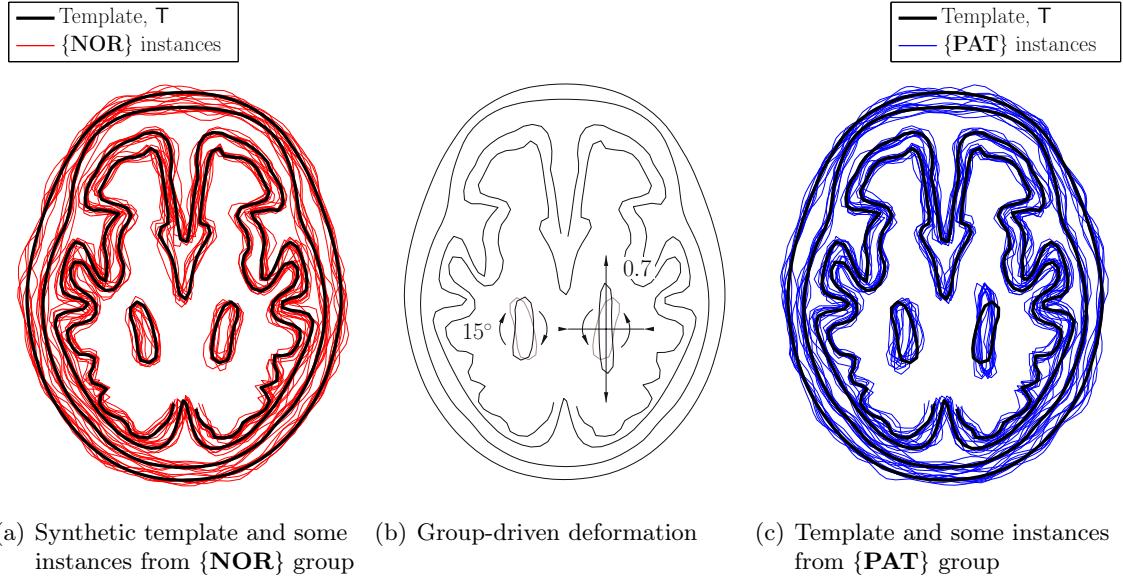
A synthetic study was designed to illustrate the results of the three right-invariant distances in a TBM application environment. In order to facilitate the visualization of the results this synthetic experiment was performed generating deformations on a 2D geometry. Two sets of 50 random deformations were generated starting from a synthetic template. The first set was designed to represent an anatomical variability within a “control” group, denoted by $\{\mathbf{NOR}\}_{n=1}^{50}$, while the second set is aimed at representing a “pathological” group with group-driven anatomical differences and intra-group variability, and denoted by $\{\mathbf{PAT}\}_{p=1}^{50}$. Figure 6.3.3 shows, at left, the synthetic template. Also, 10 instances from $\{\mathbf{NOR}\}$ group are shown to exemplify the intra-group variability which is modeled as smooth and invertible random deformation mapping. Instances from $\{\mathbf{PAT}\}$ group were modeled as the composition of a common group-driven deformation and random deformations similar to the ones used for $\{\mathbf{NOR}\}$ group. The group-driven deformation is illustrated at the central panel of Figure 6.3.3 and was designed to produce the following volume preserving changes on the ‘subcortical structures’: a clockwise rotation of the structure in the ‘left hemisphere’ of 15 degrees; a counter-clockwise rotation of the structure in the ‘right hemisphere’ of 15 degrees and a subsequent anisotropic scaling with factors 0.7 and 1/0.7 along the horizontal and vertical directions respectively. Right-panel of Figure 6.3.3 shows 10 instances from $\{\mathbf{PAT}\}$ group depicting the intra-group variability.

For this synthetic study, template and instances are contours described by equispaced knots and it will be assumed that the knots are in correspondence along all subjects. The template was registered towards all instances with a SVF registration by matching correspondent knots. Velocity fields were parameterized on a 220×280 regular grid and circular boundary conditions were imposed. Accordingly, Jacobian matrices $J(y)$ were computed by integrating up to time 1 the local deformation guided by the Jacobian of the velocity field as was explained in Section 2.5.4.

Representative examples of deformation mappings are shown in Figure 6.3.3. Note that although the group-driven deformation are simple volume preserving transformations of the ‘subcortical structures’, the surrounding regions suffer complex deformations due to the continuity of the mapping. This effect can be seen in the zoomed panel of Figure 6.3.4.

Figure 6.3.5 illustrates in detail a deformation mapping where the glyphs represent Jacobian matrices. In the left panel, Jacobian matrices are illustrated in their deformed coordinates while in the right panel Jacobian matrices are illustrated in the coordinate system of the template. Colors represent the magnitude of the Jacobian determinant in logarithmic scale.

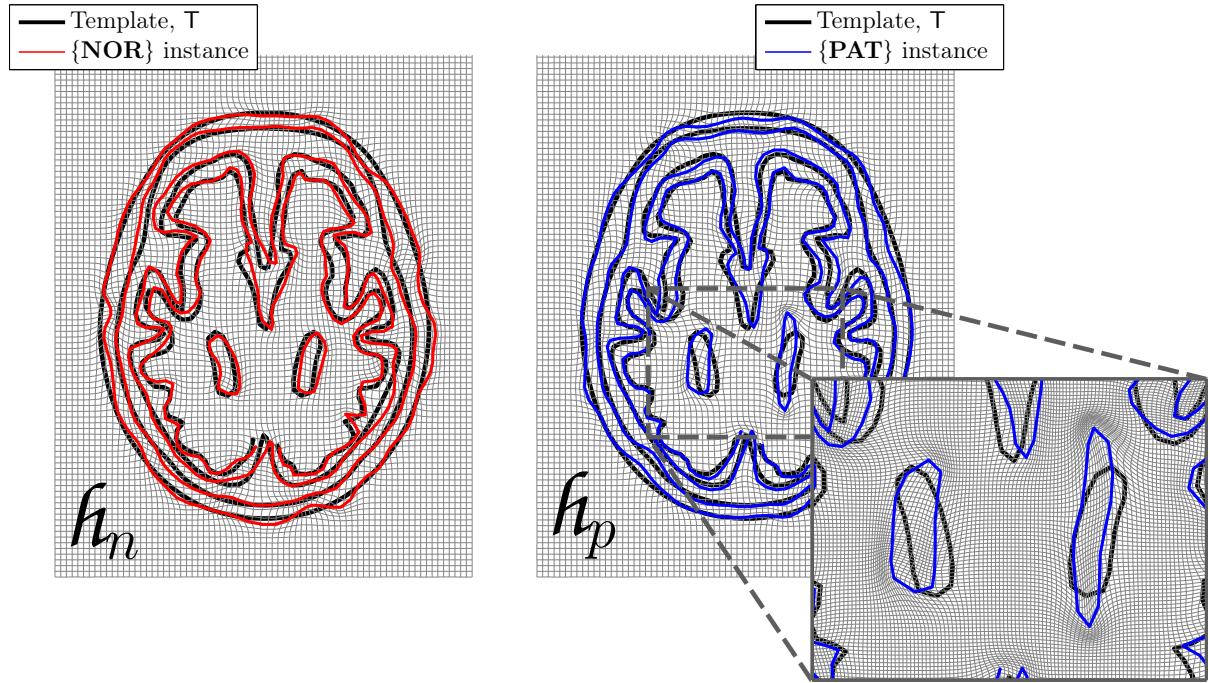
Statistical analyses were performed using the previous defined distances: \mathbf{d}_{DET} , \mathbf{d}_{AFF} and \mathbf{d}_{RI} . At each grid location of the template y , Cramér tests statistic $\sigma(\{J_n(y)\}, \{J_p(y)\})$ were computed. In order to assess its corresponding p -value, 100 000 random permutations of group



(a) Synthetic template and some instances from {NOR} group

(b) Group-driven deformation

(c) Template and some instances from {PAT} group

Figure 6.3.3. Synthetic study setup.**Figure 6.3.4.** Example of representative deformation mappings h_n and h_p between the template and {NOR} and {PAT} instances, respectively.

labels were performed by permuting the inter-element distances $D_{\kappa\kappa'}$. Finally, correction for multiple comparisons was achieved by the FDR criterion (see Section 6.2.4).

Figure 6.3.6 shows the FDR corrected significance maps corresponding to the Cramér tests based on the three distances. Using the distance \mathbf{d}_{DET} , the test was not able to detect statistically significant differences in the interior of any of the two 'subcortical structures'. However, significant differences were found at outer regions surrounding these structures. This illustrates that deformations driven by a rotation of a structure surrounded by another static structure mainly

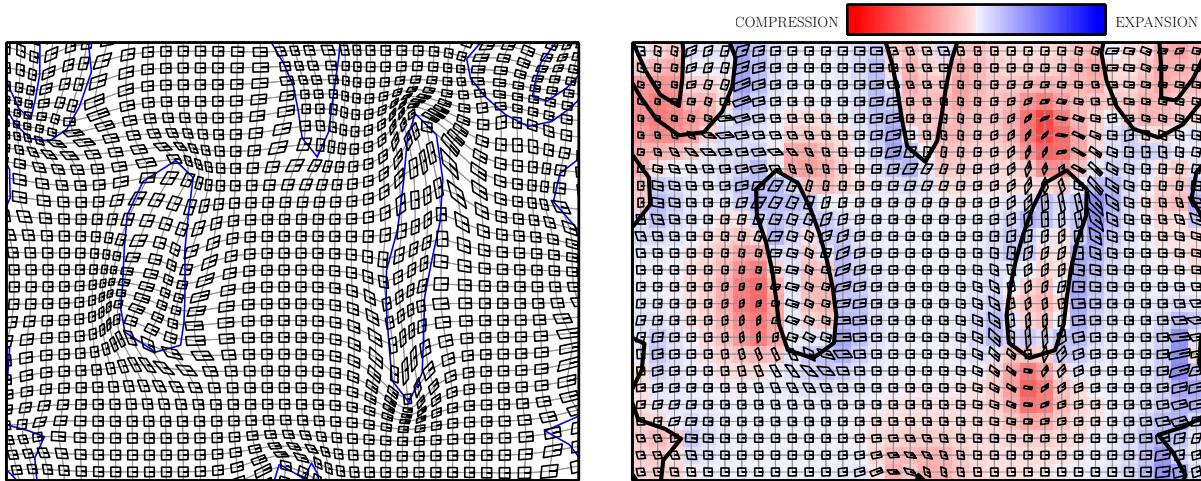


Figure 6.3.5. Glyph based illustration of the Jacobian matrix field from deformation mapping shown in zoom panel of Figure 6.3.4. Left panel: glyphs shown in the deformed coordinate system. Right panel: glyphs in the template coordinate system. Colors represent determinant magnitude in logarithmic scale.

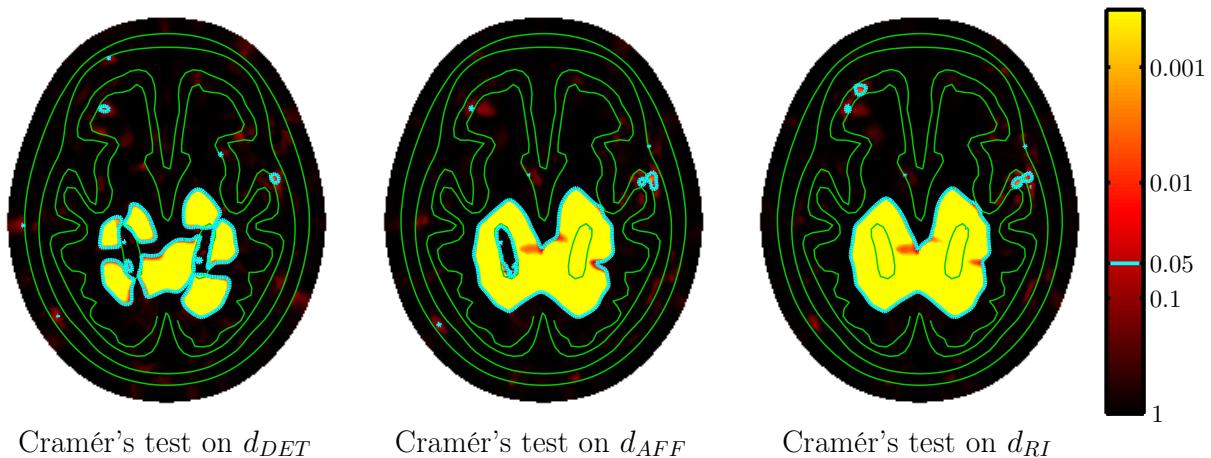


Figure 6.3.6. Statistical maps of FDR corrected significance for Cramér test based on the distances \mathbf{d}_{DET} , \mathbf{d}_{AFF} and \mathbf{d}_{RI} . Template contours are illustrated in green to aid the localization. Cyan contours outline the regions with significant differences with the criterion $FDR \leq 0.05$.

generates local volume changes outside the structure.

As was expected, the resulting statistical map using \mathbf{d}_{AFF} shows significant differences in the interior of the 'right subcortical structure', where there is an anisotropic scaling of the structure. However, no significant differences were found in the interior of the 'left subcortical structure' because deformations were mainly a rotation. Significant differences were also found in the surrounding regions.

Regarding the statistical map using \mathbf{d}_{RI} , it can be seen that significant differences were found in the interior of both 'subcortical structures'. Comparing statistical maps from \mathbf{d}_{AFF} and \mathbf{d}_{RI} , it is clearly seen that the main differences are located in the interior of the left 'subcortical structure' which is a pure rotation. In order to give a visual interpretation of the three features Figure 6.3.7 illustrate the different analysed features at two locations of the synthetic anatomy:

- Histograms shows the distribution of determinant features on both groups. In red color for {NOR} and blue color for {PAT} group; It can be seen that in both locations distributions are similar and no significative group differences is appreciated.
- The Cauchy–Green deformation tensor features are illustrate as ellipsoids glyphs. Gray glyphs are shown for {NOR} group and their intrinsic mean tensor is shown in red. Intrinsic mean tensor for {PAT} is shown in blue color. While both means look visually similar in the , at least with respect to the variance shown by the gray glyphs, at the 'left subcortical structure', differences can be appreciated from the 'right subcortical structure'.
- The full Jacobian matrices were illustrated as the linear deforming a small square. In order to represent also the rotation, a horizontal line was added in the glyphs. Again, all instances from {NOR} are shown in gray and their intrinsic mean in red color. Visual differences can be be appreciated between {NOR} mean Jacobian matrix and the mean from {PAT} (blue color). As was expected, only rotational differences are appreciated at the 'left subcortical structure'.

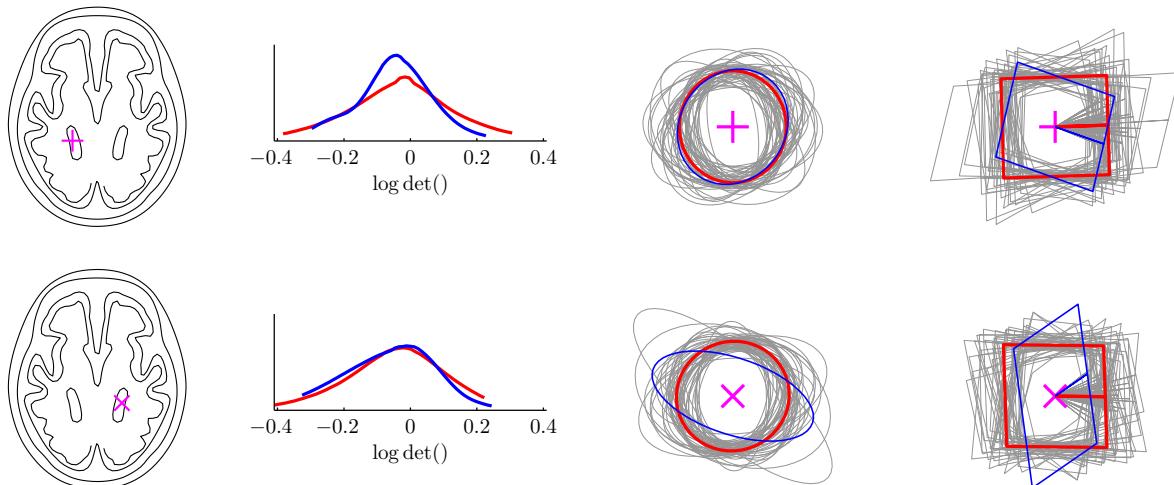


Figure 6.3.7. Glyph-based illustration of the three features used to perform the group difference test. Leftmost column shows the localization of the glyphs. From left to right: determinant histograms; elliptical glyph illustrating Cauchy–Green tensor; illustration of the Jacobian matrix (first order description of the underlying deformation). Intrinsic means are shown in red color for {NOR} group and in blue color for {PAT} group.

It can be concluded that an increasing sensitivity of the statistical test is obtained with a more complete descriptions of the local deformation.

6.3.4 Results on ADNI dataset

Twenty elderly control subjects (denoted here as {NOR} group) and twenty AD patients (denoted as {AD} group) were selected from the ADNI database. As brain atrophy is affected by factors such as age and gender, subjects were selected such that they were gender-matched and using a very narrow age interval. The demographic and clinical variables of the selected subjects are summarized in Table 6.3 where it can be seen both groups are paired in gender and age. Clinical scores, such as MMSE or the memory score in the CDR were significantly different between both groups under a classical Student's t -test.

SVF registration (Section 3.6) between the template (Section 6.2.1) and image subjects was used providing Jacobian matrices from the deformation mappings (Section 6.2.2). Afterwards,

Table 6.3. Descriptive statistics of demographic and clinical information. Format: $\text{average} \pm \text{standard deviation}$ [min, max].

Group	Gender (M/F)	Age	MMSE	memory CDR
{NOR}	10/10	72.2 ± 0.3 [70,73]	29.3 ± 1.0 [26,30]	0.0 ± 0.0 [0,0]
{AD}	10/10	72.2 ± 0.5 [70,73]	23.3 ± 1.8 [20,26]	0.9 ± 0.3 [0.5,2.0]

voxel-wise Cramér tests were performed on Jacobian matrices using the three right-invariant distances \mathbf{d}_{DET} , \mathbf{d}_{AFF} and \mathbf{d}_{RI} . Assessing of p -values were estimated by means of permutation test using 100 000 random permutations of group labels and p -value maps were corrected for multiple comparisons with the FDR criterion. Figure 6.3.8 shows illustrative coronal and sagittal slices of the corrected significance maps for each distance function.

Comparing with the result obtained in Section 6.2.4, it must be noted that the current experiment was performed on very small samples (20 subjects per group) and the statistical power is limited by this condition. Larger samples would require to covariate with fixed effects such as age and the statistical problem would become harder than the group comparison considered here.

The three distance-based tests provide statistical maps which are meaningful according to the pathophysiological knowledge of the disease. As was expected, the number of voxels with significant group differences in the \mathbf{d}_{DET} map was much smaller than in \mathbf{d}_{AFF} and \mathbf{d}_{RI} maps. The differences between \mathbf{d}_{DET} and \mathbf{d}_{AFF} – \mathbf{d}_{RI} maps might be due to the higher statistical power achieved when considering a multivariate description of the Jacobian matrix information rather than the simple scalar determinant as feature. Regarding the visual comparison between \mathbf{d}_{AFF} and \mathbf{d}_{RI} maps, they look very similar. This suggests that pure rotations were not relevant on this study.

Regarding computational issues, \mathbf{d}_{RI} is more expensive than \mathbf{d}_{DET} and \mathbf{d}_{AFF} . In our implementation, that uses the Gauss–Newton strategy for the optimization and a relative error of 10^{-10} as stopping criterion, the average computation time for \mathbf{d}_{RI} was about 50 msec. On the other hand, average computation times for \mathbf{d}_{DET} and \mathbf{d}_{AFF} were 0.02 msec and 1 msec, respectively.

6.4 Discussion

This chapter, devoted to brain morphometry and its changes occurring in the neurodegenerative process of the Alzheimer’s disease, has been the main driving application of this thesis. In a nutshell, the chapter is focused on several extensions to the classical ROI-based volumetry, which is so far the most widely used technique in clinical practice and in many medical research studies.

The first study was devoted to provide a more complete descriptor of the geometrical properties of subcortical nuclei than their volumes. The proposed descriptor characterizes the geometrical aspects of each subcortical structure by means of its pose using a small number of parameters: 3 for location, 3 for attitude, and 1 parameter for scale. This description provides a compact characterization of the anatomy that is intuitive and may be useful for several applications. Even though it describes only pose information without any detail about shape it might be good enough to describe pathology-induced changes in the anatomy. Additionally, pose information is complementary to classical high-dimensionality shape analysis studies where pose characterization is completely disregarded during the alignment stage. Pose information was described by the spatial transformation from $\mathcal{C} = \mathcal{S}^+(1) \times \mathcal{SO}(3) \times \mathcal{T}(3)$ group of centered transformations which

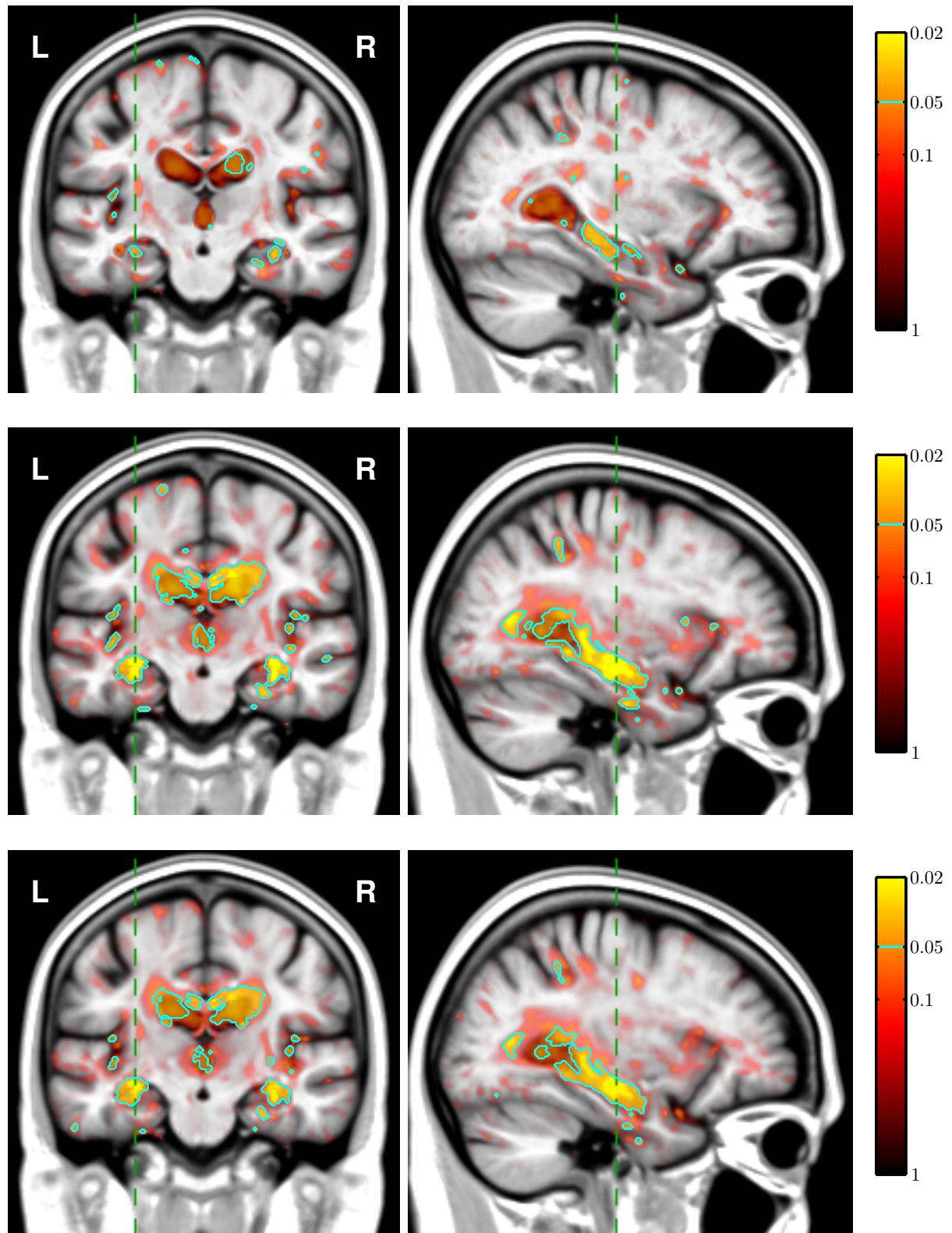


Figure 6.3.8. Illustrative coronal (left) and sagittal (right) views of FDR corrected significance maps of the Cramér test performed using three right-invariant distances: d_{DET} (top row), d_{AFF} (middle row) and d_{RI} (bottom row).

aligns a prototypical mean shape of each subcortical structure with the instance under study. A regression study on pose information versus cognitive scores provides a wealth of information about the course of anatomical changes correlated with the cognitive status given by the ADAS-cog score. On the other hand, a regression study of pose versus age on healthy and elderly subjects was also performed. Interesting, it was found that the pattern of changes related to cognitive decline are different from the ones related with the normal aging process.

The second experiment in this chapter was devoted to tensor-based morphometry studies. TBM could be understood as a volumetry study where no *a-priori* structures or ROIs are selected and the whole brain is analyzed. Using advanced tools, such as the estimation of an unbiased template and the use of SVF registration, local compression/expansion maps are computed from the determinant of Jacobian matrices of the deformation mappings. Subsequently, at each location (or at each voxel) hypothesis testing is performed in order to determine the brain regions with significant differences between a control group and a group of AD patients. As the classical volumetry is relaxed from a ROI level to the voxel level, TBM provides findings of (local) volume changes at a higher spatial and anatomical resolution.

The third study in this chapter extended the TBM methods to a multivariate characterization of the local deformation. This extension provides a more complete description of local changes in the geometry, in a similar way when ROI-based volumetry was extended with the complementary information about the pose. While standard TBM studies perform univariate hypothesis testing at each voxel on a volume change descriptor, multivariate statistical tools must be used when analyzing local deformations by means of the whole Jacobian matrix or the Green–Cauchy tensor. The proposed multivariate TBM takes the best of both previous extensions, a more complete feature than the scalar volume change and the removal of the constraints about analyze predefined ROIs. Moreover, the presented methodology takes into account the group structure and the invariance induced from the geometrical description suggested in the computational anatomy framework.

The following table summarizes the methodological tools considered throughout the chapter.

	ROI level	Voxel level
	A small number of predefined regions.	Huge number of test. Multiple comparisons must be considered.
Scalar descriptor		
Coarse and unspecific information. Intuitive and simple to analyze.	Volumetry	TBM on $\det(J)$
Multivariate descriptor	Pose analysis	Multivariate TBM on C and J
Hard to visualize the effects.		

CHAPTER

7

Conclusion and future extensions

Spatial transformations and the tools for their analysis have been mainly studied along this thesis. New methodologies have been proposed and applied to characterize brain anatomical changes induced by Alzheimer's disease. Traditionally, these analyses have been performed by directly measuring magnitudes on the anatomical structures. However, these magnitudes are, in general, unspecific and do not provide a complete description of the anatomical structures and the relationships among themselves. This thesis lays in the field of Computational anatomy in which the anatomy of each individual is geometrically described as the deformation of a given anatomical atlas. Using this description, the anatomical information remains encoded in the spatial transformation, mapping individuals to the anatomical atlas. The main advantage of this indirect description is that spatial transformations are mathematical entities with well-defined properties. However, spatial transformations generally do not have a vector space structure, and, therefore, classical statistical tools cannot directly be used. In this thesis, statistical tools from a more general framework defined for elements belonging to a metric space have been considered to perform statistical analysis on spatial transformations.

In this thesis, we have studied how to endow spatial transformation sets with a metric space structure allowing subsequent analysis and quantification. First, several interesting families of spatial transformations for medical imaging applications have been characterised. Most of these families can be considered as group structures and preserve some magnitudes from geometrical objects. Transformation group is the most prevalent algebraic structure considered in this thesis. Differential geometry and Lie group concepts were overviewed as they are key ingredients to endow transformation groups with a metric structure. The main focus has been bridging the gap between these theoretical concepts and more applied disciplines such as, computer vision, computer graphics and medical imaging. Special emphasis has been placed on illustrating these concepts with geometric transformations. In addition, there is a recent and growing interest in the analysis of manifold valued data, so the present review may be of great value for several scientific disciplines.

One of the key concepts in the thesis is the concept of geodesics. Riemannian geodesics depend on a Riemannian metric structure defined on the differentiable manifold. In order to perform geometrical analysis on a Lie group, invariant Riemannian metrics have been considered. These metrics remain invariant under the action of group elements, and, define symmetries which are relevant in several applications. Several properties and algorithms to compute invariant geodesics have been detailed and proposed in Chapters 4 and 5. Moreover, some applications for the analysis of spatial transformations have been proposed in order to illustrate the use of geodesics and invariant distances.

These "advanced" analysis tools together with the more general statistical framework for

metric spaces have been applied to several studies in the medical imaging area. In particular, the anatomical changes occurring along the Alzheimer's disease have been characterised in this thesis. To do so, several methodologies have been developed to improve the classical methodology of volumetry of anatomical structures.

The first proposal has been to consider a pose descriptor of subcortical nuclei. Complementary to the volume, pose provides information about the localization, attitude and size of the structures within the brain. Riemannian distances have been used to analyse pose information versus ADAS-cog cognitive scale by a geodesic regression. Additionally, geodesic regression study has been performed with respect to a normal aging process. Interestingly, it was found that the pattern of pose changes of subcortical nuclei induced by cognitive decline was different to the changes induced by normal aging.

On the other hand, ROI volumetry studies have been extended by analyzing the complete brain and computing volume changes at a voxel-level with TBM technique. Advanced non-rigid registration techniques based on diffeomorphisms with stationary velocity fields have been considered in this thesis. A methodology was developed to build unbiased anatomical templates with a high spatial resolution. Subsequent statistical maps obtained with TBM also identified anatomical differences between patient and control groups with a high spatial resolution. These findings were in agreement with the pathophysiological knowledge of Alzheimer's disease.

Finally, a new methodology for local deformation analysis has been proposed in this thesis. While classical TBM studies focus on local volume changes by means of the Jacobian determinant, this thesis proposes the use of the full Jacobian matrix employing analysis tools on Lie groups. The full Jacobian matrix was properly characterised as an element of $\mathbf{GL}^+(d)$ group. It has also been analyzed the issue of template invariance in TBM studies. It has been shown that a sufficient condition to ensure this fundamental requirement is using a right-invariant distance when characterizing Jacobian matrices and subsequent statistical methodologies. Voxel-wise hypothesis testing has been performed using Cramér test in two scenarios: a simulated synthetic study and a brain study for Alzheimer's disease. The main result is that the larger amount of information is considered in the analysis the higher sensitivity to detect anatomical differences. To our knowledge, it is the first time that the full Jacobian matrix is used in a template invariance TBM framework.

Future work

Along this thesis several methodologies have been proposed to analyze geometrical information. Several illustrative studies have also shown its application and performance on brain morphometry studies. Theoretical and applied aspects can be useful in several areas: to carry out new morphometric studies of the brain or others organs; to study different manifold-valued data, such as DTI; to analyze data from other technical areas, such as computer vision; and even in another areas of knowledge, such as geometric mechanics and hydrodynamics.

Just to name a few interesting extensions of the proposed methodologies and tools:

- **Geodesic regression including covariates:** in contrast to standard linear regression on a vector space using least squares where covariates can be easily considered in the analysis, the order of the effects is important when analyzing non-commutative elements, such as pose information. A detailed analysis of how to choose the most appropriate ordering of the effects should be analyzed and discussed.
- **Multivariate TBM on $\mathbf{GA}^+(d)$ group:** when comparing the results of the anatomical changes induced by Alzheimer's disease at caudate nuclei illustrated with pose information

and voxel-wise multivariate TBM, it can be concluded that significant displacements without volume change were detected by pose analysis, while TBM did not manage to find any statistical difference at caudate nuclei. This happens because only the first-order of the deformation gradient (Jacobian matrix) was used in the analysis, rejecting the information from the zero-order approximation (local displacement). It can be argued that this term should be also considered in TBM studies, such it is already done in DBM. Therefore, a more complete deformation descriptor could be analyzed considering elements from $\mathbf{GA}(d)$ group, *i.e.* considering both displacements and the full Jacobian matrix. One practical limitation of this proposal is that the computational cost would be higher than for multivariate TBM on $\mathbf{GL}^+(d)$ because it is not known a closed-form solution for the Riemannian exponential function. Again, the template invariance issue could be also solved by means of a right-invariant Riemannian metric.

Choice of the metric tensor: it has been shown that the choice of the metric tensor has a strong effect on the metric structure given to spatial transformations as well as on the resulting geodesics. In the practical examples given in this thesis, the metric tensor was selected in order to get the lowest computational cost. However, alternative choices of the metric might yield better results on a given application. In the future, it would be interesting to study methods and criteria for selecting an appropriate metric tensor for each application.

Transport of covariance matrices: In this thesis multivariate TBM was performed using Cramér test because it can be easily computed from inter-element distances. If one were interested in using a multivariate statistic, such as Hotelling's T^2 , on a differentiable manifold, a methodology to transport covariance matrices would be required in order to describe the covariance matrix of each sample on a common tangent space. While the transport of covariance matrices in a vector space is trivial, the manifold case is more complex because it depends on several properties of the particular space, such as the selected affine connection and the selected path to perform the transport.

Regularity measure of deformation fields: it is common practice to include a regularization term in non-rigid registration to favour some specific deformations. Once the space of spatial transformations is endowed with a metric structure, spatial derivatives (gradients) of Jacobian matrices can be computed, for example considering a derivative filter. A measure of the regularity of the deformation field can be obtained by integrating the norm of the gradient. Using such a regularity measure for the case of $\mathbf{GL}^+(d)$ group, affine transformations would not be penalized by the regularization term, which may be of interest when aligning objects.

Different criteria for left- and right- invariance: in this thesis invariant metrics were defined under the assumption that the invariance was respect to operations with elements from the same set which is under study. Under this assumption, many Lie groups do not have a bi-invariant metric. However, this assumption may be relaxed, for example considering Riemannian metrics on $\mathbf{GL}^+(d)$ which would be simultaneously invariant under a left-action with a $\mathbf{GL}^+(d)$ element and right-invariant with the action of an element from $\mathbf{SO}(d)$. This invariant criteria could be useful, for example, in applications where left-invariant is required to be independent of the measurement system, and right-invariant to be independent of a template choice.

Injectivity radius: Non-uniqueness of the Riemannian logarithm has been discussed in this thesis. This issue may be relevant when computing distances between spatial transformations. Further interesting research work could be done to compute the *injectivity radius*. If the length of the geodesic segment connecting two elements is shorter than the

injectivity radius, then, it could be ensured that this length corresponds to the Riemannian distance. In addition, this would allow to have a quantitative measure for vague statements, such as “instances are sufficiently concentrated to have a single mean”.

Conclusión y extensiones futuras

A lo largo de esta tesis se han estudiado las transformaciones espaciales con especial interés en las herramientas de análisis. Se han propuesto nuevas metodologías que se han aplicado a la cuantificación de los cambios anatómicos en el cerebro inducidos por la enfermedad de Alzheimer. Tradicionalmente estos análisis se llevan a cabo midiendo magnitudes medidas directamente sobre las estructuras anatómicas. Sin embargo, estas magnitudes son generalmente inespecíficas y no proporcionan una descripción completa de las estructuras anatómicas y las relaciones entre ellas. En esta tesis se ha utilizado la disciplina de Anatomía Computacional en la que las anatomías se describen de una manera geométrica y cada individuo es una versión deformada de un atlas anatómico. Con esta descripción indirecta de las estructuras anatómicas, el análisis se realiza sobre las transformaciones que relacionan a cada individuo con el atlas anatómico. Sin embargo, las transformaciones espaciales no tienen generalmente una estructura de espacio vectorial y por lo tanto las herramientas clásicas de análisis estadístico no se pueden aplicar directamente. Para realizar análisis estadísticos sobre las transformaciones espaciales, en este trabajo se han considerado herramientas estadísticas en un marco más general definido para elementos que pertenezcan a un espacio métrico.

En esta tesis se ha estudiado de forma rigurosa cómo asignar a las transformaciones espaciales de una estructura de espacio métrico que nos permita su análisis y cuantificación. Se ha empezado por la caracterización de distintas familias de transformaciones espaciales que son de interés en imagen médica. La mayoría de estas familias preservan algunas magnitudes de los objetos geométricos y se les puede asignar una estructura de grupo de transformaciones. Para dotar a los grupos de transformaciones de una estructura métrica se han revisitado conceptos de geometría diferencial y grupos de Lie. El principal objetivo de esta revisión ha sido trasladar dichos conceptos a la comunidad de visión por ordenador, gráficos por ordenador y análisis de imágenes médicas. Se ha puesto un especial interés en ilustrar estos conceptos con transformaciones geométricas. Además, en los últimos años ha crecido enormemente el interés en el análisis de datos provenientes de diversas variedades diferenciales y la revisión presentada es de mucha utilidad para diversas disciplinas científicas.

Una de los conceptos clave en esta tesis es el de geodésicas. Las geodésicas riemannianas dependen de una estructura de métrica riemanniana asignada a la variedad diferencial. Para realizar los análisis geométricos sobre un grupo de Lie se han considerado aquellas métricas riemannianas que se mantienen invariantes ante la acción de cualquier elemento de grupo y definen simetrías que son heredadas por las aplicaciones. Diversas propiedades y algoritmos para calcular estas geodésicas invariantes se han detallado y propuesto en los Capítulos 4 y 5. Además, el uso de las geodésicas y distancias invariantes se ha ejemplificado en algunas aplicaciones de análisis de transformaciones espaciales.

Estas modernas herramientas de análisis, junto con el marco estadístico general definido para espacios con estructura métrica, se han empleado para distintos estudios sobre imágenes médicas. En esta tesis, se han utilizado para la determinación de patrones de cambios anatómicos

producidos durante la enfermedad de Alzheimer. Para hacer esto, se desarrollaron metodologías de análisis que extienden los estudios volumétricos clásicos sobre estructuras anatómicas.

La primera extensión propuesta es la de considerar un descriptor de la pose de las estructuras subcorticales. Como complemento al volumen, la pose provee información sobre la localización, orientación y tamaño de las estructuras dentro del cerebro. Ademas, este descriptor pertenece a un grupo de Lie. Se ha propuesto el uso de distancias riemannianas para el análisis de la pose con respecto a la escala cognitiva ADAS-cog mediante un ajuste de regresión con geodésicas. También se ha realizado un estudio por regresión geodésica de la pose con respecto a la edad en un estudio de envejecimiento normal. Se ha concluido que ambos procesos, el deterioro cognitivo y el envejecimiento normal, generan cambios anatómicos cualitativamente diferentes a nivel subcortical.

Por otro lado los estudios volumétricos en regiones de interés se pueden generalizar al evaluar cambios locales de volumen en todo el cerebro con la técnica TBM. Por medio de herramientas avanzadas de corregistro de imágenes se ha generado un atlas anatómico de alta resolución para usar en el estudio de TBM. Con estas herramientas se han obtenido mapas estadísticos con una alta resolución espacial y que concuerdan con el conocimiento fisiopatológico disponible de la enfermedad de Alzheimer.

Por último, se ha propuesto una nueva metodología para el análisis local de los campos de deformaciones. Mientras que los estudios clásicos de TBM analizan sólo el cambio de volumen local por medio del determinante de la matriz jacobiana, se propuso el estudio de la matriz jacobiana completa utilizando las herramientas de análisis sobre grupos de Lie. Para ello se ha caracterizado correctamente la matriz jacobiana como una transformación espacial local perteneciente al grupo $\mathbf{GL}^+(d)$. Para conseguir que la elección del atlas no afecte a los resultados estadísticos, en TBM, se ha demostrado que es una condición suficiente el utilizar una distancia invariante a derecha. Utilizando como herramienta estadística el test de Cramér, se llevaron a cabo un estudio sintético ilustrativo y un estudio de TBM multivariado entre sujetos de control y pacientes de Alzheimer. Se concluye que se obtiene mayor sensibilidad cuanta más información de la matriz jacobiana se utiliza en el análisis. Para nuestro conocimiento es la primera vez que se utiliza la información completa de la matriz jacobiana junto con el requerimiento de invariancia ante la elección del atlas anatómico en un estudio de TBM.

Extensiones futuras

Los estudios y las metodologías de análisis propuestos en esta tesis han facilitado el desarrollo de herramientas para el análisis de entidades con una naturaleza geométrica. Las aportaciones metodológicas y teóricas que proporciona esta tesis en el análisis de los grupos de Lie pueden ser utilizadas en diversas áreas: nuevos estudios morfométricos, del cerebro u otros órganos; estudios en otras variedades típicamente utilizadas en imagen médica como DTI; estudios en otras áreas técnicas, como visión por ordenador; e incluso otras áreas del conocimiento, como mecánica geométrica e hidrodinámica.

Posibles usos y complementos a las herramientas propuestas que sería interesante de investigar en el futuro son:

- **Extensión de la regresión geodésica para incluir covariables:** en el caso de una regresión lineal sobre espacios vectoriales por mínimos cuadrados, las covariables se incluyen fácilmente como términos aditivos. Sin embargo, para incluirlas en una regresión geodésica hay que especificar el orden de los efectos.

TBM multivariado sobre $\mathbf{GA}^+(d)$: comparando los resultados del estudio de regresión geodésica con respecto a ADAS-cog y el resultado de TBM multivariado, podemos ver

que, mientras que en la regresión el desplazamiento de los caudados resulta notable y significativo, en el estudio de TBM multivariado no se detecta ninguna variación geométrica relevante. Esto sucede porque en la aproximación a primer orden de la deformación sólo se utilizó el término de orden 1 y se ignoró el término de orden 0. Este término se puede incluir en el análisis, y considerar como descriptor local de la deformación a un elemento del grupo $\mathbf{GA}(d)$, es decir, considerando el desplazamiento así como la matriz jacobiana. La invariancia ante el atlas anatómico se obtendría otra vez con una métrica invariante a derecha. Sin embargo, para este grupo no se conoce expresión cerrada ni siquiera para la exponencial riemanniana, requiriendo un mayor coste computacional.

Elección del tensor métrico: si bien las metodologías de análisis propuestas están fundadas en conceptos matemáticos rigurosos, hemos visto que la elección de la métrica influye drásticamente en las geodésicas y en la estructura métrica que heredan las transformaciones espaciales. En las aplicaciones prácticas, generalmente las métricas se han elegido con el objetivo de simplificar el proceso de cálculo. Sin embargo, otras elecciones del operador de inercia resultarían en análisis igualmente rigurosos y válidos. Es interesante estudiar cómo estas elecciones afectan a los resultados y definir criterios para su elección.

Transporte de matrices de covarianza: en el análisis de TBM multivariado se ha utilizado el test de Cramér. Sin embargo, es práctica común en estudios sobre espacios vectoriales utilizar el test multivariado de Hotelling T^2 . Su extensión a variedades diferenciales requiere de una metodología para trasladar las covarianzas de ambos grupos a un espacio tangente común. Mientras que en el caso vectorial el transporte de las covarianzas es trivial, en el caso de variedades es más complejo y depende de muchas propiedades del espacio en estudio, como por ejemplo de la conexión afín que se elija y del camino que se tome para realizar el transporte.

Medida de regularidad de campos de deformaciones: en el proceso de registro se incluye generalmente un término que regularice la deformación. La asignación de estructura métrica al espacio de las transformaciones permite calcular derivadas de matrices jacobianas. Dada una deformación diferenciable de espacio ambiente, utilizando una distancia sobre sus matrices jacobianas se puede calcular su gradiente espacial. Integrando la norma de este gradiente en el espacio se obtiene una medida de regularidad de la deformación. Con esta función de regularidad, las transformaciones afines no serían penalizadas, y no penalizarían el proceso de corregistro.

Diferentes invariancias a cada lado: en el escenario que se ha estudiado, las métricas invariantes se construyen de forma que sean invariantes al mismo grupo que se está estudiando. En este marco, muchos grupos de Lie no pueden ser dotados de una métrica bi-invariante. Sin embargo, estas condiciones se pueden relajar y considerar, por ejemplo, métricas riemannianas sobre $\mathbf{GL}^+(d)$ que sean al mismo tiempo invariante a izquierda ante la acción de $\mathbf{GL}^+(d)$ e invariante a derecha ante la acción de $\mathbf{SO}(d)$. Esto sería de gran utilidad, por ejemplo, en un escenario donde se deseé simultáneamente invariancia a izquierda para ser independiente al sistema de medición e invariancia a derecha para ser independiente a la elección del atlas anatómico.

Radio de inyectividad: se ha discutido sobre la no unicidad de las soluciones de la función logaritmo riemanniano. Esto es un inconveniente grave a la hora de calcular distancias entre elementos. Una línea de investigación interesante sería calcular el llamado *radio de inyectividad*. Si la longitud de una geodésica entre dos elementos es menor que el radio de inyectividad, entonces se garantizaría que es la distancia riemanniana. Además nos permitiría definir rigurosamente, qué significa que las instancias estén “lo suficientemente concentradas alrededor de la media”.

Publications derived from the thesis

Publications in Journal Articles

- E. Zacur, M. Bossa, and S. Olmos. “Left-invariant Riemannian geodesics on spatial transformation groups”. *SIAM Journal on Imaging Sciences*, In press, 2014
- E. Zacur, M. Bossa, and S. Olmos. “Multivariate tensor-based morphometry with a right-invariant Riemannian distance on $GL+(n)$ ”. *Journal of Mathematical Imaging and Vision*, In press, doi:10.1007/s10851-013-0479-7, 2013.
- M. A. Fullana, N. Cardoner, P. Alonso, M. Subirà, C. López-Solà, J. Pujol, C. Segalàs, E. Real, M. Bossa, E. Zacur, I. Martínez-Zalacaín, A. Bulbena, J.M. Menchón, S. Olmos, and C. Soriano-Mas. “Brain regions related to fear extinction in obsessive-compulsive disorder and its relation to exposure therapy outcome: a morphometric study”. *Psychological Medicine*, 44:845–856, 3 2014.
- M. Bossa, E. Zacur, and S. Olmos. “Statistical analysis of relative pose information of subcortical nuclei: application on ADNI data”. *NeuroImage*, 55(3):999–1008, Apr 2011.
- M. Bossa, E. Zacur, and S. Olmos. “Tensor-based morphometry with stationary velocity field diffeomorphic registration: application to ADNI”. *NeuroImage*, 51(3):956–969, Jul 2010.

Publications in Proceedings of international conferences

- E. Zacur, M. Bossa, and S. Olmos. “A right-invariant Riemannian distance on $GL+(n)$ and hypothesis testing on Jacobian matrices”. *Mathematical Foundations of Computational Anatomy (MFCA), MICCAI*, Nagoya (Japan), 2013.
- M. Bossa, E. Zacur, and S. Olmos. “On changing coordinate systems for longitudinal tensor-based morphometry”. In *Spatio Temporal Image Analysis Workshop (STIA), MICCAI*, Beijing (China), 2010.
- M. Bossa, E. Zacur, and S. Olmos. “Selection of registration parameters with diffeomorphic registration for TBM studies on ADNI data”. In *16th Annual Meeting of the Organization for Human Brain Mapping*, Barcelona (Spain), 2010.
- M. Bossa, E. Zacur, and S. Olmos. “On alignment techniques for statistical shape analysis of brain structures using landmark coordinate”. In *16th Annual Meeting of the Organization for Human Brain Mapping*, Barcelona (Spain), 2010.

- M. Bossa, E. Zacur, and S. Olmos. “Tensor-based morphometry with mappings parameterized by stationary velocity fields in Alzheimer’s Disease Neuroimaging Initiative”. In G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI*, pages 240–247, 2009.
- M. Bossa, E. Zacur, and S. Olmos. “Algorithms for computing the group exponential of diffeomorphisms: performance evaluation”. In *Workshops on Computer Vision and Pattern Recognition, IEEE*, pages 23–28, Jun 2008.

Other Publications

- A. Mincholé, E. Zacur, E. Pueyo, P. Laguna. “Modeling and quantification of repolarization feature dependency on heart rate”. *Methods of Information in Medicine*, In press, 2014.
- A. Mincholé, E. Pueyo, J. F. Rodríguez, E. Zacur, M. Doblaré, and P. Laguna. “Quantification of restitution dispersion from the dynamic changes of the T-wave peak to end, measured at the surface ECG”. *IEEE Transactions on Biomedical Engineering*, 58(5):1172–1182, May 2011.
- E. Meinhardt, E. Zacur, A. Frangi, and V. Caselles. “3D edge detection by selection of level surface patches”. *Journal of Mathematical Imaging and Vision*, 34(1):1–16, 2009.
- V.C. Orvalho, E. Zacur and A. Susin. “Transferring the rig and animations from a character to different face models”. *Computer Graphics Forum*, 27(8):1997–2012, 2008.
- G. Meyer, P. Arneodo Larochette, A. Baruj, F.J. Castro, P. Lacharmoise, E. Zacur, and B.A. Talagañis. “Equipment for hydrogen absorption-desorption cycling characterization of hydride forming materials”. *Review of Scientific Instruments*, 78(2):023903–, 2007.
- A. Mincholé, E. Zacur, E. Pueyo, P. Laguna. “Robust characterization of repolarization adaptation to heart rate changes”. *International Workshop on Biosignal Interpretation (BSI)*, Como (Italy), 2012.
- A. Mincholé, E. Pueyo, J.F. Rodríguez, E. Zacur, M. Doblaré, P. Laguna. “Evaluation of a method for quantification of restitution dispersion from the surface ECG”. *Procedding of Computers in Cardiology (CinC)*, Belfast (UK), 2010.
- V.C. Orvalho, E. Zacur and A. Susin. “A MAYA plugin to transfer facial expressions to different face models”. *International Digital Games Conference (iDiG)*, Portoalegre (Portugal), 2006.
- V.C. Orvalho, E. Zacur and A. Susin. “Transferring a labeled rig to animate face models”. *IV Conference on Articulated Motion and Deformable Objects (AMDO)*, Mallorca (Spain), 2006.
- V.C. Orvalho, E. Zacur and A. Susin. “Transferring facial expressions to different face models”. *3rd Ibero-American Symposium in Computer Graphics (SIACG)*, Santiago de Compostela (Spain), 2006.

Bibliography

- Abadir, K. M. and Magnus, J. R. (2005). *Matrix algebra*. Cambridge University Press. (Cited on pages 94, 96, 123, and 175.)
- Abraham, R. H. and Marsden, J. E. (1978). *Foundations of mechanics*. AMS Bookstore. (Cited on page 106.)
- Absil, P.-A., Baker, C. G., and Gallivan, K. A. (2007). Trust-region methods on Riemannian manifolds. *Foundations of Computational Mathematics*, 7(3):303–330. (Cited on page 77.)
- Absil, P.-A., Mahony, R., and Sepulchre, R. (2009). *Optimization algorithms on matrix manifolds*. Princeton University Press. (Cited on pages 77, 78, 79, and 80.)
- Absil, P.-A., Mahony, R., and Sepulchre, R. (2010). Optimization on manifolds: Methods and applications. In Diehl, M., Glineur, F., Jarlebring, E., and Michalek, W., editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 125–144. Springer Berlin Heidelberg. (Cited on page 80.)
- Adler, R. L., Dedieu, J. P., Margulies, J. Y., Martens, M., and Shub, M. (2002). Newton’s method on Riemannian manifolds and a geometric model for the human spine. *IMA Journal of Numerical Analysis*, 22(3):359–390. (Cited on page 77.)
- Akkouche, S. and Galin, E. (2001). Adaptive implicit surface polygonization using marching triangles. *Computer Graphics Forum*, 20(2):67–80. (Cited on page 27.)
- Al-Mohy, A. H. (2010). *Algorithms for the matrix exponential and its Fréchet derivative*. PhD thesis, University of Manchester. (Cited on pages 82, 96, 114, and 135.)
- Al-Mohy, A. H. and Higham, N. J. (2009). Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation. *SIAM Journal on Matrix Analysis and Applications*, 30(4):1639–1657. (Cited on page 96.)
- Aldous, D. J. (1985). Exchangeability and related topics. In Hennequin, P. L., editor, *École d’Été de Probabilités de Saint-Flour XIII–1983*, volume 1117 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg. (Cited on page 158.)
- Alexa, M. (2002). Linear combination of transformations. *ACM Transactions on Graphics*, 21(3):380–387. (Cited on pages 24, 76, 83, 133, and 135.)
- Anderson, M. J. and Robinson, J. (2001). Permutation tests for linear models. *Australian & New Zealand Journal of Statistics*, 43(1):75–88. (Cited on page 167.)
- Andruchow, E., Larotonda, G., Recht, L., and Varela, A. (2011). The left invariant metric in the general linear group. *arXiv preprint*. (Cited on pages 102 and 117.)

- Ang, W. T. and Park, Y. S. (2008). *Ordinary differential equations: methods and applications*. Universal Publishers. (Cited on page 122.)
- Angelidis, A. and Singh, K. (2006). Space deformations and their application to shape modeling. In *ACM SIGGRAPH 2006 Courses*, SIGGRAPH '06, pages 10–29. ACM. (Cited on page 24.)
- Appledorn, C. R. (1996). A new approach to the interpolation of sampled data. *IEEE Transactions on Medical Imaging*, 15(3):369–376. (Cited on page 29.)
- Arnold, D. N. and Rogness, J. (2008). Möbius transformations revealed. *Notices of the AMS*, 55:1226–1231. (Cited on page 119.)
- Arnold, V. I. (1989). *Mathematical methods of classical mechanics*. Springer. (Cited on pages 77, 80, 83, 86, 102, and 106.)
- Arnold, V. I. (1992). *Ordinary differential equations*. Springer Textbook Series. New York. (Cited on pages 9 and 35.)
- Arnold, V. I. and Khesin, B. A. (1998). *Topological methods in hydrodynamics*. Springer. (Cited on pages 77, 80, and 91.)
- Arsigny, V. (2006). *Processing data in Lie groups: an algebraic approach. Application to non-linear registration and diffusion tensor MRI*. PhD thesis, École Polytechnique. (Cited on pages 46, 76, 80, 88, 115, 116, 143, and 161.)
- Arsigny, V., Commowick, O., Pennec, X., and Ayache, N. (2006). Statistics on diffeomorphisms in a log-Euclidean framework. In Pennec, X. and Joshi, S., editors, *MICCAI Workshop on Mathematical Foundations of Computational Anatomy*. (Cited on page 45.)
- Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2005). Fast and simple calculus on tensors in the log-Euclidean framework. In Duncan, J. S. and Gerig, G., editors, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, volume 3749 of *Lecture Notes in Computer Science*, pages 115–122. Springer Berlin Heidelberg. (Cited on page 174.)
- Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2007). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347. (Cited on pages 11, 136, 170, and 174.)
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113. (Cited on page 55.)
- Ashburner, J. and Friston, K. J. (2000). Voxel-based morphometry—the methods. *NeuroImage*, 11(6):805–821. (Cited on page 18.)
- Auzias, G., Colliot, O., Glaunès, J. A., Perrot, M., Mangin, J.-F., Trouvé, A., and Baillet, S. (2011). Diffeomorphic brain registration under exhaustive sulcal constraints. *IEEE Transactions on Medical Imaging*, 30(6):1214–1227. (Cited on pages 28 and 55.)
- Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41. (Cited on pages 53 and 55.)
- Avants, B. B., Schoenemann, P. T., and Gee, J. C. (2006). Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Medical Image Analysis*, 10(3):397–412. (Cited on page 55.)

- Baerentzen, J. A. and Aanaes, H. (2005). Signed distance computation using the angle weighted pseudonormal. *IEEE Transactions on Visualization and Computer Graphics*, 11(3):243–253. (Cited on page 27.)
- Bakircioglu, M. M., Joshi, S. C., and Miller, M. I. (1999). Landmark matching on brain surfaces via large deformation diffeomorphisms on the sphere. In *Proceedings of SPIE*, volume 3661, pages 710–715. (Cited on page 24.)
- Balasubramanian, M., Polimeni, J. R., and Schwartz, E. L. (2009). Exact geodesics and shortest paths on polyhedral surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):1006–1016. (Cited on page 77.)
- Bankman, I. (2008). *Handbook of medical image processing and analysis*. Academic Press series in biomedical engineering. Elsevier Science. (Cited on page 42.)
- Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *Journal of Multivariate Analysis*, 88(1):190–206. (Cited on page 177.)
- Barr, A. H. (1984). Global and local deformations of solid primitives. *SIGGRAPH Comput. Graph.*, 18(3):21–30. (Cited on page 24.)
- Barrio, R., Blesa, F., and Lara, M. (2005). VSVO formulation of the Taylor method for the numerical solution of ODEs. *Computers & Mathematics with Applications*, 50(1):93–111. (Cited on page 122.)
- Batchelor, P. G., Moakher, M., Atkinson, D., Calamante, F., and Connelly, A. (2005). A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53(1):221–225. (Cited on pages 170 and 173.)
- Bayro-Corrochano, E. and Ortegón-Aguilar, J. (2007). Lie algebra approach for tracking and 3D motion estimation using monocular vision. *Image and Vision Computing*, 25(6):907–921. (Cited on pages 77 and 78.)
- Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. (2005). Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157. (Cited on pages 44, 55, and 77.)
- Begelfor, E. and Werman, M. (2005). Interpolating transformations: Riemannian geometry, invariant metrics and geodesics. Unpublished, accessed from: <ftp://ftp.cs.huji.ac.il/users/aristo/papers/SYGRAPH2005/sig05.pdf>. (Cited on pages 76 and 133.)
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57:289–300. (Cited on page 168.)
- Bergner, S., Moller, T., Weiskopf, D., and Muraki, D. J. (2006). A spectral analysis of function composition and its implications for sampling in direct volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1353–1360. (Cited on page 46.)
- Berkels, B., Linkmann, G., and Rumpf, M. (2010). An $SL(2)$ invariant shape median. *Journal of Mathematical Imaging and Vision*, 37(2):85–97. (Cited on pages 24 and 53.)
- Bernasconi, N., Duchesne, S., Janke, A., Lerch, J., Collins, D. L., and Bernasconi, A. (2004). Whole-brain voxel-based statistical analysis of gray matter and white matter in temporal lobe epilepsy. *NeuroImage*, 23(2):717–723. (Cited on page 18.)

- Bhattacharya, A. and Bhattacharya, R. (2012). *Nonparametric inference on manifolds: with applications to shape spaces*. Cambridge University Press. (Cited on page 136.)
- Bhattacharya, R. and Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. *The Annals of Statistics*, 31(1):1–29. (Cited on pages 136 and 137.)
- Bhattacharya, R. and Patrangenaru, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds–ii. *The Annals of Statistics*, 33(3):1225–1259. (Cited on page 136.)
- Bloch, A., Krishnaprasad, P. S., Marsden, J. E., and Ratiu, T. S. (1996). The Euler–Poincaré equations and double bracket dissipation. *Communications in Mathematical Physics*, 175(1):1–42. (Cited on page 98.)
- Bloch, A. M., Crouch, P. E., Marsden, J. E., and Ratiu, T. S. (2002). The symmetric representation of the rigid body equations and their discretization. *Nonlinearity*, 15(4):1309. (Cited on pages 102, 105, and 107.)
- Bloch, A. M., Crouch, P. E., Marsden, J. E., and Sanyal, A. K. (2008). Optimal control and geodesics on quadratic matrix Lie groups. *Foundations of Computational Mathematics*, 8(4):469–500. (Cited on page 105.)
- Bloch, A. M., Crouch, P. E., Nordkvist, N., and Sanyal, A. K. (2011). Embedded geodesic problems and optimal control for matrix Lie groups. *Journal of Geometric Mechanics*, 3(2):197–223. (Cited on pages 101, 102, and 105.)
- Blu, T., Thévenaz, P., and Unser, M. (2004). Linear interpolation revitalized. *IEEE Transactions on Image Processing*, 13(5):710–719. (Cited on page 29.)
- Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., and Ayache, N. (2008). Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Transactions on Medical Imaging*, 27(4):557–568. (Cited on pages 76 and 77.)
- Bommes, D. and Kobbelt, L. (2007). Accurate computation of geodesic distance fields for polygonal curves on triangle meshes. In *Vision, Modelling, and Visualization*, volume 7, pages 151–160. (Cited on page 77.)
- Bonfiglioli, A. and Fulci, R. (2011). *Topics in noncommutative algebra: the Theorem of Campbell, Baker, Hausdorff and Dynkin*. Lecture Notes in Mathematics. Springer. (Cited on page 71.)
- Bonnans, J.-F., Gilbert, J. C., Lemaréchal, C., and Sagastizábal, C. A. (2006). *Numerical optimization: theoretical and practical aspects*. Springer. (Cited on page 31.)
- Bookstein, F. L. (1997). *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press. (Cited on pages 17 and 28.)
- Borodin, P., Zachmann, G., and Klein, R. (2004). Consistent normal orientation for polygonal meshes. In *Computer Graphics International 2004*, pages 18–25. IEEE Computer Society. (Cited on page 27.)
- Borwein, J. M. and Crandall, R. E. (2013). Closed forms: what they are and why we care. *Notices of the AMS*, 60:50–65. (Cited on page 114.)

- Bossa, M., Hernandez, M., and Olmos, S. (2007). Contributions to 3D diffeomorphic atlas estimation: Application to brain images. In Ayache, N., Ourselin, S., and Maeder, A., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, volume 4791 of *Lecture Notes in Computer Science*, pages 667–674. Springer Berlin Heidelberg. (Cited on page 71.)
- Bossa, M. and Olmos, S. (2006). Statistical model of similarity transformations: building a multi-object pose model of brain structures. In *Computer Vision and Pattern Recognition Workshops*, pages 59–66. (Cited on pages 78 and 147.)
- Bossa, M. and Olmos, S. (2007). Multi-object statistical pose+shape models. In *IEEE International Symposium on Biomedical Imaging*, pages 1204–1207. (Cited on page 147.)
- Bossa, M., Zacur, E., and Olmos, S. (2008). Algorithms for computing the group exponential of diffeomorphisms: performance evaluation. In *Computer Vision and Pattern Recognition Workshops*, pages 23–28. IEEE. (Cited on page 45.)
- Bossa, M., Zacur, E., and Olmos, S. (2009). Tensor-based morphometry with mappings parameterized by stationary velocity fields in Alzheimer’s Disease Neuroimaging Initiative. In Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., and Taylor, C., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, volume 5762 of *Lecture Notes in Computer Science*, pages 240–247. Springer Berlin Heidelberg. (Cited on page 18.)
- Bossa, M., Zacur, E., and Olmos, S. (2010a). On changing coordinate systems for longitudinal tensor-based morphometry. In *MICCAI Workshop Spatio Temporal Image Analysis Workshop (STIA)*, volume 2010. (Cited on page 146.)
- Bossa, M., Zacur, E., and Olmos, S. (2010b). Tensor-based morphometry with stationary velocity field diffeomorphic registration: application to ADNI. *NeuroImage*, 51(3):956–969. (Cited on pages 18, 55, 146, 164, and 168.)
- Bossa, M., Zacur, E., and Olmos, S. (2011). Statistical analysis of relative pose information of subcortical nuclei: application on ADNI data. *NeuroImage*, 55(3):999–1008. (Cited on pages 76, 77, 117, and 147.)
- Böttcher, A. and Wenzel, D. (2005). How big can the commutator of two matrices be and how big is it typically? *Linear Algebra and its Applications*, 403:216–228. (Cited on pages 124 and 126.)
- Böttcher, A. and Wenzel, D. (2008). The Frobenius norm and the commutator. *Linear Algebra and its Applications*, 429(8):1864–1885. (Cited on page 124.)
- Boucher, M., Evans, A., and Siddiqi, K. (2011). Anisotropic diffusion of tensor fields for fold shape analysis on surfaces. In Székely, G. and Hahn, H. K., editors, *Information Processing in Medical Imaging*, volume 6801 of *Lecture Notes in Computer Science*, pages 271–282. Springer Berlin Heidelberg. (Cited on pages 162, 170, and 174.)
- Boumal, N. (2013). Interpolation and regression of rotation matrices. In Nielsen, F. and Barbaresco, F., editors, *Geometric Science of Information*, volume 8085 of *Lecture Notes in Computer Science*, pages 345–352. Springer Berlin Heidelberg. (Cited on page 76.)
- Boykov, Y. and Kolmogorov, V. (2003). Computing geodesics and minimal surfaces via graph cuts. In *IEEE International Conference on Computer Vision*, volume 1, pages 26–33. IEEE. (Cited on page 77.)

- Braak, H. and Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, 82(4):239–259. (Cited on page 146.)
- Bradley, W. G. (2008). History of medical imaging. *Proceedings of the American Philosophical Society*, 152(3):349–361. (Cited on page 2.)
- Brieber, S., Neufang, S., Bruning, N., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., Fink, G. R., and Konrad, K. (2007). Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, 48(12):1251–1258. (Cited on page 17.)
- Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376. (Cited on page 51.)
- Burns, K. and Gidea, M. (2005). *Differential geometry and topology: with a view to dynamical systems*. CRC Press. (Cited on pages 12, 78, and 89.)
- Bushong, S. C. (2003). *Magnetic resonance imaging*. Mosby. (Cited on page 6.)
- Buss, S. R. and Fillmore, J. P. (2001). Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, 20(2):95–126. (Cited on pages 136, 137, and 142.)
- Butcher, J. C. (2000). Numerical methods for ordinary differential equations in the 20th century. *Journal of Computational and Applied Mathematics*, 125(1):1–29. (Cited on page 121.)
- Butcher, J. C. (2008). *Numerical methods for ordinary differential equations*. Wiley. (Cited on pages 121 and 122.)
- Cachier, P. and Ayache, N. (2004). Isotropic energies, filters and splines for vector field regularization. *Journal of Mathematical Imaging and Vision*, 20(3):251–265. (Cited on pages 54 and 66.)
- Cachier, P., Bardinet, E., Dormont, D., Pennec, X., and Ayache, N. (2003). Iconic feature based nonrigid registration: the PASHA algorithm. *Computer Vision and Image Understanding*, 89(2):272–298. (Cited on page 53.)
- Cao, Y., Miller, M. I., Winslow, R. L., and Younes, L. (2005). Large deformation diffeomorphic metric mapping of vector fields. *IEEE Transactions on Medical Imaging*, 24(9):1216–1230. (Cited on page 55.)
- Cardenas, V. A., Boxer, A. L., Chao, L. L., Gorno-Tempini, M. L., Miller, B. L., Weiner, M. W., and Studholme, C. (2007a). Deformation-based morphometry reveals brain atrophy in frontotemporal dementia. *Archives of Neurology*, 64(6):873–877. (Cited on page 146.)
- Cardenas, V. A., Studholme, C., Gazdzinski, S., Durazzo, T. C., and Meyerhoff, D. J. (2007b). Deformation-based morphometry of brain changes in alcohol dependence and abstinence. *NeuroImage*, 34(3):879–887. (Cited on page 146.)
- Casas, F. and Murua, A. (2009). An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications. *Journal of Mathematical Physics*, 50(3):033513–1–033513–23. (Cited on page 71.)
- Casas, F., Murua, A., and Nadinic, M. (2012). Efficient computation of the Zassenhaus formula. *Computer Physics Communications*, 183(11):2386–2391. (Cited on page 71.)

- Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79. (Cited on pages 54 and 77.)
- Caselles, V., Meinhardt, E., and Monasse, P. (2008). Constructing the tree of shapes of an image by fusion of the trees of connected components of upper and lower level sets. *Positivity*, 12(1):55–73. (Cited on page 28.)
- Caselles, V. and Monasse, P. (2010). *Geometric description of images as topographic maps*. Lecture notes in Mathematics. Springer Berlin Heidelberg. (Cited on page 28.)
- Caselles, V., Morel, J.-M., and Sbert, C. (1998). An axiomatic approach to image interpolation. *IEEE Transactions on Image Processing*, 7(3):376–386. (Cited on page 29.)
- Cendra, H., Holm, D. D., Marsden, J. E., and Ratiu, T. S. (1998). Lagrangian reduction, the Euler–Poincaré equations, and semidirect products. *AMS Translations: Series 2*, 186:1–26. (Cited on page 99.)
- Cendra, H., Marsden, J. E., Pekarsky, S., and Ratiu, T. S. (2003). Variational principles for Lie–Poisson and Hamilton–Poincaré equations. *Moscow Mathematical Journal*, 3(3):833–867. (Cited on page 80.)
- Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277. (Cited on page 54.)
- Chang, C. T., Gorissen, B., and Melchior, S. (2011). Fast oriented bounding box optimization on the rotation group $\text{SO}(3,\mathbb{R})$. *ACM Transactions on Graphics*, 30(5):122:1–122:16. (Cited on page 77.)
- Chefd’hotel, C. (2007). A method for the transport and registration of images on implicit surfaces. In *Scale Space and Variational Methods in Computer Vision*, pages 860–870. Springer. (Cited on page 24.)
- Cheng, C. M., Vong, S. W., and Wenzel, D. (2010). Commutators with maximal Frobenius norm. *Linear Algebra and its Applications*, 432(1):292–306. (Cited on page 126.)
- Chikuse, Y. (2003). *Statistics on special manifolds*. Lecture Notes in Statistics. Springer. (Cited on page 136.)
- Choi, S.-S., Cha, S.-H., and Tappert, C. (2010). A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, 8(1):43–48. (Cited on page 52.)
- Chong, E. K. P. and Zak, S. H. (2013). *An introduction to optimization*. John Wiley & Sons. (Cited on pages 57 and 177.)
- Christensen, G. E. and Johnson, H. J. (2003). Invertibility and transitivity analysis for nonrigid image registration. *Journal of Electronic Imaging*, 12(1):106–117. (Cited on page 171.)
- Christensen, G. E., Johnson, H. J., and Vannier, M. W. (2006). Synthesizing average 3D anatomical shapes. *NeuroImage*, 32(1):146–158. (Cited on page 161.)
- Chui, H. and Rangarajan, A. (2003). A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141. (Cited on page 17.)
- Chumbley, J. R. and Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using gaussian random fields. *NeuroImage*, 44(1):62–70. (Cited on page 168.)

- Chumchob, N. and Chen, K. (2010). A variational approach for discontinuity-preserving image registration. *East-West Journal of Mathematics*, 2010:266–282. (Cited on page 55.)
- Chumchob, N., Chen, K., and Brito-Loeza, C. (2011). A fourth-order variational image registration model and its fast multigrid algorithm. *Multiscale Modeling & Simulation*, 9(1):89–128. (Cited on page 55.)
- Chung, A. C. S., Wells III, W. M., Norbush, A., and Grimson, W. E. L. (2002). Multi-modal image registration by minimising Kullback–Leibler distance. In Dohi, T. and Kikinis, R., editors, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2002*, volume 2489 of *Lecture Notes in Computer Science*, pages 525–532. Springer Berlin Heidelberg. (Cited on page 53.)
- Chung, M. K., Worsley, K. J., Paus, T., Cherif, C., Collins, D. L., Giedd, J. N., Rapoport, J. L., and Evans, A. C. (2001). A unified statistical approach to deformation-based morphometry. *NeuroImage*, 14(3):595–606. (Cited on page 18.)
- Ciarlet, P. G. (1988). *Three-dimensional elasticity*, volume 1. Elsevier. (Cited on page 173.)
- Cockrell, J. and Folstein, M. (2002). *Principles and Practice of Geriatric Psychiatry*, chapter 27. Mini-Mental State Examination, pages 140–141. John Wiley & Sons, second edition. (Cited on pages 15 and 147.)
- Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., and Marchal, G. (1995). Automated multi-modality image registration based on information theory. In *Information Processing in Medical Imaging*, volume 3, pages 263–274. (Cited on page 53.)
- Colliot, O., Chételat, G., Chupin, M., Desgranges, B., Magnin, B., Benali, H., Dubois, B., Garnero, L., Eustache, F., and Lehéricy, S. (2008). Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology*, 248(1):194–201. (Cited on page 17.)
- Commenges, D. (2003). Transformations which preserve exchangeability and application to permutation tests. *Journal of Nonparametric Statistics*, 15(2):171–185. (Cited on page 158.)
- Constantin, A. and Kolev, B. (2002). On the geometric approach to the motion of inertial mechanical systems. *Journal of Physics A: Mathematical and General*, 35(32):R51–R79. (Cited on page 77.)
- Constantinides, C. (2014). *Magnetic resonance imaging: the basics*. Taylor & Francis. (Cited on page 6.)
- Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59. (Cited on pages 17 and 149.)
- Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmeichel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L., and Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer’s disease in late onset families. *Science*, 261(5123):921–923. (Cited on page 15.)
- Cotter, C. J. (2008). The variational particle-mesh method for matching curves. *Journal of Physics A: Mathematical and Theoretical*, 41(34):344003. (Cited on page 55.)

- Craig, J. J. (2004). *Introduction to robotics: mechanics and control*. Prentice Hall. (Cited on page 40.)
- Craig, M. C., Zaman, S. H., Daly, E. M., Cutter, W. J., Robertson, D. M. W., Hallahan, B., Toal, F., Reed, S., Ambikapathy, A., Brammer, M., Murphy, C. M., and Murphy, D. G. (2007). Women with autistic-spectrum disorder: magnetic resonance imaging study of brain anatomy. *British Journal of Psychiatry*, 191(3):224–228. (Cited on page 17.)
- Criminisi, A., Sharp, T., and Blake, A. (2008). GeoS: Geodesic image segmentation. In Forsyth, D., Torr, P., and Zisserman, A., editors, *Computer Vision–ECCV 2008*, volume 5302 of *Lecture Notes in Computer Science*, pages 99–112. Springer Berlin Heidelberg. (Cited on page 77.)
- Crouch, P. E., Nordkvist, N., and Sanyal, A. K. (2010). Optimal control and geodesics on matrix Lie groups. In *Portuguese Conference on Automatic Control–CONTROLO 2010*. (Cited on pages 101 and 102.)
- Crum, W. R., Camara, O., and Hill, D. L. G. (2006). Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Transactions on Medical Imaging*, 25(11):1451–1461. (Cited on page 53.)
- Crum, W. R., Griffin, L. D., Hill, D. L. G., and Hawkes, D. J. (2003). Zen and the art of medical image registration: correspondence, homology, and quality. *NeuroImage*, 20(3):1425–1437. (Cited on page 51.)
- Crum, W. R., Hartkens, T., and Hill, D. L. G. (2004). Non-rigid image registration: theory and practice. *British Journal of Radiology*, 77(Supplement 2):S140–S153. (Cited on page 51.)
- Csernansky, J. G., Wang, L., Joshi, S., Miller, J. P., Gado, M., Kido, D., McKeel, D., Morris, J. C., and Miller, M. I. (2000). Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. *Neurology*, 55(11):1636–1643. (Cited on page 146.)
- Csernansky, J. G., Wang, L., Joshi, S. C., Ratnanather, J. T., and Miller, M. I. (2004). Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. *NeuroImage*, 23, Supplement 1:S56–68. (Cited on page 146.)
- Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194. (Cited on page 18.)
- Datar, M., Lyu, I., Kim, S.-H., Cates, J., Styner, M. A., and Whitaker, R. (2013). Geodesic distances to landmarks for dense correspondence on ensembles of complex shapes. In Mori, K., Sakuma, I., Sato, Y., Barillot, C., and Navab, N., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, volume 8150 of *Lecture Notes in Computer Science*, pages 19–26. Springer Berlin Heidelberg. (Cited on page 17.)
- Davatzikos, C., Genc, A., Xu, D., and Resnick, S. M. (2001). Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage*, 14(6):1361–1369. (Cited on pages 18 and 164.)
- Dawbarn, D. and Allen, S. J. (2007). *Neurobiology of Alzheimer's disease*. Oxford University Press. (Cited on page 14.)
- de Levie, R. (2009). An improved numerical approximation for the first derivative. *Journal of Chemical Sciences*, 121(5):935–950. (Cited on page 128.)

de Pedro-Cuesta, J., Virues-Ortega, J., Vega, S., Seijo-Martinez, M., Saz, P., Rodriguez, F., Rodriguez-Laso, A., Rene, R., de las Heras, S., Mateos, R., Martinez-Martin, P., Manubens, J., Mahillo-Fernandez, I., Lopez-Pousa, S., Lobo, A., Regla, J., Gascon, J., Garcia, F., Fernandez-Martinez, M., Boix, R., Bermejo-Pareja, F., Bergareche, A., Benito-Leon, J., de Arce, A., and del Barrio, J. (2009). Prevalence of dementia and major dementia subtypes in Spanish populations: A reanalysis of dementia prevalence surveys, 1990-2008. *BMC Neurology*, 9(1):55. (Cited on page 15.)

Dedieu, J. P. and Nowicki, D. (2005). Symplectic methods for the approximation of the exponential map and the Newton iteration on Riemannian submanifolds. *Journal of Complexity*, 21(4):487–501. (Cited on pages 77 and 92.)

Derfoul, R. and Le Guyader, C. (2013). A relaxed problem of registration based on the Saint Venant–Kirchhoff material stored energy for the mapping of mouse brain gene expression data to a neuroanatomical mouse atlas. In *MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, page 71. (Cited on page 55.)

Dey, T. K. and Levine, J. A. (2008). Delaunay meshing of isosurfaces. *The Visual Computer*, 24(6):411–422. (Cited on page 27.)

Do Carmo, M. P. (1992). *Riemannian geometry*. Birkhäuser. (Cited on pages 13, 77, 78, 88, and 108.)

Doby, T. (1997). *Origins and development of medical imaging*. SIU Press. (Cited on page 2.)

Dodgson, N. A. (1997). Quadratic interpolation for image resampling. *IEEE Transactions on Image Processing*, 6(9):1322–1326. (Cited on page 29.)

Dodwell, P. C. (1983). The Lie transformation group model of visual perception. *Perception & Psychophysics*, 34(1):1–16. (Cited on page 9.)

Drachman, D. A. (2014). The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer’s disease. *Alzheimer’s & Dementia*, 10(3):372–380. (Cited on page 16.)

Droske, M. and Rumpf, M. (2004). A variational approach to non-rigid morphological image registration. *SIAM Journal on Applied Mathematics*, 64(2):668–687. (Cited on pages 53 and 54.)

Dru, F., Fillard, P., and Vercauteren, T. (2010). An ITK implementation of the symmetric log-domain diffeomorphic demons algorithm. *Insight Journal*. (Cited on page 71.)

Drummond, T. and Cipolla, R. (2000). Application of Lie algebras to visual servoing. *International Journal of Computer Vision*, 37(1):21–41. (Cited on page 77.)

Dryden, I. L. and Mardia, K. V. (1998). *Statistical shape analysis*. Wiley. (Cited on pages 17, 24, and 149.)

Du, A. T., Schuff, N., Amend, D., Laakso, M. P., Hsu, Y. Y., Jagust, W. J., Yaffe, K., Kramer, J. H., Reed, B., Norman, D., Chui, H. C., and Weiner, M. W. (2001). Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer’s disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 71(4):441–447. (Cited on page 146.)

- Du, A. T., Schuff, N., Zhu, X. P., Jagust, W. J., Miller, B. L., Reed, B. R., Kramer, J. H., Mungas, D., Yaffe, K., Chui, H. C., and Weiner, M. W. (2003). Atrophy rates of entorhinal cortex in AD and normal aging. *Neurology*, 60(3):481. (Cited on page 146.)
- Du, J., Younes, L., and Qiu, A. (2011). Whole brain diffeomorphic metric mapping via integration of sulcal and gyral curves, cortical surfaces, and images. *NeuroImage*, 56(1):162–173. (Cited on page 32.)
- Dubrovin, B. A., Fomenko, A. T., and Novikov, S. P. (1984). *Modern geometry—Methods and applications: Part I: the geometry of surfaces, transformation groups, and fields*. Graduate Texts in Mathematics. Springer-Verlag. (Cited on page 40.)
- Dubuisson, M.-P. and Jain, A. K. (1994). A modified Hausdorff distance for object matching. In *IAPR International Conference on Pattern Recognition-Conference A: Computer Vision & Image Processing*, volume 1, pages 566–568. IEEE. (Cited on page 52.)
- Dupuis, P., Grenander, U., and Miller, M. I. (1998). Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 56(3):587. (Cited on pages 44 and 55.)
- Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E. M., Brammer, M. J., Maltezos, S., Murphy, C. M., Robertson, D., Williams, S. C., and Murphy, D. G. (2010). Describing the brain in autism in five dimensions—magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *Journal of Geometric Mechanics*, 30(32):10612–10623. (Cited on page 18.)
- Edgington, E. (1969). *Statistical inference: the distribution-free approach*. McGraw-Hill. (Cited on page 167.)
- Ewers, M., Sperling, R. A., Klunk, W. E., Weiner, M. W., and Hampel, H. (2011). Neuroimaging markers for the prediction and early diagnosis of Alzheimer’s disease dementia. *Trends in Neurosciences*, 34(8):430–442. (Cited on page 16.)
- Farin, G. E. (2001). *Curves and surfaces for computer-aided geometric design: a practical guide*. The Morgan Kaufmann Series in Computer Graphics. Morgan Kaufmann. (Cited on page 27.)
- Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., Myers, R. H., Pericak-Vance, M. A., Risch, N., and van Duijn, C. M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis. *Journal of the American Medical Association*, 278(16):1349–1356. (Cited on page 15.)
- Fazekas, F., Kleinert, R., Offenbacher, H., Schmidt, R., Kleinert, G., Payer, F., Radner, H., and Lechner, H. (1993). Pathologic correlates of incidental MRI white matter signal hyperintensities. *Neurology*, 43(9):1683–1683. (Cited on page 7.)
- Fillard, P., Arsigny, V., Ayache, N., and Pennec, X. (2005). A Riemannian framework for the processing of tensor-valued images. In Olsen, O. F., Florack, L., and Kuijper, A., editors, *Deep Structure, Singularities, and Computer Vision*, volume 3753 of *Lecture Notes in Computer Science*, pages 112–123. Springer Berlin Heidelberg. (Cited on page 76.)
- Fischer, B. and Modersitzki, J. (2008). Ill-posed medicine—an introduction to image registration. *Inverse Problems*, 24(3):034008. (Cited on page 51.)

- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355. (Cited on page 28.)
- Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., Quinn, B. T., and Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23, Supplement 1(0):S69–S84. (Cited on page 28.)
- Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207. (Cited on page 18.)
- Fleet, D. and Weiss, Y. (2006). Optical flow estimation. In *Handbook of Mathematical Models in Computer Vision*, pages 237–257. Springer US. (Cited on page 51.)
- Fletcher, P. T. (2013). Geodesic regression and the theory of least squares on Riemannian manifolds. *International Journal of Computer Vision*, 105(2):171–185. (Cited on pages 76, 152, and 155.)
- Fletcher, P. T. and Joshi, S. (2007). Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, 87(2):250–262. (Cited on pages 76, 85, 136, and 170.)
- Fletcher, P. T., Lu, C., Pizer, S. M., and Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005. (Cited on page 17.)
- Fletcher, P. T., Venkatasubramanian, S., and Joshi, S. (2009). The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1, Supplement 1):S143. (Cited on page 85.)
- Flusser, J. and Zitová, B. (2013). A comment on “a novel approach for the registration of weak affine images”. *Pattern Recognition Letters*, 34(12):1381–1385. (Cited on page 42.)
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). “mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198. (Cited on page 15.)
- Fortea, J., Sala-Llonch, R., Bartrés-Faz, D., Bosch, B., Lladó, A., Bargalló, N., Molinuevo, J. L., and Sánchez-Valle, R. (2010). Increased cortical thickness and caudate volume precede atrophy in PSEN1 mutation carriers. *Journal of Alzheimer’s Disease*, 22(3):909–922. (Cited on page 18.)
- Fox, N. C., Crum, W. R., Scahill, R. I., Stevens, J. M., Janssen, J. C., and Rossor, M. N. (2001). Imaging of onset and progression of Alzheimer’s disease with voxel-compression mapping of serial magnetic resonance images. *The Lancet*, 358(9277):201–205. (Cited on page 18.)
- Fox, N. C. and Freeborough, P. A. (1997). Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer’s disease. *Journal of Magnetic Resonance Imaging*, 7(6):1069–1075. (Cited on page 17.)
- Frangi, A. F., Rueckert, D., Schnabel, J. A., and Niessen, W. J. (2002). Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. *IEEE Transactions on Medical Imaging*, 21(9):1151–1166. (Cited on pages 17 and 160.)

- Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. In *Annales de l'institut Henri Poincaré*, volume 10, pages 215–310. Presses universitaires de France. (Cited on page 137.)
- Freeborough, P. A. and Fox, N. C. (1997). The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Transactions on Medical Imaging*, 16(5):623–629. (Cited on page 17.)
- Freifeld, O. (2013). *Statistics on Manifolds with Applications to Modeling Shape Deformations*. PhD thesis, Brown University. (Cited on pages 136 and 154.)
- Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., and Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nature Reviews Neurology*, 6(2):67–77. (Cited on page 16.)
- Gabay, D. (1982). Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219. (Cited on page 77.)
- Gaens, T., Maes, F., Vandermeulen, D., and Suetens, P. (1998). Non-rigid multimodal image registration using mutual information. In Wells, W. M., Colchester, A., and Delp, S., editors, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 1998*, volume 1496 of *Lecture Notes in Computer Science*, pages 1099–1106. Springer Berlin Heidelberg. (Cited on page 53.)
- Gallier, J. (2011). Logarithms and square roots of real matrices. Existence, uniqueness and applications in medical imaging. Unpublished, accessed from: <http://www.cis.upenn.edu/~cis610/matlog.pdf>. (Cited on page 83.)
- Gallier, J. (2013). Notes on differential geometry and Lie groups. Unpublished, accessed from: <http://www.cis.upenn.edu/~jean/gbooks/manif.html>. (Cited on pages 77, 80, 82, 87, and 88.)
- Gallier, J. and Xu, D. (2002). Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices. *International Journal of Robotics and Automation*, 17(4). (Cited on pages 40 and 116.)
- Gallot, S., Hulin, D., and Lafontaine, J. (2004). *Riemannian geometry*. Springer-Verlag. (Cited on pages 13, 77, 78, 117, and 119.)
- Garrison, D. H. and Hast, M. H. (2013). *The Fabric of the Human Body: an Annotated Translation of the 1543 and 1555 Editions of “De Humani Corporis Fabrica Librum Septem”*. S. Karger, 1 edition. (Cited on page 2.)
- Gaser, C., Nenadic, I., Buchsbaum, B. R., Hazlett, E. A., and Buchsbaum, M. S. (2001). Deformation-based morphometry and its relation to conventional volumetry of brain lateral ventricles in MRI. *NeuroImage*, 13(6):1140–1145. (Cited on page 18.)
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878. (Cited on page 168.)
- Giachetti, A. (2000). Matching techniques to compute image motion. *Image and Vision Computing*, 18(3):247–260. (Cited on page 53.)

- Gibby, W. A. (2005). Basic principles of magnetic resonance imaging. *Neurosurgery Clinics of North America*, 16(1):1–64. (Cited on page 6.)
- Gilmore, R. (2005). *Lie groups, Lie algebras, and some of their applications*. Dover books on mathematics. Dover Publications. (Cited on page 36.)
- Glaunès, J., Trouvé, A., and Younes, L. (2004a). Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages II–712–II–718. IEEE. (Cited on page 55.)
- Glaunès, J., Vaillant, M., and Miller, M. I. (2004b). Landmark matching via large deformation diffeomorphisms on the sphere. *Journal of Mathematical Imaging and Vision*, 20(1-2):179–200. (Cited on page 24.)
- Glöckner, H. (2006). Fundamental problems in the theory of infinite-dimensional Lie groups. *arXiv preprint*. (Cited on page 71.)
- Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J., and Frackowiak, R. S. J. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, 14(1):21–36. (Cited on page 161.)
- Good, P. I. (2000). *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Series in Statistics. Springer. (Cited on page 158.)
- Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 285–339. (Cited on page 149.)
- Gorczowski, K., Styner, M., Jeong, J. Y., Marron, J. S., Piven, J., Hazlett, H. C., Pizer, S. M., and Gerig, G. (2010). Multi-object analysis of volume, pose, and shape using statistical discrimination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):652–661. (Cited on pages 76, 77, and 147.)
- Gori, P., Colliot, O., Worbe, Y., Marrakchi-Kacem, L., Lecomte, S., Poupon, C., Hartmann, A., Ayache, N., and Durrleman, S. (2013). Bayesian atlas estimation for the variability analysis of shape complexes. In Mori, K., Sakuma, I., Sato, Y., Barillot, C., and Navab, N., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, volume 8149 of *Lecture Notes in Computer Science*, pages 267–274. Springer Berlin Heidelberg. (Cited on pages 28 and 32.)
- Goshtasby, A. A. (2005). *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. Wiley. (Cited on page 51.)
- Goshtasby, A. A. (2012). *Image registration: Principles, tools and methods*. Springer. (Cited on pages 7, 9, 51, and 54.)
- Govindu, V. (2006). Robustness in motion averaging. In Narayanan, P. J., Nayar, S. K., and Shum, H.-Y., editors, *Asian Conference on Computer Vision–ACCV 2006*, volume 3852 of *Lecture Notes in Computer Science*, pages 457–466. Springer Berlin Heidelberg. (Cited on page 78.)
- Graff-Radford, N. R., Green, R. C., Go, R. C. P., Hutton, M. L., Edeki, T., Bachman, D., Adamson, J. L., Griffith, P., Willis, F. B., Williams, M., Hipps, Y., Haines, J. L., Cupples, L. A., and Farrer, L. A. (2002). Association between apolipoprotein E genotype and Alzheimer

- disease in African American subjects. *Archives of Neurology*, 59(4):594–600. (Cited on page 15.)
- Gramkow, C. (2001). On averaging rotations. *Journal of Mathematical Imaging and Vision*, 15(1):7–16. (Cited on pages 76 and 125.)
- Grassia, F. S. (1998). Practical parameterization of rotations using the exponential map. *Journal of Graphics Tools*, 3(3):1–13. (Cited on page 40.)
- Grenander, U. (1970). A unified approach to pattern analysis. *Advances in Computers*, 10:175–216. (Cited on page 4.)
- Grenander, U. (2008). *Probabilities on algebraic structures*. Courier Dover Publications. (Cited on page 154.)
- Grenander, U. and Miller, M. I. (1998). Computational anatomy: an emerging discipline. *Quarterly of Applied Mathematics*, 56(4):617–694. (Cited on pages 4 and 24.)
- Guetter, C., Xu, C., Sauer, F., and Hornegger, J. (2005). Learning based non-rigid multi-modal image registration using Kullback–Leibler divergence. In Duncan, J. S. and Gerig, G., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, volume 3750 of *Lecture Notes in Computer Science*, pages 255–262. Springer Berlin Heidelberg. (Cited on page 53.)
- Guggenheim, H. W. (1967). *Plane geometry and its groups*. Holden Day. (Cited on page 40.)
- Gur, Y. and Sochen, N. (2009). Regularizing flows over Lie groups. *Journal of Mathematical Imaging and Vision*, 33(2):195–208. (Cited on page 76.)
- Gwak, S., Kim, J., and Park, F. C. (2003). Numerical optimization on the Euclidean group with applications to camera calibration. *IEEE Transactions on Robotics and Automation*, 19(1):65–74. (Cited on page 77.)
- Haber, E. and Modersitzki, J. (2006). Intensity gradient based registration and fusion of multi-modal images. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, volume 4191 of *Lecture Notes in Computer Science*, pages 726–733. Springer Berlin Heidelberg. (Cited on page 53.)
- Hairer, E., Nørsett, S. P., and Wanner, G. (2011). *Solving ordinary differential equations I: nonstiff problems*. Springer. (Cited on page 121.)
- Hajnal, J. V. and Hill, D. L. G. (2010). *Medical image registration*. CRC Press. (Cited on pages 7, 51, and 54.)
- Hall, B. C. (2003). *Lie groups, Lie algebras, and representations: an elementary introduction*. Springer. (Cited on pages 36, 78, and 81.)
- Hamann, B., Trott, I. J., and Farin, G. E. (1997). On approximating contours of the piecewise trilinear interpolant using triangular rational quadratic Bezier patches. *IEEE Transactions on Visualization and Computer Graphics*, 3(3):215–227. (Cited on page 27.)
- Hampel, H., Lista, S., Teipel, S. J., Garaci, F., Nisticò, R., Blennow, K., Zetterberg, H., Bertram, L., Duyckaerts, C., Bakardjian, H., Drzezga, A., Colliot, O., Epelbaum, S., Broich, K., Lehéricy, S., Brice, A., Khachaturian, Z. S., Aisen, P. S., and Dubois, B. (2013). Perspective on future role of biological markers in clinical therapy trials of Alzheimer’s disease: A long-range point of view beyond 2020. *Biochemical Pharmacology*. (Cited on page 15.)

- Han, Q. and Hong, J.-X. (2006). *Isometric embedding of Riemannian manifolds in Euclidean spaces*. American Mathematical Society. (Cited on page 92.)
- Hanson, A. J. (2006). *Visualizing quaternions*. Morgan Kaufmann. (Cited on page 40.)
- Hartley, R., Trumpf, J., Dai, Y., and Li, H. (2013). Rotation averaging. *International Journal of Computer Vision*, 103(3):267–305. (Cited on pages 116, 136, 137, and 142.)
- Hartley, R. and Zisserman, A. (2004). *Multiple view geometry in computer vision*. Cambridge University Press. (Cited on page 118.)
- Hartley, R. I. and Kahl, F. (2009). Global optimization through rotation space search. *International Journal of Computer Vision*, 82(1):64–79. (Cited on page 77.)
- Hawkes, D. J. (1998). Algorithms for radiological image registration and their clinical application. *Journal of Anatomy*, 193(3):347–361. (Cited on page 51.)
- Hawkins, A. and Grimm, C. M. (2005). Keyframing using linear interpolation of matrices. In *ACM SIGGRAPH 2005 Posters*, SIGGRAPH '05. ACM. (Cited on page 135.)
- Heckemann, R. A., Hajnal, J. V., Aljabar, P., Rueckert, D., and Hammers, A. (2006). Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115–126. (Cited on page 160.)
- Helgason, S. (1979). *Differential geometry, Lie groups, and symmetric spaces*. Academic Press. (Cited on page 117.)
- Heo, G. and Small, C. G. (2006). Form representations and means for landmarks: A survey and comparative study. *Computer Vision and Image Understanding*, 102(2):188–203. (Cited on page 28.)
- Hermosillo, G., Chefd'Hotel, C., and Faugeras, O. (2002). Variational methods for multimodal image matching. *International Journal of Computer Vision*, 50(3):329–343. (Cited on page 53.)
- Hernandez, M., Bossa, M., and Olmos, S. (2009). Registration of anatomical images using paths of diffeomorphisms parameterized with stationary vector field flows. *International Journal of Computer Vision*, 85(3):291–306. (Cited on page 55.)
- Hibar, D. P., Stein, J. L., Kohannim, O., Jahanshad, N., Saykin, A. J., Shen, L., Kim, S., Pankratz, N., Foroud, T., Huettel, M. J., Potkin, S. G., Jr., C. R. J., Weiner, M. W., Toga, A. W., and Thompson, P. M. (2011). Voxelwise gene-wide association study (vGeneWAS): Multivariate gene-based association testing in 731 elderly subjects. *NeuroImage*, 56(4):1875–1891. (Cited on page 146.)
- Higham, N. (2005). The scaling and squaring method for the matrix exponential revisited. *SIAM Journal on Matrix Analysis and Applications*, 26(4):1179–1193. (Cited on page 46.)
- Higham, N. J. and Al-Mohy, A. H. (2010). Computing matrix functions. *Acta Numerica*, 19:159–208. (Cited on pages 82 and 83.)
- Higham, N. J. and Relton, S. D. (2013). Higher order Fréchet derivatives of matrix functions and the level-2 condition number. Unpublished, accessed from: <http://eprints.ma.man.ac.uk/2066/>. (Cited on page 97.)
- Hill, D. L. G., Batchelor, P. G., Holden, M., and Hawkes, D. J. (2001). Medical image registration. *Physics in Medicine and Biology*, 46(3):R1. (Cited on page 51.)

- Hinkle, J., Fletcher, P. T., and Joshi, S. (2013). Intrinsic polynomials for regression on Riemannian manifolds. *Journal of Mathematical Imaging and Vision*, doi:10.1007/s10851-013-0489-5. (Cited on pages 76, 91, and 98.)
- Hirsch, M. W. (2011). Smooth actions of Lie groups and Lie algebras on manifolds. *Journal of Fixed Point Theory and Applications*, 10(2):219–232. (Cited on page 24.)
- Hirsch, M. W., Smale, S., and Devaney, R. L. (2004). *Differential equations, dynamical systems and an introduction to chaos*. Elsevier. (Cited on page 44.)
- Ho, J., Cheng, G., Salehian, H., and Vemuri, B. (2013). Recursive Karcher expectation estimators and geometric law of large numbers. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 325–332. (Cited on page 138.)
- Hoffman, K. and Kunze, R. (1977). *Algebra lineal*. Prentice/Hall Internacional. (Cited on page 36.)
- Holden, M. (2008). A review of geometric transformations for nonrigid body registration. *IEEE Transactions on Medical Imaging*, 27(1):111–128. (Cited on page 34.)
- Holm, D. D., Schmah, T., Stoica, C., and Ellis, D. C. P. (2009). *Geometric mechanics and symmetry: from finite to infinite dimensions*. Oxford University Press. (Cited on pages 77, 80, 98, 106, and 140.)
- Hömke, L., Frohn-Schauf, C., Henn, S., and Witsch, K. (2007). Total variation based image registration. In Tai, X.-C., Lie, K.-A., Chan, T. F., and Osher, S., editors, *Image Processing Based on Partial Differential Equations*, Mathematics and Visualization, pages 343–361. Springer Berlin Heidelberg. (Cited on page 55.)
- Honea, R., Crow, T. J., Passingham, D., and Mackay, C. E. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *American Journal of Psychiatry*, 162(12):2233–2245. (Cited on page 18.)
- Hormann, K. and Agathos, A. (2001). The point in polygon problem for arbitrary polygons. *Computational Geometry*, 20(3):131–144. (Cited on page 27.)
- Horn, B. K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642. (Cited on page 56.)
- Horn, B. K. P., Hilden, H. M., and Negahdaripour, S. (1988). Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America A*, 5(7):1127–1135. (Cited on page 56.)
- Horner, J., Norman, M. L., and Ripich, D. N. (2007). *Medical Speech-Language Pathology: A Practitioner's Guide*, chapter 7. Dementia: diagnostic approaches and current taxonomies, pages 94–109. Thieme. (Cited on page 14.)
- Hou, H. S. and Andrews, H. (1978). Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(6):508–517. (Cited on page 30.)
- Hou, Z. (2006). A review on MR image intensity inhomogeneity correction. *International Journal of Biomedical Imaging*, 2006. (Cited on pages 10 and 53.)

- Hsu, W. M., Hughes, J. F., and Kaufman, H. (1992). Direct manipulation of free-form deformations. *SIGGRAPH Computer Graphics*, 26(2):177–184. (Cited on page 43.)
- Hua, X., Lee, S., Yanovsky, I., Leow, A. D., Chou, Y., Ho, A. J., Gutman, B., Toga, A. W., Jack, C. R., Bernstein, M. A., Reiman, E. M., Harvey, D. J., Kornak, J., Schuff, N., Alexander, G. E., Weiner, M. W., and Thompson, P. M. (2009). Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects. *NeuroImage*, 48(4):668–681. (Cited on pages 18 and 146.)
- Hua, X., Leow, A. D., Parikshak, N., Lee, S., Chiang, M., Toga, A. W., Jack, C. R., Weiner, M. W., and Thompson, P. M. (2008). Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. *NeuroImage*, 43(3):458–469. (Cited on pages 18, 146, and 164.)
- Huckemann, S., Hotz, T., and Munk, A. (2010). Intrinsic shape analysis: Geodesic PCA for Riemannian manifolds modulo isometric Lie group actions. *Statistica Sinica*, 20(1):1–100. (Cited on page 136.)
- Hulshoff Pol, H. E., Schnack, H. G., Mandl, R. C. W., Brans, R. G. H., van Haren, N. E. M., Baaré, W. F. C., van Oel, C. J., Collins, D. L., Evans, A. C., and Kahn, R. S. (2006). Gray and white matter density changes in monozygotic and same-sex dizygotic twins discordant for schizophrenia using voxel-based morphometry. *NeuroImage*, 31(2):482–488. (Cited on page 18.)
- Huynh, D. Q. (2009). Metrics for 3D rotations: Comparison and analysis. *Journal of Mathematical Imaging and Vision*, 35(2):155–164. (Cited on page 116.)
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P., and Zanna, A. (2000). Lie-group methods. *Acta Numerica*, 9(1):215–365. (Cited on page 121.)
- Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., L. Whitwell, J., Ward, C., Dale, A. M., Felmlee, J. P., Gunter, J. L., Hill, D. L. G., Killiany, R., Schuff, N. T., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C. S., Krueger, G., Ward, H. A., Metzger, G. J., Scott, K. T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J. P., Fleisher, A. S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., and Weiner, M. W. (2008a). The Alzheimer's disease neuroimaging initiative (adni): MRI methods. *Journal of Magnetic Resonance Imaging*, 27(4):685–691. (Cited on page 19.)
- Jack, C. R., Weigand, S. D., Shiung, M. M., Przybelski, S. A., O'Brien, P. C., Gunter, J. L., Knopman, D. S., Boeve, B. F., Smith, G. E., and Petersen, R. C. (2008b). Atrophy rates accelerate in amnestic mild cognitive impairment. *Neurology*, 70(19, Part 2):1740–1752. (Cited on page 17.)
- Jack Jr, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., and Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*, 9(1):119–128. (Cited on page 15.)
- Jack Jr, C. R., Petersen, R. C., Xu, Y. C., O'brien, P. C., Smith, G. E., Ivnik, R. J., Boeve, B. F., Waring, S. C., Tangalos, E. G., and Kokmen, E. (1999). Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology*, 52(7):1397. (Cited on page 146.)

- Jackson, K. R. and Nedialkov, N. S. (2002). Some recent advances in validated methods for IVPs for ODEs. *Applied Numerical Mathematics*, 42(1):269–284. (Cited on page 122.)
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). {FSL}. *NeuroImage*, 62(2):782–790. (Cited on page 148.)
- Jeuris, B., Vandebril, R., and Vandereycken, B. (2012). A survey and comparison of contemporary algorithms for computing the matrix geometric mean. *Electronic Transactions on Numerical Analysis*, 39:379–402. (Cited on page 136.)
- Jia, C. and Evans, B. L. (2013). 3D rotational video stabilization using manifold optimization. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2493–2497. IEEE. (Cited on page 76.)
- Johnson, H. J. and Christensen, G. E. (2002). Consistent landmark and intensity-based image registration. *IEEE Transactions on Medical Imaging*, 21(5):450–461. (Cited on page 32.)
- Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, 2(4). (Cited on page 16.)
- Jorba, A. and Zou, M. (2005). A software package for the numerical integration of ODEs by means of high-order Taylor methods. *Experimental Mathematics*, 14(1):99–117. (Cited on page 123.)
- Joshi, S., Davis, B., Jomier, M., and Gerig, G. (2004). Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160. (Cited on page 161.)
- Joshi, S. C. and Miller, M. I. (2000). Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370. (Cited on page 55.)
- Jurdjevic, V. (1996). *Geometric control theory*. Cambridge University Press. (Cited on page 103.)
- Kadosh, A., Cohen-Or, D., and Yagel, R. (2003). Tricubic interpolation of discrete surfaces for binary volumes. *IEEE Transactions on Visualization and Computer Graphics*, 9(4):580–586. (Cited on page 30.)
- Kang, I. G. and Park, F. C. (1999). Cubic spline algorithms for orientation interpolation. *International Journal for Numerical Methods in Engineering*, 46(1):45–64. (Cited on page 76.)
- Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541. (Cited on page 137.)
- Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331. (Cited on page 54.)
- Kassubek, J., Juengling, F. D., Kioschies, T., Henkel, K., Karitzky, J., Kramer, B., Ecker, D., Andrich, J., Saft, C., Kraus, P., Aschoff, A. J., Ludolph, A. C., and Landwehrmeyer, G. B. (2004). Topography of cerebral atrophy in early Huntington’s disease: a voxel based morphometric MRI study. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(2):213–220. (Cited on page 18.)
- Keller, H. B. (1976). *Numerical solution of two point boundary value problems*. SIAM. (Cited on page 126.)

- Keys, R. (1981). Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(6):1153–1160. (Cited on page 30.)
- Khan, A. R., Wang, L., and Beg, M. F. (2008). FreeSurfer-initiated fully-automated subcortical brain segmentation in MRI using Large Deformation Diffeomorphic Metric Mapping. *NeuroImage*, 41(3):735–746. (Cited on page 28.)
- Khesin, B. (2005). Topological fluid dynamics. *Notices of the AMS*, 52:9–19. (Cited on page 77.)
- Khesin, B. (2008). Groups and topology in the Euler hydrodynamics and KdV. In Craig, W., editor, *Hamiltonian Dynamical Systems and Applications*, NATO Science for Peace and Security Series, pages 93–102. Springer Netherlands. (Cited on page 77.)
- Kim, J. and Fessler, J. A. (2004). Intensity-based image registration using robust correlation coefficients. *IEEE Transactions on Medical Imaging*, 23(11):1430–1444. (Cited on page 53.)
- Kimia, B. B., Frankel, I., and Popescu, A.-M. (2003). Euler spiral for shape completion. *International Journal of Computer Vision*, 54(1-3):159–182. (Cited on page 27.)
- Kimmel, R., Amir, A., and Bruckstein, A. M. (1995). Finding shortest paths on surfaces using level sets propagation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(6):635–640. (Cited on page 77.)
- Kindermann, L. and Georgiev, P. (2002). Modelling iterative roots of mappings in multidimensional spaces. In *Proceedings of the 9th International Conference on Neural Information Processing–ICONIP*, volume 5, pages 2655–2659. (Cited on page 45.)
- Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., and Parseya, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3):786–802. (Cited on page 53.)
- Klein, F. (1893). A comparative review of recent researches in geometry. *Bulletin of the AMS*, 2(10):215–249. (Cited on pages 13 and 150.)
- Klein, S. (2008). *Optimisation methods for medical image registration*. PhD thesis, Image Sciences Institute, UMC Utrecht. (Cited on page 60.)
- Knabner, P., Korotov, S., and Summ, G. (2003). Conditions for the invertibility of the isoparametric mapping for hexahedral finite elements. *Finite Elements in Analysis and Design*, 40(2):159–172. (Cited on page 30.)
- Kochunov, P., Lancaster, J., Thompson, P., Toga, A. W., Brewer, P., Hardies, J., and Fox, P. (2002). An optimized individual target brain in the Talairach coordinate system. *NeuroImage*, 17(2):922–927. (Cited on page 161.)
- Kolar, I., Michor, P. W., and Slovak, J. (1993). *Natural operations in differential geometry*. Electronic Library of Mathematics. Springer. (Cited on page 79.)
- Kolev, B. (2004). Lie groups and mechanics: an introduction. *Journal of Nonlinear Mathematical Physics*, 11:480–498. (Cited on pages 80 and 83.)

- Kolev, B. (2007). Groupes de Lie et mécanique. Unpublished, accessed from: <http://www.cmi.univ-mrs.fr/~kolev/pdf/GLM.pdf>. (Cited on pages 91 and 106.)
- Kovačević, N., Henderson, J. T., Chan, E., Lifshitz, N., Bishop, J., Evans, A., Henkelman, R. M., and Chen, X. J. (2005). A three-dimensional MRI atlas of the mouse brain with estimates of the average and variability. *Cerebral Cortex*, 15(5):639–645. (Cited on page 161.)
- Krasuski, J. S., Alexander, G. E., Horwitz, B., Daly, E. M., Murphy, D. G. M., Rapoport, S. I., and Schapiro, M. B. (1998). Volumes of medial temporal lobe structures in patients with Alzheimer’s disease and mild cognitive impairment (and in healthy controls). *Biological Psychiatry*, 43(1):60–68. (Cited on page 146.)
- Kriegel, A. and Michor, P. W. (1997). *The convenient setting of global analysis*. Ams/IP Studies in Advanced Mathematics. American Mathematical Society. (Cited on page 71.)
- Kuffner, J. J. (2004). Effective sampling and distance metrics for 3D rigid body path planning. In *IEEE International Conference on Robotics and Automation*, volume 4, pages 3993–3998. (Cited on page 76.)
- Kwon, J., Lee, H., Park, F., and Lee, K. (2014). A geometric particle filter for template-based visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(4):625–643. (Cited on page 78.)
- Kwon, J. and Park, F. C. (2010). Visual tracking via particle filtering on the affine group. *International Journal of Robotics Research*, 29(2-3):198–217. (Cited on pages 77 and 78.)
- Laakso, M. P., Soininen, H., Partanen, K., Helkala, E. L., Hartikainen, P., Vainio, P., Hallikainen, M., Hänninen, T., and Riekkinen Sr, P. J. (1995). Volumes of hippocampus, amygdala and frontal lobes in the MRI-based diagnosis of early Alzheimer’s disease: Correlation with memory functions. *Journal of Neural Transmission - Parkinson’s Disease and Dementia Section*, 9(1):73–86. (Cited on page 146.)
- Laquer, H. T. (1992). Invariant affine connections on Lie groups. *Transactions of the AMS*, 331:541–551. (Cited on page 91.)
- Lara, M., Elipe, A., and Palacios, M. (1999). Automatic programming of recurrent power series. *Mathematics and Computers in Simulation*, 49(4-5):351–362. (Cited on page 123.)
- Le, H. (1998). On the consistency of procrustean mean shapes. *Advances in Applied Probability*, pages 53–63. (Cited on page 149.)
- Le Guyader, C. and Vese, L. A. (2009). A combined segmentation and registration framework with a nonlinear elasticity smoother. In Tai, X.-C., Mørken, K., Lysaker, M., and Lie, K.-A., editors, *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Science*, pages 600–611. Springer Berlin Heidelberg. (Cited on page 54.)
- Lee, S., Choi, M., Kim, H., and Park, F. C. (2007). Geometric direct search algorithms for image registration. *IEEE Transactions on Image Processing*, 16(9):2215–2224. (Cited on pages 77, 125, and 176.)
- Lehmann, T. M., Gonner, C., and Spitzer, K. (1999). Survey: interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18(11):1049–1075. (Cited on page 29.)

- Lekien, F. and Marsden, J. (2005). Tricubic interpolation in three dimensions. *International Journal for Numerical Methods in Engineering*, 63(3):455–471. (Cited on page 30.)
- Lenglet, C., Rousson, M., Deriche, R., and Faugeras, O. (2006). Statistics on the manifold of multivariate normal distributions: theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision*, 25(3):423–444. (Cited on page 76.)
- Lepore, N., Brun, C., Chou, Y. Y., Chiang, M. C., Dutton, R. A., Hayashi, K. M., Luders, E., Lopez, O. L., Aizenstein, H. J., Toga, A. W., Becker, J. T., and Thompson, P. M. (2008). Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. *IEEE Transactions on Medical Imaging*, 27(1):129–141. (Cited on pages 18, 170, and 174.)
- Leporé, N., Brun, C., Pennec, X., Chou, Y.-Y., Lopez, O. L., Aizenstein, H. J., Becker, J. T., Toga, A. W., and Thompson, P. M. (2007). Mean template for tensor-based morphometry using deformation tensors. In Ayache, N., Ourselin, S., and Maeder, A., editors, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007*, volume 4792 of *Lecture Notes in Computer Science*, pages 826–833. Springer Berlin Heidelberg. (Cited on page 161.)
- Lepore, N., Brun, C. A., Chiang, M., Chou, Y., Dutton, R. A., Hayashi, K. M., Lopez, O. L., Aizenstein, H. J., Toga, A. W., Becker, J. T., and Thompson, P. M. (2006). Multivariate statistics of the Jacobian matrices in tensor based morphometry and their application to HIV/AIDS. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, volume 4190 of *Lecture Notes in Computer Science*, pages 191–198. Springer Berlin Heidelberg. (Cited on page 170.)
- Lester, H. and Arridge, S. R. (1999). A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32(1):129–149. (Cited on page 51.)
- Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., Schuff, N., Fox, N. C., and Ourselin, S. (2010a). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer’s disease. *NeuroImage*, 51(4):1345–1359. (Cited on page 17.)
- Leung, K. K., Clarkson, M. J., Bartlett, J. W., Clegg, S., Jr., C. R. J., Weiner, M. W., Fox, N. C., and Ourselin, S. (2010b). Robust atrophy rate measurement in Alzheimer’s disease using multi-site serial MRI: Tissue-specific intensity normalization and parameter selection. *NeuroImage*, 50(2):516–523. (Cited on page 17.)
- Lewis, F. L., Vrabie, D., and Syrmos, V. L. (2012). *Optimal control*. Wiley. (Cited on pages 103 and 104.)
- Li, C., Sheng, Y., and Wang, M. (2010). An effective method to compute Fréchet derivative of matrix exponential and its error analysis. *Journal of Information & Computational Science*, 7:1854–1859. (Cited on page 96.)
- Li, G., Liu, Y., Yin, J., and Shi, Z. (2009). Newton geodesic optimization on special linear group. In *IEEE Conference on Decision and Control*, pages 4346–4351. (Cited on page 117.)
- Li, H. and Hartley, R. (2007). The 3D-3D registration problem revisited. In *IEEE International Conference on Computer Vision*, pages 1–8. (Cited on page 77.)
- Li, J. and Hao, P. (2006). Smooth interpolation on homogeneous matrix groups for computer animation. *Journal of Zhejiang University SCIENCE A*, 7(7):1168–1177. (Cited on pages 24, 76, 83, and 133.)

- Li, M., Tan, T., Chen, W., and Huang, K. (2012). Efficient object tracking by incremental self-tuning particle filtering on the affine group. *IEEE Transactions on Image Processing*, 21(3):1298–1313. (Cited on page 78.)
- Liao, S. and Chung, A. C. S. (2012). Nonrigid brain MR image registration using uniform spherical region descriptor. *IEEE Transactions on Image Processing*, 21(1):157–169. (Cited on page 53.)
- Lin, D., Grimson, E., and Fisher, J. (2009). Learning visual flows: A Lie algebraic approach. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 747–754. IEEE Computer Society. (Cited on pages 77 and 78.)
- Liu, J., Chen, Y. Q., Maisog, J. M., and Luta, G. (2010). A new point containment test algorithm based on preprocessing and determining triangles. *Computer-Aided Design*, 42(12):1143–1150. (Cited on page 27.)
- Lombaert, H. (2012). *Atlas Construction for Measuring the Variability of Complex Anatomical Structures*. PhD thesis, University of Montreal. (Cited on page 54.)
- Lombaert, H., Sporring, J., and Siddiqi, K. (2013). Diffeomorphic spectral matching of cortical surfaces. In Gee, J. C., Joshi, S., Pohl, K. M., Wells, W. M., and Zöllei, L., editors, *Information Processing in Medical Imaging*, volume 7917 of *Lecture Notes in Computer Science*, pages 376–389. Springer Berlin Heidelberg. (Cited on page 24.)
- Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *SIGGRAPH Comput. Graph.*, 21(4):163–169. (Cited on page 27.)
- Lorenzen, P., Prastawa, M., Davis, B., Gerig, G., Bullitt, E., and Joshi, S. (2006). Multi-modal image set registration and atlas formation. *Medical Image Analysis*, 10(3):440–451. (Cited on page 161.)
- Lorenzi, M., Ayache, N., Frisoni, G. B., and Pennec, X. (2013). LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. *NeuroImage*, 81:470–483. (Cited on page 53.)
- Lorenzi, M., Ayache, N., and Pennec, X. (2011). Schild's ladder for the parallel transport of deformations in time series of images. In Szekely, G. and Hahn, H., editors, *Proceedings of Information Processing in Medical Imaging (IPMI'11)*, volume 6801 of *Lecture Notes in Computer Science*, pages 463–474. (Cited on page 146.)
- Lorenzi, M. and Pennec, X. (2013a). Efficient parallel transport of deformations in time series of images: from schild's to pole ladder. *Journal of Mathematical Imaging and Vision*. (Cited on page 146.)
- Lorenzi, M. and Pennec, X. (2013b). Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International Journal of Computer Vision*, 105(2):111–127. (Cited on pages 55, 90, and 91.)
- Lorenzo-Valdés, M., Sanchez-Ortiz, G. I., Mohiaddin, R., and Rueckert, D. (2002). Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration. In Dohi, T. and Kikinis, R., editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2002*, volume 2488 of *Lecture Notes in Computer Science*, pages 642–650. Springer Berlin Heidelberg. (Cited on page 160.)

- Lubliner, J. (2008). *Plasticity theory*. Dover books on engineering. Dover Publications. (Cited on page 173.)
- Lui, Y. M. (2012). Advances in matrix manifolds for computer vision. *Image and Vision Computing*, 30(6–7):380–388. (Cited on pages 76 and 136.)
- Lyoo, I. K., Sung, Y. H., Dager, S. R., Friedman, S. D., Lee, J.-Y., Kim, S. J., Kim, N., Dunner, D. L., and Renshaw, P. F. (2006). Regional cerebral cortical thinning in bipolar disorder. *Bipolar Disorders*, 8(1):65–74. (Cited on page 18.)
- Machado, L., Silva Leite, F., and Krakowski, K. (2010). Higher-order smoothing splines versus least squares problems on Riemannian manifolds. *Journal of Dynamical and Control Systems*, 16(1):121–148. (Cited on page 76.)
- Maeland, E. (1988). On the comparison of interpolation methods. *IEEE Transactions on Medical Imaging*, 7(3):213–217. (Cited on page 29.)
- Magnus, J. R. and Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Wiley. (Cited on page 94.)
- Maintz, J. B. and Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36. (Cited on page 51.)
- Makadia, A., Patterson, A. I., and Daniilidis, K. (2006). Fully automatic registration of 3D point clouds. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR’06, pages 1297–1304. IEEE Computer Society. (Cited on page 76.)
- Mangin, J.-F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Collins, D. L., Evans, A. C., and Régis, J. (2004). Object-based morphometry of the cerebral cortex. *IEEE Transactions on Medical Imaging*, 23(8):968–982. (Cited on page 28.)
- Manjón, J. V., Lull, J. J., Carbonell-Caballero, J., García-Martí, G., Martí-Bonmatí, L., and Robles, M. (2007). A nonparametric MRI inhomogeneity correction method. *Medical Image Analysis*, 11(4):336–345. (Cited on page 53.)
- Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in biology*. CRC Press. (Cited on page 158.)
- Manzano, J. M., Llorca, J., Ledesma, A., and Lopez-Ibor, J. J. (1994). Adaptación española de la Alzheimer’s Disease Assessment Scale (ADAS). *Actas Luso Españolas de Neurología y Psiquiatría*, 22:64–70. (Cited on page 15.)
- Mardia, K. V. and Jupp, P. E. (2009). *Directional sStatistics*. Wiley Series in Probability and Statistics. Wiley. (Cited on page 136.)
- Marsden, J. E. (1992). *Lectures on mechanics*. Cambridge University Press. (Cited on pages 98 and 106.)
- Marsland, S. and Twining, C. J. (2004). Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. *IEEE Transactions on Medical Imaging*, 23(8):1006–1020. (Cited on pages 55, 64, and 66.)
- Mathias, R. (1992). Evaluating the Fréchet derivative of the matrix exponential. *Numerische Mathematik*, 63(1):213–226. (Cited on page 96.)

- Mathias, R. (1996). A chain rule for matrix functions and applications. *SIAM Journal on Matrix Analysis and Applications*, 17(3):610–620. (Cited on page 96.)
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Goualher, G. L., Boomsma, D., Cannon, T., Kawashima, R., and Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1412):1293–1322. (Cited on page 161.)
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer’s disease report of the NINCDS-ADRDA work group* under the auspices of Department of Health and Human Services Task Force on Alzheimer’s disease. *Neurology*, 34(7):939–939. (Cited on pages 15 and 147.)
- McLachlan, R. I. and Quispel, G. R. W. (2006). Geometric integrators for ODEs. *Journal of Physics A: Mathematical and General*, 39(19):5251. (Cited on page 121.)
- Meek, D. S. and Walton, D. J. (2004). An arc spline approximation to a clothoid. *Journal of Computational and Applied Mathematics*, 170(1):59–77. (Cited on page 27.)
- Miao, X. and Rao, R. P. N. (2007). Learning the Lie groups of visual invariance. *Neural Computation*, 19(10):2665–2693. (Cited on page 9.)
- Michel, F. and Paragios, N. (2010). Image transport regression using mixture of experts and discrete Markov random fields. In *IEEE International Symposium on Biomedical Imaging*, pages 1229–1232. IEEE. (Cited on page 53.)
- Millán, R. D., Dempere-Marco, L., Pozo, J. M., Cebral, J. R., and Frangi, A. F. (2007). Morphological characterization of intracranial aneurysms using 3-D moment invariants. *IEEE Transactions on Medical Imaging*, 26(9):1270–1282. (Cited on page 149.)
- Miller, M., Banerjee, A., Christensen, G., Joshi, S., Khaneja, N., Grenander, U., and Matejic, L. (1997). Statistical methods in computational anatomy. *Statistical Methods in Medical Research*, 6(3):267–299. (Cited on page 4.)
- Miller, M. I. (2004). Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage*, 23, Supplement 1(0):S19–S33. (Cited on page 4.)
- Miller, M. I., Trouvé, A., and Younes, L. (2002). On the metrics and Euler–Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4(1):375–405. (Cited on pages 24 and 55.)
- Milnor, J. (1976). Curvatures of left invariant metrics on Lie groups. *Advances in Mathematics*, 21(3):293–329. (Cited on pages 77, 88, 109, and 143.)
- Misra, C., Fan, Y., and Davatzikos, C. (2009). Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *NeuroImage*, 44(4):1415–1422. (Cited on page 146.)
- Moakher, M. (2002). Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16. (Cited on pages 76, 116, and 136.)

- Moakher, M. (2005). A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(3):735–747. (Cited on pages 170 and 173.)
- Modersitzki, J. (2004). *Numerical methods for image registration*. Oxford University Press. (Cited on pages 7, 9, 24, 29, and 51.)
- Modersitzki, J. (2009). *FAIR: flexible algorithms for image registration*. SIAM. (Cited on page 51.)
- Modin, K., Perlmutter, M., Marsland, S., and McLachlan, R. (2011). On Euler–Arnold equations and totally geodesic subgroups. *Journal of Geometry and Physics*, 61(8):1446–1461. (Cited on pages 87 and 98.)
- Mohs, R. C. (1994). Administration and scoring manual for the Alzheimer’s disease assessment scale. Research report, New York: Mount Sinai School of Medicine. (Cited on pages 15 and 147.)
- Mohs, R. C. and Cohen, L. (1988). Alzheimer’s Disease Assessment Scale (ADAS). *Psychopharmacology Bulletin*, 24:627–628. (Cited on page 15.)
- Moler, C. and Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45:3–49. (Cited on pages 82, 114, and 135.)
- Moore, D. S., McCabe, G. P., and Craig, B. A. (2012). *Introduction to the practice of statistics*. W. H. Freeman & Company. (Cited on page 158.)
- Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to interval analysis*. SIAM. (Cited on page 122.)
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*, 43(11):2412–2414. (Cited on pages 15 and 147.)
- Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., and Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Archives of Neurology*, 58(3):397–405. (Cited on page 15.)
- Mortenson, M. E. (1997). *Geometric modeling*. Wiley. (Cited on page 24.)
- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., Jagust, W., Trojanowski, J. Q., Toga, A. W., and Beckett, L. (2005). The Alzheimer’s disease neuroimaging initiative. *Neuroimaging Clinics of North America*, 15(4):869–877. (Cited on page 19.)
- Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685. (Cited on page 28.)
- Murray, R. M., Li, Z., and Sastry, S. S. (1994). *A mathematical introduction to robotic manipulation*. CRC Press. (Cited on pages 40 and 77.)
- Myronenko, A. and Song, X. (2010). Intensity-based image registration by minimizing residual complexity. *IEEE Transactions on Medical Imaging*, 29(11):1882–1891. (Cited on page 53.)
- Nagel, H.-H. and Enkelmann, W. (1986). An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (5):565–593. (Cited on page 54.)

- Najfeld, I. and Havel, T. F. (1995). Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics*, 16(3):321–375. (Cited on page 96.)
- Nesvåg, R., Lawyer, G., Varnäs, K., Fjell, A. M., Walhovd, K. B., Frigessi, A., Jönsson, E. G., and Agartz, I. (2008). Regional thinning of the cerebral cortex in schizophrenia: effects of diagnosis, age and antipsychotic medication. *Schizophrenia Research*, 98(1):16–28. (Cited on page 18.)
- Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1):1–25. (Cited on pages 158 and 168.)
- Noakes, L. and Popiel, T. (2007). Geometry for robot path planning. *Robotica*, 25(6):691–702. (Cited on page 77.)
- Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer. (Cited on pages 57 and 177.)
- Nordkvist, N., Crouch, P. E., Bloch, A. M., and Sanyal, A. K. (2011). Embedded optimal control problems. In *IEEE Conference on Decision and Control and European Control Conference*, pages 7311–7316. IEEE. (Cited on page 102.)
- Nugent, A. C., Milham, M. P., Bain, E. E., Mah, L., Cannon, D. M., Marrett, S., Zarate, C. A., Pine, D. S., Price, J. L., and Drevets, W. C. (2006). Cortical abnormalities in bipolar disorder investigated with MRI and voxel-based morphometry. *NeuroImage*, 30(2):485–497. (Cited on page 18.)
- Oliveira, F. P. M. and Tavares, J. M. R. S. (2014). Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2):73–93. (Cited on page 51.)
- Olsson, C., Kahl, F., and Oskarsson, M. (2006). The registration problem revisited: optimal solutions from points, lines and planes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1206–1213. (Cited on page 77.)
- Ono, M., Kubick, S., and Albernathay, C. D. (1990). *Atlas of the cerebral sulci*. Thieme. (Cited on pages 5 and 28.)
- Orchard, J. (2005). Image deformation using velocity fields: An exact solution. In *Proceedings of the Second International Conference on Image Analysis and Recognition*, ICIAR’05, pages 439–446, Berlin, Heidelberg. Springer-Verlag. (Cited on page 45.)
- Organization, W. H. and International, A. D. (2012). Dementia: a public health priority. World Health Organization. Accessed from: <http://www.alz.co.uk/WHO-dementia-report>. (Cited on page 15.)
- Ou, Y., Sotiras, A., Paragios, N., and Davatzikos, C. (2011). DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Medical Image Analysis*, 15(4):622–639. (Cited on page 53.)
- Pantazis, D., Leahy, R. M., Nichols, T. E., and Styner, M. (2004). Statistical surface-based morphometry using a nonparametric approach. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, volume 2, pages 1283–1286. (Cited on page 168.)

- Pantazis, D., Nichols, T. E., Baillet, S., and Leahy, R. M. (2005). A comparison of random field theory and permutation methods for the statistical analysis of MEG data. *NeuroImage*, 25(2):383–394. (Cited on page 169.)
- Park, F. C. (1995). Distance metrics on the rigid-body motions with applications to mechanism design. *Journal of Mechanical Design*, 117(1):48–54. (Cited on pages 13, 77, 86, 88, and 116.)
- Park, F. C., Bobrow, J. E., and Ploen, S. R. (1995). A Lie group formulation of robot dynamics. *International Journal of Robotics Research*, 14(6):609–618. (Cited on page 77.)
- Park, F. C. and Kim, M. W. (2000). Lie theory, Riemannian geometry, and the dynamics of coupled rigid bodies. *Zeitschrift für angewandte Mathematik und Physik*, 51(5):820–834. (Cited on page 77.)
- Park, F. C. and Ravani, B. (1997). Smooth invariant interpolation of rotations. *ACM Transactions on Graphics*, 16(3):277–295. (Cited on pages 24, 76, and 116.)
- Parker, J. A., Kenyon, R. V., and Troxel, D. (1983). Comparison of interpolating methods for image resampling. *IEEE Transactions on Medical Imaging*, 2(1):31–39. (Cited on page 29.)
- Paviour, D. C., Price, S. L., Jahanshahi, M., Lees, A. J., and Fox, N. C. (2006). Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy. *Brain*, 129(4):1040–1049. (Cited on page 17.)
- Peña-Casanova, J. (1997). Alzheimer’s Disease Assessment Scale—Cognitive in clinical practice. *International Psychogeriatrics*, 9(S1):105–114. (Cited on pages 15 and 147.)
- Pennanen, C., Kivipelto, M., Tuomainen, S., Hartikainen, P., Hänninen, T., Laakso, M. P., Hallikainen, M., Vanhanen, M., Nissinen, A., Helkala, E.-L., Vainio, P., Vanninen, R., Partanen, K., and Soininen, H. (2004). Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. *Neurobiology of Aging*, 25(3):303–310. (Cited on page 146.)
- Pennec, X. (1999). Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In Cetin, A. E., Akarun, L., Ertuzun, A., Gurcan, M. N., and Yardimci, Y., editors, *Nonlinear Signal and Image Processing*, volume 1, pages 194–198. IEEE-EURASIP. (Cited on page 136.)
- Pennec, X. (2006a). Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154. (Cited on pages 76, 85, 137, and 154.)
- Pennec, X. (2006b). Left-invariant Riemannian elasticity: a distance on shape diffeomorphisms? In *MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pages 1–13. (Cited on pages 55 and 173.)
- Pennec, X. (2013). Bi-invariant means on Lie groups with Cartan–Schouten connections. In Nielsen, F. and Barbaresco, F., editors, *Geometric Science of Information*, volume 8085 of *Lecture Notes in Computer Science*, pages 59–67. Springer Berlin Heidelberg. (Cited on pages 91, 136, and 161.)
- Pennec, X. and Arsigny, V. (2012). Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups. In Barbaresco, F., Mishra, A., and Nielsen, F., editors, *Matrix Information Geometry*, pages 123–166. Springer. (Cited on pages 90, 91, 136, and 137.)

- Pennec, X. and Ayache, N. (1998). Uniform distribution, distance and expectation problems for geometric features processing. *Journal of Mathematical Imaging and Vision*, 9(1):49–67. (Cited on pages 13 and 76.)
- Pennec, X. and Fillard, P. (2014). Statistical computing on non-linear spaces for computational anatomy. In Paragios, N., Duncan, J., and Ayache, N., editors, *Handbook of Biomedical Imaging: Methodologies and Clinical Research*. Springer. To Appear. (Cited on page 2.)
- Pennec, X., Fillard, P., and Ayache, N. (2006). A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66. (Cited on pages 140 and 170.)
- Pennec, X., Stefanescu, R., Arsigny, V., Fillard, P., and Ayache, N. (2005). Riemannian elasticity: A statistical regularization framework for non-linear registration. In Duncan, J. S. and Gerig, G., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, volume 3750 of *Lecture Notes in Computer Science*, pages 943–950. Springer Berlin Heidelberg. (Cited on page 55.)
- Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., and Hawkes, D. J. (1998). A comparison of similarity measures for use in 2D-3D medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–595. (Cited on page 77.)
- Perrot, M., Rivière, D., and Mangin, J.-F. (2011). Cortical sulci recognition and spatial normalization. *Medical Image Analysis*, 15(4):529–550. (Cited on page 28.)
- Perrot, M., Riviere, D., Tucholka, A., and Mangin, J.-F. (2009). Joint Bayesian cortical sulci recognition and spatial normalization. In Prince, J. L., Pham, D. L., and Myers, K. J., editors, *Information Processing in Medical Imaging*, volume 5636 of *Lecture Notes in Computer Science*, pages 176–187. Springer Berlin Heidelberg. (Cited on page 28.)
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology*, 56(3):303–308. (Cited on page 15.)
- Peyré, G. and Cohen, L. D. (2006). Geodesic remeshing using front propagation. *International Journal of Computer Vision*, 69(1):145–156. (Cited on pages 77 and 92.)
- Pham, M. T., Woodford, O. J., Perbet, F., Maki, A., Stenger, B., and Cipolla, R. (2011). A new distance for scale-invariant 3D shape recognition and registration. In *IEEE International Conference on Computer Vision*, pages 145–152. (Cited on page 24.)
- Pianykh, O. S. (2012). *Digital Imaging and Communications in Medicine (DICOM): a practical introduction and survival guide*. Springer. (Cited on page 3.)
- Pizer, S. M., Fletcher, P. T., Thall, A., Styner, M., Gerig, G., and Joshi, S. (2003). Object models in multiscale intrinsic coordinates via m-reps. *Image and Vision Computing*, 21(1):5–15. (Cited on page 17.)
- Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2003). Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004. (Cited on pages 10 and 53.)
- Porikli, F., Tuzel, O., and Meer, P. (2006). Covariance tracking using model update based on Lie algebra. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 728–735. (Cited on page 78.)

- Postelnicu, G., Zollei, L., and Fischl, B. (2009). Combined volumetric and surface registration. *IEEE Transactions on Medical Imaging*, 28(4):508–522. (Cited on page 32.)
- Postnikov, M. M. (1967). *The variational theory of geodesics*. Dover Publications. (Cited on page 90.)
- Postnikov, M. M. (2001). *Geometry VI: Riemannian geometry*. Springer. (Cited on pages 90 and 91.)
- Pozo, J. M., Villa-Uriol, M.-C., and Frangi, A. F. (2011). Efficient 3d geometric and zernike moments computation from unstructured surface meshes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(3):471–484. (Cited on page 149.)
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical recipes: the art of scientific computing*. Cambridge University Press. (Cited on pages 126 and 128.)
- Pripoae, G. T. (2006). Vector fields dynamics as geodesic motion on Lie groups. *Comptes Rendus Mathematique*, 342(11):865–868. (Cited on page 102.)
- Pruessner, J. C., Li, L. M., Serles, W., Pruessner, M., Collins, D., Kabani, N., Lupien, S., and Evans, A. (2000). Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cerebral Cortex*, 10(4):433–442. (Cited on page 17.)
- Querbes, O., Aubry, F., Pariente, J., Lotterie, J.-A., Démonet, J.-F., Duret, V., Puel, M., Berry, I., Fort, J.-C., and Celsis, P. (2009). Early diagnosis of Alzheimer’s disease using cortical thickness: impact of cognitive reserve. *Brain*, 132(Pt 8):2036–47. (Cited on page 146.)
- Rabbitt, R. D., Weiss, J. A., Christensen, G. E., and Miller, M. I. (1995). Mapping of hyperelastic deformable templates using the finite element method. In *International Symposium on Optical Science, Engineering, and Instrumentation*, pages 252–265. International Society for Optics and Photonics. (Cited on page 54.)
- Rahman, I. U., Drori, I., Stodden, V. C., Donoho, D. L., and Schröder, P. (2005). Multiscale representations for manifold-valued data. *Multiscale Modeling & Simulation*, 4(4):1201–1232. (Cited on pages 76, 78, and 83.)
- Ramani, S., Thévenaz, P., and Unser, M. (2010). Regularized interpolation for noisy images. *IEEE Transactions on Medical Imaging*, 29(2):543–558. (Cited on page 29.)
- Ramírez-Ruiz, B., Martí, M. J., Tolosa, E., Bartrés-Faz, D., Summerfield, C., Salgado-Pineda, P., Gómez-Ansón, B., and Junqué, C. (2005). Longitudinal evaluation of cerebral morphological changes in Parkinson’s disease with and without dementia. *Journal of Neurology*, 252:1345–1352. (Cited on page 18.)
- Rao, A., Aljabar, P., and Rueckert, D. (2008). Hierarchical statistical shape analysis and prediction of sub-cortical brain structures. *Medical Image Analysis*, 12(1):55–68. (Cited on page 147.)
- Ridgway, G. (2009). *Statistical analysis for longitudinal MR imaging of dementia*. PhD thesis, University College London. (Cited on pages 167, 170, and 177.)
- Roche, A., Malandain, G., and Ayache, N. (2000). Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology*, 11(1):71–80. (Cited on pages 53 and 68.)

- Roche, A., Pennec, X., Malandain, G., and Ayache, N. (2001). Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049. (Cited on pages 53 and 77.)
- Rohlfing, T. (2012). Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE Transactions on Medical Imaging*, 31(2):153–163. (Cited on pages 7, 53, 54, and 172.)
- Rohlfing, T. and Avants, B. (2012). “Nonparametric local smoothing” is not image registration. *BMC Research Notes*, 5(1):610. (Cited on pages 7, 9, 51, and 54.)
- Röntgen, W. C. (1895). Über eine neue art von strahlen. *sitzungsberichte der physik.-med. gesellschaft zu würzburg. Annalen der Physik und Chemie*, 64(1898):132–141. (Cited on page 2.)
- Rosman, G., Wang, Y., Tai, X.-C., Kimmel, R., and Bruckstein, A. M. (2012). Fast regularization of matrix-valued images. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., and Schmid, C., editors, *Computer Vision–ECCV 2012*, volume 7574 of *Lecture Notes in Computer Science*, pages 173–186. Springer Berlin Heidelberg. (Cited on page 76.)
- Rossmann, W. (2002). *Lie groups: an introduction through linear groups*. Oxford University Press. (Cited on page 96.)
- Rouchdy, Y., Pousin, J., Schaerer, J., and Clarysse, P. (2007). A hyperelastic deformable template for cardiac segmentation in MRI. In Sachse, F. B. and Seemann, G., editors, *Functional Imaging and Modeling of the Heart*, volume 4466 of *Lecture Notes in Computer Science*, pages 443–452. Springer Berlin Heidelberg. (Cited on page 54.)
- Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326. (Cited on page 136.)
- Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., and Hammers, A. (2006). Diffeomorphic registration using B-splines. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, volume 4191 of *Lecture Notes in Computer Science*, pages 702–709. Springer Berlin Heidelberg. (Cited on page 43.)
- Rueckert, D., Frangi, A. F., and Schnabel, J. A. (2003). Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, 22(8):1014–1025. (Cited on page 160.)
- Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., and Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721. (Cited on page 43.)
- Rüsch, N., Elst, L., Ludaescher, P., Wilke, M., Huppertz, H.-J., Thiel, T., Schmahl, C., Bohus, M., Lieb, K., Hesslinger, B., Henning, J., and Ebert, D. (2003). A voxel-based morphometric MRI study in female patients with borderline personality disorder. *NeuroImage*, 20(1):385–392. (Cited on page 18.)
- Sabuncu, M. R., Yeo, B. T. T., Van Leemput, K., Vercauteren, T., and Golland, P. (2009). Asymmetric image-template registration. In Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., and Taylor, C., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, volume 5761 of *Lecture Notes in Computer Science*, pages 565–573. Springer Berlin Heidelberg. (Cited on page 61.)

- Saccon, A., Hauser, J., and Aguiar, A. P. (2013). Optimal control on Lie groups: the projection operator approach. *IEEE Transactions on Automatic Control*, 58(9):2230–2245. (Cited on pages 90 and 91.)
- Sachkov, Y. L. (2009). Control theory on Lie groups. *Journal of Mathematical Sciences*, 156(3):381–439. (Cited on page 103.)
- Said, S., Courty, N., Le Bihan, N., and Sangwine, S. J. (2007). Exact principal geodesic analysis for data on SO(3). In *European Signal Processing Conference*, pages 1700–1705. EURASIP. (Cited on page 76.)
- Sattinger, D. H. and Weaver, O. L. (1986). *Lie groups and algebras with applications to physics, geometry, and mechanics*. Springer-Verlag. (Cited on pages 81 and 99.)
- Schneider, P. and Eberly, D. H. (2002). *Geometric tools for computer graphics*. The Morgan Kaufmann Series in Computer Graphics. Elsevier Science. (Cited on page 36.)
- Scholz, E. (1999). *History of Topology*, chapter 2. The concept of manifold, 1850–1950, pages 25–64. Elsevier Science. (Cited on page 79.)
- Schott, J. M., Price, S. L., Frost, C., Whitwell, J. L., Rossor, M. N., and Fox, N. C. (2005). Measuring atrophy in Alzheimer disease. A serial MRI study over 6 and 12 months. *Neurology*, 65(1):119–124. (Cited on page 17.)
- Schreiner, J., Scheiclegger, C. E., and Silva, C. T. (2006). High-quality extraction of isosurfaces from regular and irregular grids. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):1205–1212. (Cited on page 27.)
- Schwartzman, A. (2006). *Random ellipsoids and false discovery rates: statistics for diffusion tensor imaging data*. PhD thesis, Stanford University. (Cited on page 170.)
- Seong, J.-K., Jeong, W.-K., and Cohen, E. (2008). Anisotropic geodesic distance computation for parametric surfaces. In *IEEE International Conference on Shape Modeling and Applications*, pages 179–186. (Cited on page 77.)
- Sethian, J. (1999). Fast marching methods. *SIAM Review*, 41(2):199–235. (Cited on pages 77 and 92.)
- Shen, D. and Davatzikos, C. (2002). HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, 21(11):1421–1439. (Cited on page 53.)
- Shoemake, K. (1985). Animating rotation with quaternion curves. *SIGGRAPH Computer Graphics*, 19(3):245–254. (Cited on pages 24, 40, 76, and 133.)
- Sifri, O., Sheffer, A., and Gotsman, C. (2003). Geodesic-based surface remeshing. In *Proceedings 12th International Meshing Roundtable*. (Cited on page 77.)
- Skerl, D., Likar, B., and Pernus, F. (2006). A protocol for evaluation of similarity measures for rigid registration. *IEEE Transactions on Medical Imaging*, 25(6):779–791. (Cited on page 77.)
- Škrinjar, O., Bistoquet, A., and Tagare, H. (2008). Symmetric and transitive registration of image sequences. *Journal of Biomedical Imaging*, 2008:14. (Cited on pages 171 and 172.)

- Sled, J. G., Zijdenbos, A. P., and Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Transactions on Medical Imaging*, 17(1):87–97. (Cited on page 53.)
- Small, C. (1996). *The statistical theory of shape*. Springer. (Cited on pages 17 and 28.)
- Small, G. W. (2002). *Neuropsychopharmacology: The Fifth Generation of Progress: an Official Publication of the American College of Neuropsychopharmacology*, chapter 86. Structural and Functional Brain Imaging Of Alzheimer Disease, pages 1231–1242. M - Medicine Series. Lippincott Williams & Wilkins. (Cited on page 16.)
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, Supplement 1:S208–19. (Cited on page 148.)
- Smith, S. T. (1994). Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, 3(3):113–135. (Cited on page 77.)
- Smith, W. A. P. (2012). Representing, storing and visualizing 3d data. In Pears, N., Liu, Y., and Bunting, P., editors, *3D Imaging, Analysis and Applications*, pages 139–182. Springer London. (Cited on page 27.)
- Snedecor, G. W. and Cochran, W. G. (1991). *Statistical methods*. Wiley. (Cited on page 164.)
- Snell, R. S. (2007). *Neuroanatomia Clinica/Clinical Neuroanatomy*. Ed. Médica Panamericana. (Cited on page 5.)
- Sommer, S., Lauze, F., Hauberg, S., and Nielsen, M. (2010). Manifold valued statistics, exact principal geodesic analysis and the effect of linear approximations. In Daniilidis, K., Maragos, P., and Paragios, N., editors, *Computer Vision–ECCV 2010*, volume 6316 of *Lecture Notes in Computer Science*, pages 43–56. Springer Berlin Heidelberg. (Cited on page 136.)
- Sommer, S., Lauze, F., and Nielsen, M. (2013). Optimization over geodesics for exact principal geodesic analysis. *Advances in Computational Mathematics*, pages 1–31. (Cited on pages 76, 85, and 108.)
- Sotiras, A., Davatzikos, C., and Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, 32(7):1153–1190. (Cited on pages 7, 34, 51, and 54.)
- Stahl, S. (2007). *A gateway to modern geometry: the Poincaré half-plane*. Jones and Bartlett Publishers. (Cited on pages 115 and 143.)
- Storey, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*, pages 2013–2035. (Cited on page 168.)
- Studholme, C., Cardenas, V., Blumenfeld, R., Schuff, N., Rosen, H. J., Miller, B., and Weiner, M. (2004). Deformation tensor morphometry of semantic dementia with quantitative validation. *NeuroImage*, 21(4):1387–1398. (Cited on pages 146 and 161.)
- Studholme, C., Cardenas, V., Schuff, N., Rosen, H., Miller, B., and Weiner, M. (2001). Detecting spatially consistent structural differences in Alzheimer’s and fronto temporal dementia using

- deformation morphometry. In Niessen, W. J. and Viergever, M. A., editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, volume 2208 of *Lecture Notes in Computer Science*, pages 41–48. Springer Berlin Heidelberg. (Cited on page 146.)
- Studholme, C., Drapaca, C., Iordanova, B., and Cardenas, V. (2006). Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE Transactions on Medical Imaging*, 25(5):626–639. (Cited on page 164.)
- Styner, M., Gerig, G., Lieberman, J., Jones, D., and Weinberger, D. (2003). Statistical shape analysis of neuroanatomical structures based on medial models. *Medical Image Analysis*, 7(3):207–20. (Cited on page 146.)
- Styner, M., Gorczowski, K., Fletcher, T., Jeong, J. Y., Pizer, S. M., and Gerig, G. (2006a). Statistics of pose and shape in multi-object complexes using principal geodesic analysis. In Yang, G.-Z., Jiang, T. Z., Shen, D., Gu, L., and Yang, J., editors, *Medical Imaging and Augmented Reality*, volume 4091 of *Lecture Notes in Computer Science*, pages 1–8. Springer Berlin Heidelberg. (Cited on page 147.)
- Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J. J., Shenton, M. E., and Gerig, G. (2006b). Framework for the statistical shape analysis of brain structures using SPHARM-PDM. *The Insight Journal*, (1071):242–250. (Cited on page 168.)
- Styner, M., Oguz, I., Xu, S., Pantazis, D., and Gerig, G. (2007). Statistical group differences in anatomical shape analysis using hotelling T2 metric. In *Medical Imaging*, pages 65123Z–65123Z. International Society for Optics and Photonics. (Cited on page 168.)
- Subbarao, R. (2008). *Robust Statistics Over Riemannian Manifolds for Computer Vision*. PhD thesis, State University of New Jersey. (Cited on page 136.)
- Subbarao, R. and Meer, P. (2009). Nonlinear mean shift over Riemannian manifolds. *International Journal of Computer Vision*, 84(1):1–20. (Cited on pages 78 and 85.)
- Sun, W., Niessen, W. J., van Stralen, M., and Klein, S. (2013). Simultaneous multiresolution strategies for nonrigid image registration. *IEEE Transactions on Image Processing*, 22(12):4905–4917. (Cited on page 60.)
- Surazhsky, V., Surazhsky, T., Kiryanov, D., Gortler, S. J., and Hoppe, H. (2005). Fast exact and approximate geodesics on meshes. *ACM Transactions on Graphics*, 24(3):553–560. (Cited on page 77.)
- Symms, M., Jäger, H. R., Schmierer, K., and Yousry, T. A. (2004). A review of structural magnetic resonance neuroimaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(9):1235–1244. (Cited on page 5.)
- Székely, G. J. and Rizzo, M. L. (2004). Testing for equal distributions in high dimension. *InterStat*, (Nov). (Cited on page 177.)
- Szeliski, R. (2006). Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2(1):1–104. (Cited on page 51.)
- Talairach, J. and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging*. Thieme. (Cited on pages 5 and 7.)

- Tao, X., Prince, J. L., and Davatzikos, C. (2002). Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. *IEEE Transactions on Medical Imaging*, 21(5):513–524. (Cited on page 24.)
- Tapp, K. (2005). *Matrix groups for undergraduates*. AMS. (Cited on pages 36, 76, 78, 79, and 117.)
- Tenenbaum, J. B. B., Silva, V. d., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323. (Cited on page 136.)
- Thévenaz, P., Blu, T., and Unser, M. (2000). Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758. (Cited on page 29.)
- Thompson, D. W. (1917). *On Growth and Form*. Cambridge University Press. (Cited on pages 4, 24, and 28.)
- Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., and Toga, A. W. (2000a). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*, 404(6774):190–193. (Cited on page 18.)
- Thompson, P. M., Hayashi, K. M., Dutton, R. A., Chiang, M.-C., Leow, A. D., Sowell, E. R., De Zubicaray, G., Becker, J. T., Lopez, O. L., Aizenstein, H. J., and Toga, A. W. (2007). Tracking Alzheimer’s disease. *Annals of the New York Academy of Sciences*, 1097(310):183–214. (Cited on page 146.)
- Thompson, P. M. and Toga, A. W. (1997). Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. *Medical Image Analysis*, 1(4):271–294. (Cited on page 171.)
- Thompson, P. M., Woods, R. P., Mega, M. S., and Toga, A. W. (2000b). Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. *Human Brain Mapping*, 9(2):81–92. (Cited on page 160.)
- Trouvé, A. (1995). An infinite dimensional group approach for physics based models in patterns recognition. Unpublished, accessed from: http://www.cis.jhu.edu/publications/papers_in_database/alain/trouve1995.pdf. (Cited on page 44.)
- Trouvé, A. (1998). Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221. (Cited on page 55.)
- Tuzel, O., Porikli, F., and Meer, P. (2008). Learning on Lie groups for invariant detection and tracking. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8. (Cited on pages 77 and 78.)
- Tuzel, O., Subbarao, R., and Meer, P. (2005). Simultaneous multiple 3D motion estimation via mode finding on Lie groups. In *IEEE International Conference on Computer Vision*, volume 1, pages 18–25. (Cited on page 78.)
- Twining, C. J. and Marsland, S. (2008). Constructing an atlas for the diffeomorphism group of a compact manifold with boundary, with application to the analysis of image registrations. *Journal of Computational and Applied Mathematics*, 222(2):411–428. (Cited on page 66.)
- Umeyama, S. (1991). Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380. (Cited on page 56.)

- Unser, M., Aldroubi, A., and Eden, M. (1993). B-spline signal processing. i. theory. *IEEE Transactions on Signal Processing*, 41(2):821–833. (Cited on page 29.)
- Vaccaro, C. (2012). Heat kernel methods in finance: the SABR model. *arXiv preprint*. (Cited on pages 95 and 115.)
- Vaillant, M. and Glaunès, J. (2005). Surface matching via currents. In Christensen, G. E. and Sonka, M., editors, *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 381–392. Springer Berlin Heidelberg. (Cited on page 55.)
- Vaillant, M., Miller, M. I., Younes, L., and Trouvé, A. (2004). Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23, Supplement 1:S161–S169. (Cited on pages 55 and 136.)
- Van Brunt, B. (2004). *The calculus of variations*. Springer New York. (Cited on page 94.)
- Van den Elsen, P. A., Maintz, J. B. A., Pol, E.-J. D., and Viergever, M. A. (1995). Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Transactions on Medical Imaging*, 14(2):384–396. (Cited on page 53.)
- Vandereycken, B. (2010). *Riemannian and multilevel optimization for rank-constrained matrix problems*. PhD thesis, Department of Computer Science, KU Leuven. (Cited on page 136.)
- Vandereycken, B., Absil, P.-A., and Vandewalle, S. (2013). A Riemannian geometry with complete geodesics for the set of positive semidefinite matrices of fixed rank. *IMA Journal of Numerical Analysis*, 33:481–514. (Cited on pages 102 and 117.)
- Vargas, J. R. (2002). *Anatomía topográfica*. UACJ. (Cited on pages 2 and 5.)
- Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2008). Symmetric log-domain diffeomorphic registration: A demons-based approach. In Metaxas, D., Axel, L., Fichtinger, G., and Székely, G., editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2008*, volume 5241 of *Lecture Notes in Computer Science*, pages 754–761. Springer Berlin Heidelberg. (Cited on pages 55 and 71.)
- Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2009). Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage*, 45(1):S61–S72. (Cited on page 55.)
- Verma, N. (2013). Distance preserving embeddings for general n-dimensional manifolds. *Journal of Machine Learning Research*, 14(1):2415–2448. (Cited on page 92.)
- Vesalius, A. (1543). *Andreae Vesalii Bruxellensis, scholae medicorum Patauinae professoris De humani corporis fabrica libri septem*. Basileae: Ex officina Joannis Oporini. (Cited on page 2.)
- Vince, J. (2011a). *Quaternions for computer graphics*. Springer. (Cited on page 40.)
- Vince, J. (2011b). *Rotation Transforms for Computer Graphics*. Springer. (Cited on pages 36 and 40.)
- Viola, P. and Wells III, W. M. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154. (Cited on page 53.)

- Vogeley, K., Hobson, T., Schneider-Axmann, T., Honer, W. G., Bogerts, B., and Falkai, P. (1998). Compartmental volumetry of the superior temporal gyrus reveals sex differences in schizophrenia—a post-mortem study. *Schizophrenia Research*, 31(2):83–87. (Cited on page 17.)
- Vong, S. W. and Jin, X. Q. (2008). Proof of Böttcher and Wenzel’s conjecture. *Operators and Matrices*, 2(3):435–442. (Cited on page 124.)
- W., R. (1997). An exponential formula for polynomial vector fields. *Advances in Mathematics*, 128(1):190–216. (Cited on page 45.)
- Wahba, G. (1990). *Spline models for observational data*. SIAM. (Cited on page 29.)
- Wang, H.-C. (1969). Discrete nilpotent subgroups of Lie groups. *Journal of Differential Geometry*, 3(3-4):481–492. (Cited on page 117.)
- Wang, Q. and Ward, R. K. (2007). A new orientation-adaptive interpolation method. *IEEE Transactions on Image Processing*, 16(4):889–900. (Cited on page 29.)
- Warfield, S. K., Zou, K. H., and Wells, W. M. (2006). Validation of image segmentation by estimating rater bias and variance. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, pages 839–847. Springer. (Cited on page 53.)
- Wein, W., Brunke, S., Khamene, A., Callstrom, M. R., and Navab, N. (2008). Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis*, 12(5):577–585. (Cited on page 53.)
- Weishaupt, D., Köchli, V. D., and Marincek, B. (2003). *How does MRI work?: an introduction to the physics and function of magnetic resonance imaging*. Springer. (Cited on page 6.)
- Whitcher, B., Wisco, J. J., Hadjikhani, N., and Tuch, D. S. (2007). Statistical group comparison of diffusion tensors via multivariate hypothesis testing. *Magnetic Resonance in Medicine*, 57(6):1065–1074. (Cited on page 177.)
- Whitwell, J. L., Jack, C. R., Parisi, J. E., Knopman, D. S., Boeve, B. F., Petersen, R. C., Ferman, T. J., Dickson, D. W., and Josephs, K. A. (2007). Rates of cerebral atrophy differ in different degenerative pathologies. *Brain*, 130(4):1148–1158. (Cited on page 16.)
- Wild, E. J. and Fox, N. C. (2009). Serial volumetric MRI in Parkinsonian disorders. *Movement Disorders*, 24(S2):S691–S698. (Cited on page 17.)
- Woods, R. P. (2003). Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. *NeuroImage*, 18(3):769–788. (Cited on pages 40 and 86.)
- Woods, R. P., Grafton, S. T., Watson, J. D. G., Sicotte, N. L., and Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22(1):153–165. (Cited on page 34.)
- Woolson, R. F. and Clarke, W. R. (2011). *Statistical Methods for the Analysis of Biomedical Data*. Wiley Series in Probability and Statistics. Wiley. (Cited on page 164.)
- Wüstner, M. (2003). A connected Lie group equals the square of the exponential image. *Journal of Lie Theory*, 13(1):307–309. (Cited on page 117.)

- Xue, Z., Shen, D., Karacali, B., Stern, J., Rottenberg, D., and Davatzikos, C. (2006). Simulating deformations of MR brain images for validation of atlas-based segmentation and registration algorithms. *NeuroImage*, 33(3):855–866. (Cited on page 160.)
- Yaglom, I. M., Grant, H., and Shenitzer, A. (1988). *Felix Klein and Sophus Lie: evolution of the idea of symmetry in the nineteenth century*. Birkhäuser. (Cited on pages 13 and 150.)
- Yanovsky, I. (2008). *Unbiased Nonlinear Image Registration*. PhD thesis, University of California. (Cited on page 58.)
- Yanovsky, I., Le Guyader, C., Leow, A., Toga, A. W., Thompson, P. M., and Vese, L. (2008). Unbiased volumetric registration via nonlinear elastic regularization. In *MICCAI Workshop on Mathematical Foundations of Computational Anatomy*. (Cited on page 55.)
- Yavari, A., Marsden, J. E., and Ortiz, M. (2006). On spatial and material covariant balance laws in elasticity. *Journal of Mathematical Physics*, 47(4):042903–1–042903–53. (Cited on page 173.)
- Yeo, B. T. T., Sabuncu, M. R., Vercauteren, T., Ayache, N., Fischl, B., and Golland, P. (2010). Spherical demons: Fast diffeomorphic landmark-free surface registration. *IEEE Transactions on Medical Imaging*, 29(3):650–668. (Cited on page 24.)
- Younes, L. (1998). Computable elastic distances between shapes. *SIAM Journal on Applied Mathematics*, 58(2):565–586. (Cited on page 77.)
- Younes, L. (2010). *Shapes and diffeomorphisms*. Springer-Verlag. (Cited on pages 44 and 55.)
- Younes, L., Arrate, F., and Miller, M. I. (2009). Evolutions equations in computational anatomy. *NeuroImage*, 45(1):S40–S50. (Cited on page 31.)
- Zach, C., Pock, T., and Bischof, H. (2007). A duality based approach for realtime TV-L1 optical flow. In *Pattern Recognition*, pages 214–223. Springer-Verlag. (Cited on page 55.)
- Zacur, E., Bossa, M., and Olmos, S. (2013). Multivariate tensor-based morphometry with a right-invariant Riemannian distance on $GL+(n)$. *Journal of Mathematical Imaging and Vision*, doi:10.1007/s10851-013-0479-7. (Cited on pages 13, 76, and 128.)
- Zannis, V. I., Just, P. W., and Breslow, J. L. (1981). Human apolipoprotein E isoprotein subclasses are genetically determined. *American Journal of Human Genetics*, 33(1):11. (Cited on page 15.)
- Žefran, M. and Kumar, V. (1998). Interpolation schemes for rigid body motions. *Computer-Aided Design*, 30(3):179–189. (Cited on page 77.)
- Žefran, M., Kumar, V., and Croke, C. (1999). Metrics and connections for rigid-body kinematics. *International Journal of Robotics Research*, 18(2):242. (Cited on pages 77 and 88.)
- Žefran, M., Kumar, V., and Croke, C. B. (1998). On the generation of smooth three-dimensional rigid body motions. *IEEE Transactions on Robotics and Automation*, 14(4):576–589. (Cited on page 77.)
- Zhang, S. (2005a). Invertible Jacobian for hexahedral finite elements. part 1. bijectivity. Unpublished, accessed from: <http://www.math.udel.edu/~s Zhang/research/p/bijective1.ps>. (Cited on page 30.)

- Zhang, S. (2005b). Invertible Jacobian for hexahedral finite elements. part 2. global positivity. Unpublished, accessed from: <http://www.math.udel.edu/~szhang/research/p/bijective2.ps>. (Cited on page 30.)
- Zhou, H., Zheng, J., and Yang, X. (2012). Euler arc splines for curve completion. *Computers & Graphics*, 36(6):642–650. (Cited on page 27.)
- Zitova, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000. (Cited on page 51.)
- Zou, G., Hua, J., and Muzik, O. (2007). Non-rigid surface registration using spherical thin-plate splines. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 367–374. Springer. (Cited on page 24.)