

Statistical analysis of organs' shapes and deformations: the Riemannian and the affine settings in computational anatomy

Xavier Pennec

Abstract Computational anatomy is an emerging discipline at the interface of geometry, statistics and medicine that aims at analyzing and modeling the biological variability of organs' shapes at the population level. Shapes are equivalence classes of images, surfaces or deformations of a template under rigid body (or more general) transformations. Thus, they belong to non-linear manifolds. In order to deal with multiple samples in non-linear spaces, a consistent statistical framework on Riemannian manifolds has been designed over the last decade. We detail in this chapter the extension of this framework to Lie groups endowed with the affine symmetric connection, a more invariant (and thus more consistent) but non-metric structure on transformation groups. This theory provides strong theoretical bases for the use of one-parameter subgroups and diffeomorphisms parametrized by stationary velocity fields (SVF), for which efficient image registration methods like log-Demons have been developed with a great success from the practical point of view. One can further reduce the complexity with locally affine transformations, leading to parametric diffeomorphisms of low dimension encoding the major shape variability. We illustrate the methodology with the modeling of the evolution of the brain with Alzheimer's disease and the analysis of the cardiac motion from MRI sequences of images.

1 Introduction

At the interface of geometry, statistics, image analysis and medicine, computational anatomy aims at analyzing and modeling the biological variability of the organs' shapes and their dynamics at the population level. The goal is to model the mean anatomy, its normal variation, its motion / evolution and to

Xavier Pennec
Université Côte d'Azur and Inria, France, e-mail: xavier.pennec@inria.fr

discover morphological differences between normal and pathological groups. For instance, the analysis of population-wise structural brain changes with aging in Alzheimer’s disease requires first the analysis of longitudinal morphological changes for a specific subject, which can be done using non-linear registration-based regression, followed by a longitudinal group-wise analysis where the subject-specific longitudinal trajectories are transported in a common reference [32, 20]. **In both steps, it is desirable that the longitudinal and the inter-subject transformations smoothly preserve the spatial organization of the anatomical tissues by avoiding intersections, foldings or tearing. Simply encoding deformations with a vector space of displacement fields is not sufficient to preserve the topology: one needs to require diffeomorphic transformations (differentiable one-to-one transformations with differentiable inverse). Space of diffeomorphisms are examples of infinite dimensional manifolds. Informally, manifolds are spaces that locally (but not globally) resemble a given Euclidean space. The simplest example is the sphere or the earth surface which looks locally flat at a scale which is far below the curvature radius but exhibit curvature and a non-linear behaviour at larger scales.**

Likewise, shape analysis most often relies on the identification of features describing locally the anatomy such as landmarks, curves, surfaces, intensity patches, full images, etc. Modeling their statistical distribution in the population requires to first identify point-to-point anatomical correspondences between these geometric features across subjects. **This may be feasible for landmark points, but not for curves or surfaces. Thus, one generally considers relabelled point-sets or reparametrized curve/surface/image as equivalent objects. With this geometric formulation, shapes spaces are the quotient the original space of features by their reparametrization group.** One also often wants to remove a global rigid or affine transformation. One considers in this case the equivalence classes of images, surfaces or deformations under the action of this space transformation group, and shape spaces are once again quotient spaces. Unfortunately, even if we start from features belonging to a nice Euclidean space, taking the quotient generally endows the shape space with a non-linear manifold structure. For instance, equivalence classes of k -tuples of points under rigid or similarity transformations result in non-linear Kendall’s shape spaces (see e.g. [15] for a recent account on that subject). The quotient of curves, surfaces and higher dimensional objects by their reparametrizations (diffeomorphisms of their domains) produces in general even more complex infinite dimensional shape spaces [6].

Thus, shapes and deformations belong in general to non-linear manifolds, while statistics were essentially developed for linear and Euclidean spaces. For instance, adding or subtracting two curves does not really make sense. It is thus not easy to average several shapes. Likewise, averaging unit vectors (resp. rotation matrices) do not lead to a unit vector (resp. a rotation matrix).

It is thus necessary to define a consistent statistical framework on manifolds and Lie groups. This has motivated the development of Geometric Statistics during the last decade [52]. We summarize below the main features of the theory of statistics on manifolds, before generalizing it in the next section to more general affine connection spaces.

1.1 Riemannian manifolds

While being non-linear, manifolds are locally Euclidean, and an infinitesimal measure of the distance (a metric) allows to endow them with a Riemannian manifold structure. More formally, a Riemannian metric on a manifold \mathcal{M} is a continuous collection of scalar products on the tangent space $T_x\mathcal{M}$ at each point x of the manifold. The metric measures the dot product of two infinitesimal vectors at a point of our space: this allows to measure directions and angles in the tangent space. One can also measure the length of a curve on our manifold by integrating the norm of its tangent vector. **The minimal length among all the curves joining two given points defines the intrinsic distance between these two points. The curves realizing these shortest paths are called geodesics, generalizing the geometry of our usual flat 3D space to curved spaces among which the flat torus, the sphere and the hyperbolic space are the simplest examples.**

The calculus of variations shows that geodesics are the solutions of a system of second order differential equations depending on the Riemannian metric. Thus, the geodesic curve $\gamma_{(x,v)}(t)$ starting at a given point x with a given tangent vector $v \in T_x\mathcal{M}$ always exists for some short time. When the time-domain of all geodesics can be extended to infinity, the manifold is said to be geodesically complete. This means that the manifold has no boundary nor any singular point that we can reach in a finite time. As an important consequence, the Hopf-Rinow-De Rham theorem states that there always exists at least one minimizing geodesic between any two points of the manifold (i.e. whose length is the distance between the two points) [12]. Henceforth, we implicitly assume that all Riemannian manifolds are geodesically complete.

The function $\exp_x(v) = \gamma_{(x,v)}(1)$ mapping the tangent space $T_x\mathcal{M}$ at x to the manifold \mathcal{M} is called the exponential map at the point x . It is defined on the whole tangent space but it is diffeomorphic only locally. Its inverse $\log_x(y)$ is a vector rooted at x . It maps each point y of a neighborhood of x to the shortest tangent vector that allows to join x to y geodesically. The maximal definition domain of the log is called the injectivity domain. It covers all the manifold except a set of null measure called the cut-locus of the point. For statistical purposes, we can thus safely neglect this set in many cases. The Exp and Log maps \exp_x and \log_x are defined at any point x of the manifold (x is called the foot-point in differential geometry). They realize a continuous family of very convenient charts of the manifold where geodesics

starting from the foot-point are straight lines, and along which the distance to the foot-point is conserved. These charts are somehow the “most linear” chart of the manifold with respect to their foot-point (Fig.1)

In practice, we can identify a tangent vector $v \in T_x\mathcal{M}$ within the injectivity domain to the end-points of the geodesic segment $[x, y = \exp_x(v)]$ thanks to the exponential maps. Conversely, almost any bi-point (x, y) on the manifold where y is not in the cut-locus of x can be mapped to the vector $\vec{xy} = \log_x(y) \in T_x\mathcal{M}$ by the log map. In a Euclidean space, we would write $\exp_x(v) = x + v$ and $\log_x(y) = y - x$. This reinterpretation of addition and subtraction using logarithmic and exponential maps is very powerful to generalize algorithms working on vector spaces to algorithms on Riemannian manifolds. It is also very powerful in terms of implementation since we can express many of the geometric operations in these terms: the implementation of the exp and log maps at each point is thus the basis of programming on Riemannian manifolds.

1.2 Statistics on Riemannian manifolds

The Riemannian metric induces an infinitesimal volume element on each tangent space, denoted $d\mathcal{M}$, that can be used to measure random events on the manifold and to define intrinsic probability density functions (pdf). It is worth noticing that the measure $d\mathcal{M}$ represents the notion of uniformity according to the chosen Riemannian metric. With the probability measure of a random element, we can integrate functions from the manifold to any vector space, thus defining the expected value of this function. However, we

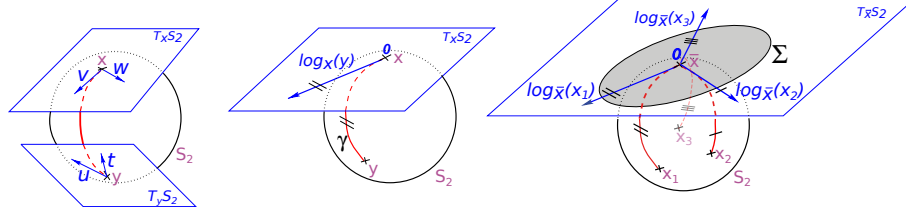


Fig. 1 Riemannian geometry and statistics on the sphere. Left: The tangent planes at points x and y of the sphere S_2 are different: the tangent vectors v and w at the point x cannot be compared to the vectors t and u that tangent at the point y . Thus, it is natural to define the scalar product on each tangent plane. Middle: Geodesics starting at x are straight lines in a normal coordinate system at x and the distance is conserved up to the cut-locus. Right: the Fréchet mean \bar{x} is the point minimizing the mean squared Riemannian distance to the data points. It corresponds to the point for which the development of the geodesics to the data points on the tangent space is optimally centered (the mean $\sum_i \log_{\bar{x}}(x_i) = 0$ in that tangent space is zero). The covariance matrix is then defined in that tangent space. Figure adapted from [45].

generally cannot integrate manifold-valued functions since an integral is a linear operator. Thus, one cannot define the mean or expected “value” of a random manifold element using a weighted sum or an integral as usual.

The main solution to this problem is to redefine the mean as the minimizer of an intrinsic quantity: the Fréchet (resp. Karcher) mean minimizes globally (resp. locally) the sum of squared Riemannian distance to our samples. As the mean is now defined through a minimization procedure, its existence and uniqueness may be questioned. In practice, one mean value almost always exists, and it is unique as soon as the distribution is sufficiently peaked. The properties of the mean are very similar to those of the modes of a distribution in the Euclidean case. The Fréchet mean was used since the 1990s in medical image analysis for redefining simple statistical methods on Riemannian manifolds [43, 49, 44, 45, 17].

To compute the Fréchet mean, one can follow the Riemannian gradient of the variance with an iteration of the type:

$$\bar{x}_{t+1} = \exp_{\bar{x}_t} \left(\alpha \frac{1}{n} \sum_i \log_{\bar{x}_t}(x_i) \right).$$

The algorithm essentially alternates the computation of the tangent mean in the tangent space at the current estimation of the mean, and a geodesic marching step towards the computed tangent mean. The value $\alpha = 1$ corresponding to a Gauss-Newton scheme is usually working very well, although there are examples where it should be reduced due to the curvature of the space. An adaptive time-step in the spirit of Levenberg-Marquardt is easily solving this problem.

When the Fréchet mean is determined, one can pull back our distribution of data points on the tangent space at the mean to define higher order moments like the covariance matrix $\Sigma = \frac{1}{n} \sum_{i=1}^n \log_{\bar{x}}(x_i) \log_{\bar{x}}(x_i)^T$. Seen for the most central point (the Fréchet mean), we have somehow corrected the non-linearity of our Riemannian manifold. Based on this mean \bar{x} and this covariance matrix Σ , we can define the Mahalanobis distance in the tangent space by:

$$\mu_{(\bar{x}, \Sigma)}^2(y) = \log_{\bar{x}}(y)^T \Sigma^{(-1)} \log_{\bar{x}}(y).$$

It is worth noticing that the expected Mahalanobis distance of a random point is independent of the distribution and is equal to the dimension of the manifold when its mean and covariance are known, as in the vector case [43, 45]. A very simple extension of Principle Component Analysis (PCA) consists in diagonalizing the covariance matrix Σ and defining the modes using the eigenvectors of decreasing eigenvalues in the tangent space at the mean. This method usually works very well for sufficiently concentrated data. More complex methods like Principal Geodesic Analysis (PGA), geodesic PCA or Barycentric Subspace Analysis (BSA) may be investigated for data distributions with a larger support [46].

A notion of Gaussian may also be defined on a manifold by choosing the distribution that minimizes the entropy knowing the mean and the covariance. It was shown in [43, 45] that this amounts to consider a truncated Gaussian distribution on the tangent space at the mean point which only covers the injectivity domain (i.e. truncated at the tangential cut locus): the pdf (with respect to the Riemannian measure) is $N_{(\bar{x}, \Sigma)}(y) = Z(\bar{x}, \Sigma) \exp(-\frac{1}{2} \log_{\bar{x}}(y)^T \Gamma \log_{\bar{x}}(y))$. However, we should be careful that the relation between the concentration matrix Γ and the covariance matrix Σ is more complex than the simple inversion of the Euclidean case since it has to be corrected for the curvature of the manifold.

Based on this truncated Gaussian distribution, one can generalize the multivariate Hotelling T-squared test using the Mahalanobis distance. When the distribution is Gaussian with a known mean and covariance matrix, the law generalize the χ^2 law and [45] showed that it has the same density as in the vector case up to order 3. This opens the way to the generalization of many other statistical tests, as we should obtain similarly simple approximations for sufficiently centered distributions.

Notice that the reformulation of the (weighted) mean as an intrinsic minimization problem allows to extend quite a number of other image processing algorithms to manifold-valued signal and images, like interpolation, diffusion and restoration of missing data (extrapolation). This is the case for instance of diffusion tensor imaging for which manifold-valued image processing was pioneered in [50].

2 An affine symmetric space structure for Lie groups

A classical way to perform statistics on shapes in computational anatomy is to estimate or assume a template shape and then to encode other shapes by diffeomorphic transformations of that template. This lifts the problem from statistics on manifolds to statistics on smooth transformation groups, i.e. Lie groups. The classical Riemannian methodology consists in endowing the Lie group with a left (or right) invariant metric which turns the transformation group into a Riemannian manifold. This means that the metric at a point x of the group is obtained by the left translation $L_x(y) = x \circ y$ of the metric at identity, or in a more computational way, that the scalar product of two tangent vectors at x is obtained by left-translating them back to identity using $DL_{x^{-1}}$ and taking the scalar product there. A right-invariant metric is obtained if we use the differential of the right translation $R_x(y) = y \circ x$ to identify the tangent space at x to the tangent space at identity. However, this Riemannian approach is consistent with the inversion operations of the group only if the metric is both left and right invariant. This is the case for compact or commutative groups, such as rotations. But as soon as the Lie group is a non direct product of simpler compact or commutative ones, such as rigid

body transformations in 2D or 3D, there does not exist a bi-invariant metric: left-invariant metrics are not right-invariant. Since the inversion exchanges left and right, such metrics are not inverse consistent either. This means that the Fréchet mean for a left (resp. right) invariant metric is not consistent with inversion and right (resp. left) composition. In particular, the mean of inverse transformations is not the inverse of the mean.

One can wonder if there exists a more general framework, obviously non-Riemannian, to realize consistent statistics on these Lie groups. Indeed, numerous methods in Lie groups are based on pure group properties, independently of the action of transformations on objects. These methods rely in particular on one-parameter subgroups, realized in finite dimensional matrix Lie groups by the matrix exponential. There exist particularly efficient algorithms to compute the matrix exponential like the scaling and squaring procedure [23] or for integrating differential equations on Lie groups in geometric numerical integration theory [21, 25]. In infinite dimension, one parameter subgroups are deformations realized by the flow of Stationary Velocity Fields (SVFs), as we will see in Section 3.1. Parametrizing diffeomorphisms with SVFs was proposed for medical image registration by [3] and very quickly adopted by many other authors [5, 59, 22, 42]. The group structure was also used to obtain efficient low-dimensional parametric locally affine diffeomorphisms as we will see in Section 5.1.

In fact, these one-parameter subgroups (matrix exponential and flow of SVF) are the geodesics of the Cartan-Schouten connection, a more invariant and thus more consistent but non-metric structure on transformation groups. We detail in this section the extension of the computing and statistical framework to Lie groups endowed with the affine symmetric connection. In the medical imaging and geometric statistics communities, these notions were first developed in [48, 34]. A more complete account on the theory appeared recently in [51]. We refer the reader to this chapter for more explanations and mathematical details.

2.1 Affine geodesics

Geodesics, exponential and log maps are among the most fundamental tools to work on differential manifolds. In order to define a notion of geodesics in non-Riemannian spaces, we cannot rely on the shortest path as there is no Riemannian metric to measure length. The main idea is to define straight lines as curves with vanishing acceleration, or equivalently curves whose tangent vectors remains parallel to themselves (auto-parallel curves). In order to compare vectors living in different tangent spaces (even at points which are infinitesimally close), we need to provide a notion of parallel transport from one tangent space to the other. Likewise, computing accelerations requires a notion of infinitesimal parallel transport that is called a connection.

In a local coordinate system, a connection is completely determined by its coordinates on the basis vector fields: $\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k$. The n^3 coordinates Γ_{ij}^k of the connection are called the Christoffel symbols. A curve $\gamma(t)$ is a geodesic if its tangent vector $\dot{\gamma}(t)$ remains parallel to itself, i.e. if the covariant derivative $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ of γ is zero. In a local coordinate system, the equation of the geodesics is thus: $\ddot{\gamma}^k + \Gamma_{ij}^k \dot{\gamma}^i \dot{\gamma}^j = 0$, exactly as in a Riemannian case. The difference is that the affine connection case starts with the Christoffel symbols, while these are determined by the metric in the Riemannian case, giving a natural connection called the Levi-Civita connection. Unfortunately, the converse is not always possible: many affine connection spaces do not accept a compatible Riemannian metric. Riemannian manifolds are only a subset of affine connection spaces.

What is remarkable is that we conserve many properties of the Riemannian exponential map in affine connection spaces. For instance, the geodesic $\gamma_{(x,v)}(t)$ starting at any point x with any tangent vector v is defined for a sufficiently small time, which means that we can define the affine exponential map $\exp_x(v) = \gamma_{(x,v)}(1)$ for a sufficiently small neighborhood. Moreover, there exists at each point a *normal convex neighborhood* (NCN) in which any couple of points (x, y) is connected by a unique geodesic $\gamma(t)$ entirely contained in this neighborhood. We can thus define the log-map locally without ambiguity.

2.2 An affine symmetric space structure for Lie groups

In the case of Lie groups, the Symmetric Cartan-Schouten (SCS) connection is a canonical torsion free connection introduced in 1926 shortly after the invention of the notion of connection by Cartan [13]. This is also the unique affine connection induced by the canonical symmetric space structure of the Lie groups with the symmetry $s_g(h) = gh^{(-1)}g$. The SCS connection exists on all Lie groups, and it is left and right-invariant. When there exists a bi-invariant metric on the Lie group (i.e. when the group is the direct product of Abelian and compact groups), the SCS connection is the Levi-Civita connection of that metric. However, the SCS connection still exists when there is no bi-invariant metric.

Geodesics of the SCS connection are called group geodesics. The ones going through the identity are the flow of left-invariant vector fields. They are also called one-parameter subgroups since $\gamma(s+t) = \gamma(s) \circ \gamma(t)$ is an isomorphism of Lie groups, that is a mapping that preserves the Lie group structure. In matrix Lie groups, one-parameter subgroups are described by the exponential $\exp(M) = \sum_{k=0}^{\infty} M^k/k!$ of square matrices. Conversely, if there exists a square matrix M such that $\exp(M) = A$, then M is said to be a logarithm of the invertible square matrix A . In general, the logarithm of a real invertible matrix is not unique and may fail to exist. However, when this

matrix has no (complex) eigenvalue on the (closed) half line of negative real numbers, then it has a unique real logarithm $\log(M)$, called the principal logarithm whose (complex) eigenvalues have an imaginary part in $(-\pi, \pi)$ [27, 18]. Moreover, matrix \exp and \log can be very efficiently numerically computed with the ‘Scaling and Squaring Method’ [23] and ‘Inverse Scaling and Squaring Method’ [24].

Group geodesics starting from other point can be obtained very simply by left or right translation: $\gamma(t) = A \exp(tA^{(-1)}M) = \exp(tMA^{(-1)})A$ is the geodesic starting at A with tangent vector M . In finite dimension, the group exponential is a chart around the identity. In infinite dimensional Fréchet manifolds, the absence of an inverse function theorem prevents the straightforward extension of this property to general groups of diffeomorphisms and one can show that there exists diffeomorphisms as close as we want to the identity that cannot be reached by one-parameter subgroups [28]. In practice, though, the diffeomorphisms that we cannot reach have not yet proved to be of practical use for any real-world application.

Thus, everything looks very similar to the Riemannian case, except that group geodesics are defined from group properties only and do not require any Riemannian metric. One should be careful that they are generally different from the Riemannian exponential map associated to a Riemannian metric on the Lie group.

2.3 Statistics in affine connection spaces

In order to generalize the Riemannian statistical tools to affine connection spaces, the Fréchet / Karcher means have to be replaced by the weaker notion of exponential barycenters, which are the critical points of the variance in Riemannian manifolds. In an affine connection space, the exponential barycenters of a set of points $\{x_1 \dots x_n\}$ are implicitly defined as the points x for which the tangent mean field vanishes:

$$\mathfrak{M}(x) = \frac{1}{n} \sum_{i=1}^n \log_x(x_i) = 0. \quad (1)$$

While this definition is close to the Riemannian center of mass [19], it uses the logarithm of the affine connection instead of the Riemannian logarithm.

For sufficiently concentrated distributions with compact support, typically in a normal convex neighborhood, there exists at least one exponential barycenter. Moreover, exponential barycenters are stable by affine diffeomorphisms (connection preserving maps). For distributions whose support is too large, exponential barycenters may not exist. This should be related to the classical non-existence of the mean for heavy tailed distributions in Euclidean spaces. The uniqueness of the exponential barycenter can be shown with ad-

ditional assumptions, either on the derivatives of the curvature [9] or on a stronger notion of convexity (semi-local convex geometry [1]).

Higher order moments can also be defined locally. For instance, the empirical covariance field is the 2-contravariant tensor $\Sigma(x) = \frac{1}{n} \sum_{i=1}^n \log_x(x_i) \otimes \log_x(x_i)$ and its value $\Sigma = \Sigma(\bar{x})$ at the exponential barycenter \bar{x} is called the empirical covariance. Notice that this definition depends on the chosen basis and that diagonalizing the matrix makes no sense since we do not know what are orthonormal unit vectors. Thus, tangent PCA is not easily generalized. Despite the absence of a canonical reference metric, the Mahalanobis distance of a point y to a distribution can be defined locally as in the Riemannian case with the inverse of the covariance matrix. This definition is independent of the basis chosen for the tangent space and is actually invariant under affine diffeomorphisms of the manifold. This simple extension of the Mahalanobis distance suggests that it might be possible to extend much more statistical definitions and tools on affine connection spaces in a consistent way.

2.4 The case of Lie groups with the canonical Cartan-Schouten connection

Thanks to the bi-invariance properties of the SCS connection, the exponential barycenters of Eq.(1) define bi-invariant group means. Let $\{A_i\}$ be a set of transformations from the group (we can think of matrices here). Then a transformation \bar{A} verifying $\sum_i \log(\bar{A}^{(-1)} A_i) = \sum_i \log(A_i \bar{A}^{(-1)}) = 0$ is a group mean which exists and is unique for sufficiently concentrated data $\{A_i\}$. Moreover, the fixed point iteration $\bar{A}_{t+1} = \sum_i \log(\bar{A}_t^{(-1)} A_i)$ converges to the bi-invariant mean at least linearly (still under a sufficient concentration condition), which provides a very useful algorithm to compute it in practice.

The bi-invariant mean turns out to be globally unique in a number of Lie groups which do not support any bi-invariant metric, for instance nilpotent or some specific solvable groups [48, 34, 51]. For rigid-body transformations, the bi-invariant mean is unique when the mean rotation is unique, so that we do not lose anything with respect to the Riemannian setting. Thus, the group mean appears to be a very general and natural notion on Lie groups.

3 The SVF framework for shape and deformation modeling

In the context of medical image registration, diffeomorphic registration was introduced with the “Large Deformation Diffeomorphic Metric Mapping (LD-DMM)” framework [57, 7], which parametrizes deformations with the flow of *time varying velocity fields* $v(x, t)$ with a right-invariant Riemannian met-

ric (see [61] for a complete mathematical description). In view of reducing the computational and memory costs, [3] subsequently proposed to restrict this parametrization to the subset of diffeomorphisms parametrized by the flow of stationary velocity fields (SVFs), for which efficient image registration methods like log-Demons have been developed with a great success from the practical point of view. The previous theory of statistics on Lie groups with the canonical symmetric Cartan-Schouten connection provides strong theoretical bases for the use of these one-parameter subgroups.

3.1 Diffeomorphisms parametrized by stationary velocity fields

To construct our group of diffeomorphisms, one first restricts the Lie algebra to sufficiently regular velocity fields according to the regularization term of the SVF registration algorithms [59, 22] or to the spline parametrization of the SVF in [5, 42]. The flow of these stationary velocity fields and their finite composition generates a group of diffeomorphisms that we endow with the affine symmetric Cartan-Schouten connection. The geodesics are then exactly the one-parameter subgroups generated by the flow of SVFs: the deformation $\phi = \exp(v)$ is parametrized by the Lie group exponential of a smooth SVF $v : \Omega \rightarrow \mathbb{R}^3$ through the Ordinary Differential Equation (ODE) $\frac{\partial \phi(x,t)}{\partial t} = v(\phi(x,t))$ with initial condition $\phi(x,0) = x$. It is known that not all diffeomorphisms can be reached by such a one-parameter subgroup (we might have to compose several ones to reach them all) but in practice this does not seem to be a limitation.

Many of the techniques developed for the matrix case can be adapted to SVFs. This is the case of the scaling and squaring algorithm, which integrates the previous ODE very effectively thanks to the iterative composition of successive exponentials: $\exp(v) = \exp(v/2) \exp(v/2) = (\exp(v/2^n))^n$. Inverting a deformation is usually quite difficult or at least computationally intensive as we have to find ψ such that $\psi(\phi(x)) = \phi(\psi(x)) = x$. This is generally performed using the least-square minimization of the error on the above equation integrated over the image domain. In the SVF setting, such a computation can be performed seamlessly since $\phi^{(-1)} = \exp(-v)$.

In order to measure volume changes induced by the deformation, one usually computes the Jacobian matrix $d\phi = \nabla \phi^T$ using finite differences, and then takes its determinant. However, finite-differences schemes are highly sensitive to noise. In the SVF framework, the log-Jacobian can be reliably estimated by finite differences for the scaled velocity field $v/2^n$, and then recursively computed thanks to the chain rule in the scaling-and-squaring scheme thanks to the additive property of the one-parameter subgroups. The Jacobian determinant that we obtain is therefore fully consistent with the exponential path taken to compute the diffeomorphism.

Last but not least, one often needs to compose two deformations, for instance to update the current estimation in an image registration algorithm. The Baker Campbell Hausdorff (BCH) formula is a series expansion that approximates the SVF

$$BCH(v, u) = \log(\exp(v)\exp(u)) = v + u + \frac{1}{2}[v, u] + \frac{1}{12}[v, [v, u]] + \dots$$

as a power series in the two SVFs u and v . In this formula, the Lie bracket of vector fields is $v: [v, u] = dvu - duv = \partial_u v - \partial_v u$. In the context of diffeomorphic image registration, this trick to do all the computations in the Lie algebra was introduced by [8].

3.2 SVF-Based Diffeomorphic Registration with the Log-Demons

The encoding of diffeomorphisms via the flow of SVF [3] inspired several SVF-based image registration algorithms [58, 8, 5, 59, 22, 42, 41, 29]. Among them, the log-demons registration algorithm [58, 59, 60, 29] found a considerable interest in the medical image registration community with many successful applications to clinical problems [53, 36, 32, 55].

Given a pair of images $I, J : \mathbb{R}^3 \mapsto \mathbb{R}$, the log-demons algorithm aims at estimating a SVF v parametrizing diffeomorphically the spatial correspondences that minimize a similarity functional $Sim[I, J \circ \exp(v)]$. A classically used similarity criterion is the sum of square differences (SSD) $Sim[I, J] = \int (I(x) - J(x))^2 dx$. In order to symmetrize the criterion and ensure inverse consistency, one can add the symmetric similarity criterion $Sim[I \circ \exp(-v), J]$ [59] or more simply measure the discrepancy at the mid-deformation point using $Sim[I \circ \exp(-v/2), J \circ \exp(v/2)]$. This last formulation allows to easily symmetrize a similarity functional that is more complex than the SSD, such as the local correlation coefficient (LCC) [29].

In order to prevent overfitting, a regularization term that promotes spatially more regular solutions is added to the similarity criterion. In the log-demons framework, this regularization is naturally performed on the SVF v rather than on the deformation $\phi = \exp(v)$. A feature of the demons' type algorithms is also to introduce an auxiliary variable encoding for the correspondences, here a SVF v_c , in addition to the SVF v encoding for the transformation [11]. The two variables are linked using a coupling criterion that prevents the two from being too far away from each other. The criterion optimized by the log-demons is then:

$$E(v, v_c, I, J) = \frac{1}{\sigma_i^2} Sim(I, J, v_c) + \frac{1}{\sigma_x^2} \|v_c - v\|_{L_2}^2 + \frac{1}{\sigma_T^2} Reg(v). \quad (2)$$

The interest of the auxiliary variable is to decouple a non-linear and non-convex optimization problem into two simpler optimization problems that are respectively local and quadratic. The classical criterion is obtained at the limit when the typical scale of the error σ_x^2 between the transformation and the correspondences tends to zero.

The minimization of (2) is alternatively performed with respect to the correspondence SVF v_c and the transformation SVF v . The first step is a non-convex but purely local problem which is usually optimized via gradient descent using Gauss-Newton or Levenberg Marquardt algorithms. To simplify the second step, one can choose $Reg(\cdot)$ to be an Isotropic Differential Quadratic Form (IDQF [10]), which leads to a closed form solution by convolution. In most cases, one chooses this convolution to be Gaussian: $v = G_\sigma * v_c$, which can be computed very efficiently using separable recursive filters.

4 Modeling longitudinal deformation trajectories in Alzheimer’s disease

With the log-demons algorithm, we can register two longitudinal images of the same subject. When more images are available at multiple time-points, we can regress the geodesic that best describes the different registrations to obtain a longitudinal deformation trajectory encoded by a single SVF [32, 20]. We should notice that while such a geodesic is a linear model in the space of SVFs, it is a highly non-linear model on the displacement field and on the space of images.

However, follow-up imaging studies usually require to transport this subject-specific longitudinal trajectories in a common reference for group-wise statistical analysis. A typical example is the analysis of structural brain changes with aging in Alzheimer’s disease versus normal controls. It is quite common in neuroimaging to transport a scalar summary of the changes over time like the Jacobian or the log-Jacobian encoding for local volume changes. This is easy and numerically stable as we just have to resample the scalar map. However, this does not allow to compute the ”average” group-wise deformation and its variability, nor to transport it back at the subject level to predict what will be the future deformation. To realize such a generative model of the longitudinal deformations, we should normalize the deformations as a geometric object and not just its components independently. This involves defining a method of transport of the longitudinal deformation parameters along the inter-subject change of coordinate system.

4.1 Parallel transport in Riemannian and affine spaces

Depending on the considered parametrization of the transformation (displacement fields, stationary velocity fields, initial momentum field...), different approaches have been proposed in the literature to transport longitudinal deformations. In the Riemannian and affine connection space setting, where longitudinal deformations are encoded by geodesics parametrized by their initial tangent vector, it is natural to consider the parallel transport of this initial tangent vector (describing the longitudinal deformation) along the inter-subject deformation curve. Parallel transport is an isometry of tangent spaces in the Riemannian case, so that the norm is conserved. In the affine connection case, this is an affine transformation of tangent spaces. Instead of defining properly the parallel transport in the continuous setting and approximating it in an inconsistent discrete setting, it was proposed in [30] to rely on a carefully designed discrete construction that intrinsically respects all the symmetries on the problem: the Schild's Ladder. This algorithm was initially introduced in the 1970s by the physicist Alfred Schild [16] in the field of the general relativity. The method was refined with the pole ladder in [33] to minimize the number of steps when the transport is made along geodesics. Schild's and pole ladders only require the computation of exponentials and logarithms, and thus can easily and consistently be implemented for any manifold provided that we have these basic algorithmic bricks.

In this process, the numerical accuracy of parallel transport algorithm is the key to preserve the statistical information. The analysis of pole ladder in [47] actually showed that the scheme is of order three in general affine connection spaces with a symmetric connection, an order higher than expected. Moreover, the fourth order error term vanishes in affine symmetric spaces since the curvature is covariantly constant. In fact, the error terms vanish completely in a symmetric affine connection space: one step of pole ladder realizes a transvection, which is an exact parallel transport (provided that geodesics and mid-points are computed exactly of course) [47]. These properties make pole ladder a very attractive alternative for parallel transport in the framework of diffeomorphisms parametrized by SVFs. In particular, parallel transport has a closed form expression $\Pi^v(u) = \log(\exp(v/2)\exp(u)\exp(-v/2))$ [33]. In practice, the symmetric reformulation of the pole ladder scheme using the composition of two central symmetries (a transvection) gives numerically more stable results and was recently shown to be better than the traditional Euclidean point-distribution model on cardiac ventricular surfaces [26].

4.2 Longitudinal modeling of Alzheimer's progression

Parallel transport allows us to compute a mean deformation trajectory at the group level and to differentiate populations on the basis of their full deformation features and not only according to local volume change as in traditional tensor-based morphometry (TBM). We illustrate in this Section an application of this framework to the statistical modeling of the longitudinal changes in a group of patients affected by Alzheimer's disease (AD). In this disease, it was shown that the brain atrophy that one can measure using the registration of time sequences of magnetic resonance images (MRI) is strongly correlated to cognitive performance and neuropsychological scores. Thus, deformation-based morphometry provides an interesting surrogate image biomarker for the progression of the disease from pre-clinical to pathological stages.

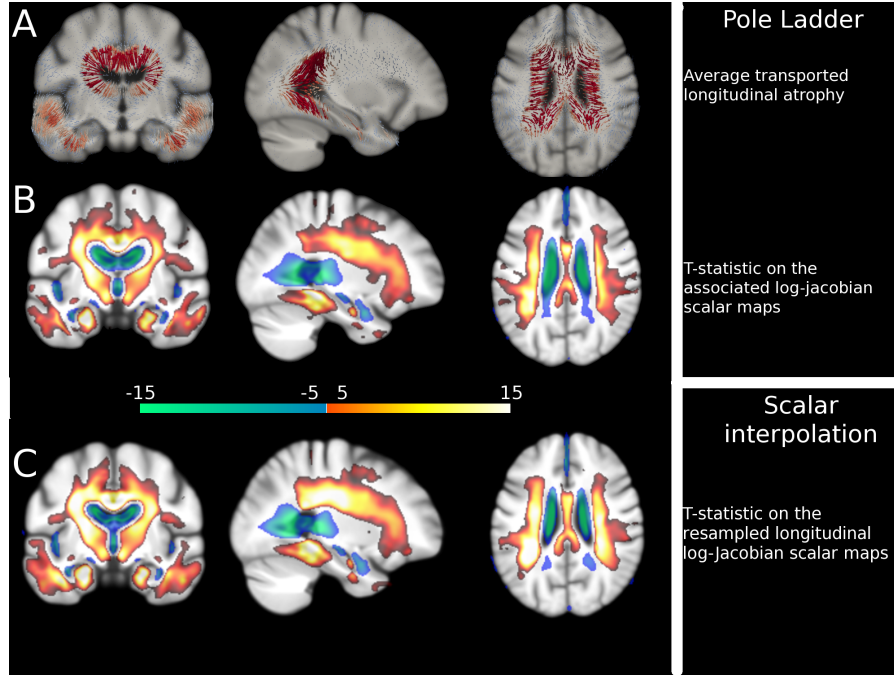


Fig. 2 One year structural changes for 135 Alzheimer's patients. A) Mean of the longitudinal SVFs transported in the template space with the pole ladder. We notice the lateral expansion of the ventricles and the contraction in the temporal areas. B) T-statistic for the corresponding log-Jacobian values significantly different from 0 ($p < 0.001$ FDR corrected). C) T-statistic for longitudinal log-Jacobian scalar maps resampled from the subject to the template space. Blue color: significant expansion, Red color: significant contraction (Figure reproduced from [33] with permission).

The study that we summarize here was published in [33]. We took 135 Alzheimer’s subjects of the ADNI database with images at baseline and one year later. The SVFs v_i parametrizing the longitudinal deformation trajectory $\phi_i = \exp(v_i)$ between the two time-points was estimated with the LCC log-demons. These SVFs were then transported with the pole ladder from their subject-specific space to the template reference T along the subject-to-template geodesic, also computed using the LCC log-demons. The mean \bar{v} of the transported SVFs in the template space parametrizes our model of the group-wise longitudinal progression $\exp(t\bar{v})$. The spatial localization of significant longitudinal changes (expansion or contraction) was established using one-sample t-test on the log-Jacobian scalar maps after parallel transport. In order to compare with the traditional method used in tensor-based morphometry, another one-sample t-test was computed on the subject-specific log-Jacobian scalar maps resampled in the template space.

Results are presented in Figure 2. Row A illustrates the mean SVF of the transported one-year longitudinal trajectories. It shows a pronounced enlargement of the ventricles, an expansion of their temporal horns and a consistent contracting flow in the temporal areas. It is impressive that the extrapolation of the deformation along the geodesic from 1 year to 15 years produces a sequence of very realistic images going from a young brain at $t = -7$ years to a quite old AD brain with very large ventricles and almost no hippocampus at $t = 8$ years. This shows that a linear model in a carefully designed non-linear manifold of diffeomorphisms can handle realistically very large shape deformations. Such a result is definitely out of sight with a statistical model on the displacement vector field or even with a classical point distribution model (PDM), as is often done in classical medical shape analysis.

Evaluating the volumetric changes (here computed with the log-Jacobian) leads to areas of significant expansion around the ventricles with a spread in the Cerebrospinal Fluid (CSF, row B). Areas of significant contraction are located as expected in the temporal lobes, hippocampi, parahippocampal gyrus and in the posterior cingulate. These results are in agreement with the classical resampling of the subject-specific log-Jacobian maps done in TBM presented in row C. It is striking that there is no substantial loss of localization power for volume changes by transporting SVFs instead of resampling the scalar log-Jacobian maps. In contrast to TBM, we also preserve the full multidimensional information about the transformation, which allows to make more powerful multivariate voxel-by-voxel comparisons than the ones obtained with the classical univariate tests. For example, we could show for the first time in [?] a statistically significant different brain shape evolutions depending on the level of $A\beta_{1-42}$ protein in the CSF. As the level of $A\beta_{1-42}$ is sometimes considered as pre-symptomatic of Alzheimer’s disease, we could be observing the very first morphological impact of the disease. More generally, a normal longitudinal deformation model allows to disentangle normal aging

component from the pathological atrophy even with one time-point only per patient (cross-sectional design) [35].

The SVF describing the trajectory can also be decomposed using Helmholtz’ decomposition into a divergent part (the gradient of a scalar potential) that encodes the local volume changes and a divergence free reorientation pattern [31]. This allows to consistently define anatomical regions of longitudinal brain atrophy in multiple patients, leading to improved measurements of the quantification of the longitudinal hippocampal and ventricular atrophy in AD. This method provided very reliable results during the MIRIAD atrophy challenge for the regional atrophy quantification in the brain, with really state of the art performances (first and second rank on deep structures) [14].

5 The SVF framework for Cardiac motion analysis

Cardiac motion plays an important role in the function of the heart and abnormalities in the cardiac motion can be the cause of multiple diseases and complications. Modeling cardiac motion can therefore provide precious information. Unfortunately, the outputs from cardiac motion models are complex. Therefore, they are hard to analyze, compare and personalize. The approach described below relies on a polyaffine projection applied to the whole cardiac motion and results in a few parameters that are physiologically relevant.

5.1 Parametric diffeomorphisms with locally affine transformations

The polyaffine framework assumes that the image domain is divided into regions defined by smooth normalized weights ω_i (i.e summing up to one over all regions). The transformation of each region is modeled by a locally affine transformation expressed by a 4×4 matrix A_i in homogeneous coordinates. Using the principal logarithm $M_i = \log(A_i)$ of these matrices, we compute the SVF at any voxel x (expressed in homogeneous coordinates) as the weighted sum of these locally affine transformation [4, 2]:

$$v_{poly}(x) = \sum_i \omega_i(x) M_i x.$$

The polyaffine transformation is then obtained by taking the flow of SVF using the previous scaling and squaring algorithm for the exponential. This leads to a very flexible locally affine diffeomorphism parametrized by very few parameters. In this formulation, taking the log in homogeneous coordinates ensures that the inverse of the polyaffine transformations is also a polyaffine

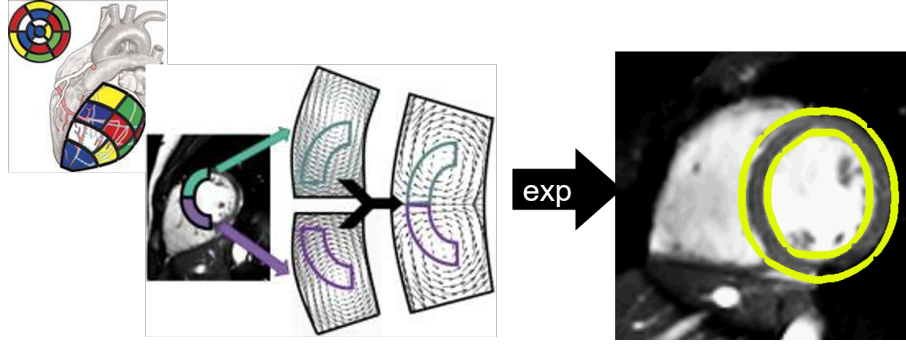


Fig. 3 A low dimensional parametrization of diffeomorphisms for tracking cardiac motion in cine-MRI: the flow of an affine transformation with 12 parameters (middle) is generating a local velocity field around each of the 17 AHA regions (on the left). The weighted average of these 17 affine velocity fields produces a global velocity field whose flow (the group exponential) is parametrizing the heart deformation (on the right). In this context, motion tracking consists in optimizing the $12 \times 17 = 204$ regional parameters, which is easily done in the log-demons framework.

transformation. This property is necessary to create generative motion models.

As shown in [56], the log affine matrix parameters M_i can be estimated explicitly by a linear least squares projection of an observed velocity field $v(x)$ into the space of Log-Euclidean Polyaffine Transformations (LEPT's). Denoting $\Sigma_{ij} = \int_{\Omega} \omega_i(x) \omega_j(x) x x^T dx$ and $B_i = \int_{\Omega} \omega_i(x) v(x) x^T dx$, the optimal matrix of log-affine transformation parameters $M = [M_1, M_2, \dots, M_n]$, minimizing the criterion $C(M) = \int_{\Omega} \|\sum_i \omega_i(x) M_i x - v(x)\|^2 dx$ is given by $M = B \Sigma^{-1}$. The solution is unique when the Gram matrix Σ of the basis vectors of our polyaffine SVF is invertible. This gives rise to the polyaffine log-demons algorithm where the estimated SVF at each step of the log-demons algorithm is projected into this low dimensional parameter space instead of being regularized.

A cardiac-specific version of this model was proposed in [38] by choosing regions corresponding to the standard American Heart Association (AHA) regions for the left ventricle (Fig.3). The weights ω_i are normalized Gaussian functions around the barycenter of each regions. Furthermore, an additional regularization between neighboring regions was added to account for the connectedness of cardiac tissue among neighboring regions, as well as an incompressibility penalization to account for the low volume change in cardiac tissue over the cardiac cycle.

5.2 Towards intelligible population-based cardiac motion features

The interpretability of the affine parameters of each region can be considerably increased by expressing the local affine transformations in a local coordinate system having radial, longitudinal vector and circumferential axes for each region. The resulting parameters can be related to physiological deformation: the translations parameters correspond to the motion along the radial, longitudinal, and circumferential axes while the linear part of the transformation encodes the circumferential twisting, radial thickening, and longitudinal shrinking. In a first study, the parameters were further reduced by assuming the linear part of each matrix M_i to be diagonal, thus reducing the number of parameters to 6 per region. These intelligible parameters were then used by supervised learning algorithms to classify a database of 200 cases with equal number of infarcted and non-infarcted subjects (the STACOM statistical shape modeling). A 10-fold cross-validation showed that the method was achieving more than 95% of correct classification on yet-unseen data [54].

In [37, 38, 39], relevant factors discriminating between the motion patterns of healthy and unhealthy subjects were identified thanks to a Tucker decomposition on Polyaffine motion parameters with a constraint on the sparsity of the core tensor (which essentially defines the loadings of each mode combination). The key idea is to consider that the parameters resulting from the tracking of the motion over cardiac image sequences of a population can be stacked in a 4-way tensor along motion parameters \times region \times time \times subject. Performing the decomposition on the full tensor directly using 4-way Tucker Decomposition has the advantage of describing how all the components interact (as opposed to matricising the tensor and performing 2-way decomposition using classical singular value decomposition). The Tucker tensor decomposition method is a higher-order extension of PCA which computes orthonormal subspaces associated with each axis of the data tensor. Thus, we get modes that independently describe: a reduced basis of transformations (common to all regions, all time-points of the sequence and all subjects); a spatial basis (region weights) that localize deformations on the heart; a set of modes along time that triggers the deformation; and discriminative factors across clinical conditions. In order to minimize the number of interactions between all these modes along all the tensor axes, sparsity constraints were added on the core tensor. The sparsity of the discriminating factors and their individual intelligibility appears to be a key for a clear and intuitive interpretation of differences between populations in order to gain insight into pathology-specific functional behavior.

The method was applied to a data-set of 15 healthy subjects and 10 Tetralogy of Fallot patients with short axis cine MRI sequences of 12 to 16 slices (slice thickness of 8mm) and 15 to 30 image frames. The decomposition was performed with 5 modes per axis and the core tensor loadings for each sub-

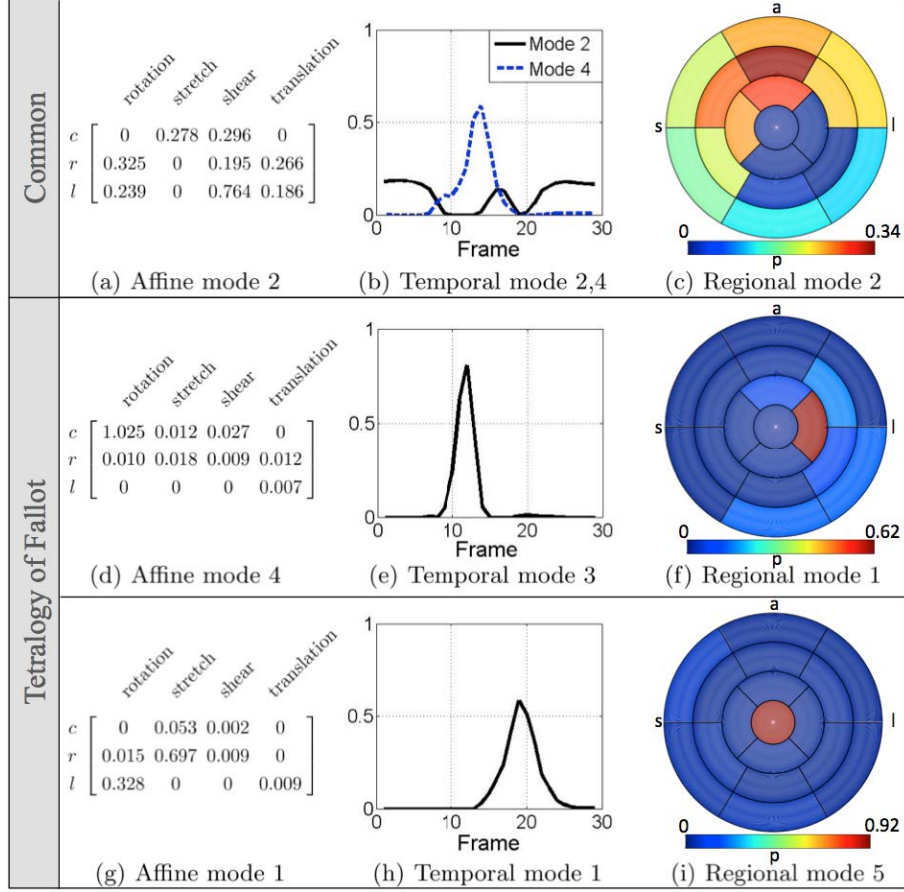


Fig. 4 Dominant mode combinations common to healthy and ToF cohorts: affine mode 2 (a), temporal modes 2 and 4 (b), and regional mode 2 (c). Key - a: anterior, p: posterior, s: septal, l: lateral. Figure reproduced from [37] with permission.

ject were averaged for the different groups. This showed that the two groups share some common dominant loadings. As expected, the Tetralogy of Fallot group also has some additional dominant loadings representing the abnormal motion patterns in these patients.

The common dominant mode combinations are plotted in Fig.4 (top row). The affine mode for the dominant mode combinations (Fig.4, a) shows predominant stretching in the circumferential direction related to the twisting motion in the left ventricle. The temporal modes (Fig.4, b) show a dominant pattern around the end- and mid-diastolic phases for mode 2, which may be due to the end of relaxation and end of filling. The dominant regions for these mode combinations are anterior (Fig.4, c). The dominant mode combinations for the Tetralogy of Fallot group are plotted in Fig.4. The affine mode for

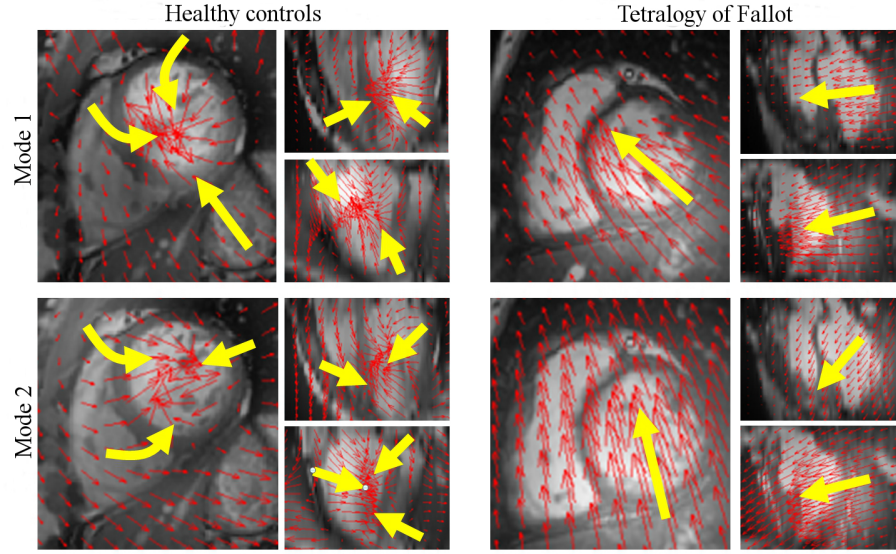


Fig. 5 Three views of the first (top row) and second (bottom row) spatial modes for the healthy controls (left) and for the Tetralogy of Fallot patients (right). The modes for the healthy controls represent the radial contraction and circumferential motion, whereas the modes for the Tetralogy of Fallot patients represent the translation towards the right ventricle. Yellow arrows indicate the general direction of motion. Figure ©IEEE 2015, reproduced from [38] with permission

the first dominant combination (Fig.4, d) indicates little longitudinal motion. The corresponding temporal mode (Fig.4, e) represents a peak at the end systolic frame (around one third of the length of the cardiac cycle). The corresponding regional mode (Fig.4, f) indicates that there is a dominance in the motion in the lateral wall. This is an area with known motion abnormalities in these patients given that the motion in the free wall of the left ventricle is dragged towards the septum. The temporal mode for the second dominant mode (Fig.4, h) has instead a peak around mid-systole, with corresponding regional mode (Fig.4, i), indicating dominance around the apex, which may be due to poor resolution at the apex. The SVF corresponding to the first two spatial modes are shown in Fig.5. The first mode for the healthy controls appears to capture both the radial contraction and the circumferential motion (shown in block yellow arrows). The Tetralogy of Fallot modes, on the other hand, appear to capture a translation of the free wall and septal wall towards the right ventricle (RV). This abnormal motion is evident in the image sequences of these patients.

6 Conclusion

We have presented in this Chapter an overview of the theory of statistics on non-linear spaces and of its application to the modeling of shapes and deformations in medical image analysis. When the variability of shapes becomes important, linear methods like point distribution models for shapes or linear statistics on displacement vector fields for images and deformations become ill-posed as they authorize self intersections. Considering non-linear spaces that are locally Euclidean (i.e. Riemannian manifolds) solves this issue. The cost to pay is that we have to work locally with tangent vectors and geodesics. However, once the exponential and log maps are implemented at any point of our shape space, many algorithms and statistical methods can be generalized quite seamlessly to these non-linear spaces.

For statistics on deformations, we have to consider smooth manifolds that have an additional structure: transformations form a Lie group under composition and inversion. One difficulty to use the Riemannian framework is that there often does not exist a metric which is completely invariant with respect to all the group operations (composition on the left and on the right, inversion). As a consequence, the statistics that we compute with left or right-invariant metrics are not fully consistent with the group structure. We present in this chapter an extension of the Riemannian framework to affine connection spaces that solves this problem. In this new setting, all the computations continue to be related to geodesics using the exp and log maps. Here, geodesics are defined with the more general notion of straight lines (zero acceleration or auto-parallel curves) instead of being shortest paths. Every Riemannian manifold is an affine connection space with the Levi-Civita connection, but the reverse is not true. This is why we can find a canonical connection on every Lie group (the symmetric Cartan-Schouten connection) that is consistent with left and right composition as well as inversion while there is generally no bi-invariant Riemannian metric.

We have drafted the generalization of the statistical theory to this affine setting, and we have shown that it can lead to a very powerful framework for diffeomorphisms where geodesics starting from the identity are simply the flow of stationary velocity fields (SVFs). Very well known non-linear registration algorithms based on this parametrization of diffeomorphisms are: the log-demons [59], Ashburner's DARTEL toolbox in SPM8 [5] and the NiftyReg registration package [42, 41]. The combination of these very efficient algorithm with the well-posed geometrical and statistical framework allows to develop new methods for the analysis of longitudinal data. Furthermore, the affine symmetric structure of our group of deformation provides parallel transport algorithms that are numerically more stable and efficient than in the Riemannian case. We showed on two brain and cardiac applications that this allows to construct not only statistically more powerful analysis tools, but also generative models of shape motion and evolution.

With polyaffine deformations in the cardiac example, we have also shown that the deformation parameters can be localized and aligned with biophysical reference frames to produce a diffeomorphism parametrized by low dimensional and intelligible parameters. Such a sensible vectorization of deformations is necessary for sparse decomposition methods: each parameter has to make sense individually as an atom of deformation if we want to describe the observed shape changes with an extremely low number of meaningful variables. This opens the way to very promising factor analysis methods dissociating the influence of the type of local deformation, the localization, the time trigger and the influence of the disease as we have shown with the sparse Tucker tensor decomposition. There is no doubt that these methods will find many other applications in medical image analysis.

For efficiency, the medical applications shown in this chapter were implemented using C++ software dedicated to 3D image registration parameterized by SVFs. An open-source implementation of the symmetric logdemons integrated into the Insight Toolkit (ITK) is available at <http://hdl.handle.net/10380/3060>. A significant improvement of this software including the more robust LCC similarity measure and symmetric confidence masks is available at <https://team.inria.fr/epione/fr/software/lcclogdemons/> [29], along with additional standalone tools to work on SVFs including the pole ladder algorithm [33]. The code for the polyaffine logdemons is also available as an open-source ITK package at <https://github.com/ChristofSeiler/PolyaffineTransformationTrees> [56]. For Riemannian geometric data which are less computationally demanding than the very large 3D images, it can be more comfortable to work in python. The recent *Geomstat* Python toolbox <https://github.com/geomstats/geomstats> provides an efficient and user friendly interface for computing the exponential and logarithmic maps, geodesics, parallel transport on nonlinear manifolds such as hyperbolic spaces, spaces of symmetric positive definite matrices, Lie groups of transformations, and many more. The package provides methods for statistical estimation and learning algorithms, clustering and dimension reduction on manifolds with support for different execution backends, namely NumPy, PyTorch and TensorFlow, enabling GPU acceleration [40].

Acknowledgements:

This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant G-Statistics No 786854).

References

1. M. Araudon and X.-M. Li. Barycenters of measures transported by stochastic flows. *The annals of probability*, 33(4):1509–1543, 2005.
2. V. Arsigny, O. Commowick, N. Ayache, and X. Pennec. A Fast and Log-Euclidean Polyaffine Framework for Locally Linear Registration. *Journal of Mathematical Imaging and Vision*, 33(2):222–238, 2009.
3. V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-Euclidean framework for statistics on diffeomorphisms. In *Proc. of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I*, number 4190 in LNCS, pages 924–931, 2–4 October 2006. PMID: 17354979.
4. V. Arsigny, X. Pennec, and N. Ayache. Polyrigid and Polyaffine Transformations: a Novel Geometrical Tool to Deal with Non-Rigid Deformations - Application to the registration of histological slices. *Medical Image Analysis*, 9(6):507–523, 2005. PMID: 15948656.
5. J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113, Oct. 2007.
6. M. Bauer, M. Bruveris, and P. W. Michor. Overview of the Geometries of Shape Spaces and Diffeomorphism Groups. *Journal of Mathematical Imaging and Vision*, 50(1-2):60–97, Sept. 2014. arXiv: 1305.1150.
7. M. F. Beg, M. I. M.I. Miller, A. Trounev, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. Journal of Computer Vision*, 61(2):139–157, 2005.
8. M. Bossa, M. Hernandez, and S. Olmos. Contributions to 3D diffeomorphic atlas estimation: Application to brain images. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*, volume 4792 of LNCS, pages 667–674. Springer-Verlag, 2007.
9. P. Buser and H. Karcher. *Gromov's almost flat manifolds*. Number 81 in Astérisque. Société mathématique de France, 1981.
10. P. Cachier and N. Ayache. Isotropic energies, filters and splines for vector field regularization. *Journal of Mathematical Imaging and Vision*, 20(3):251–265, 2004.
11. P. Cachier, E. Bardinet, D. Dormont, X. Pennec, and N. Ayache. Iconic feature based nonrigid registration: The pasha algorithm. *Computer Vision and Image Understanding*, 89(2-3):272–298, Feb.-march 2003. Special Issue on Nonrigid Registration.
12. M. P. d. Carmo. *Riemannian geometry*. Mathematics. Theory & applications. Birkhäuser, Boston, 1992.
13. E. Cartan and J. A. Schouten. On the geometry of the group-manifold of simple and semi-simple groups. *Proc. Akad. Wetensch, Amsterdam*, 29:803–815, 1926.
14. D. M. Cash, C. Frost, L. O. Iheme, D. Ünay, M. Kandemir, J. Fripp, O. Salvado, P. Bourgeat, M. Reuter, B. Fischl, M. Lorenzi, G. B. Frisoni, X. Pennec, R. K. Pierson, J. L. Gunter, M. L. Senjem, C. R. Jack, N. Guizard, V. S. Fonov, D. L. Collins, M. Modat, M. J. Cardoso, K. K. Leung, H. Wang, S. R. Das, P. A. Yushkevich, I. B. Malone, N. C. Fox, J. M. Schott, and S. Ourselin. Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. *NeuroImage*, 123:149–164, Dec. 2015.
15. I. L. Dryden and K. V. Mardia. *Statistical shape analysis with applications in R*. Wiley series in probability and statistics. Wiley, Chichester, UK ; Hoboken, NJ, second edition edition, 2016.
16. J. Ehlers, F. Pirani, and A. Schild. *The geometry of free fall and light propagation, in O’Raifeartaigh. General Relativity, Papers in Honor of J. L. Synge*. Oxford University Press, 1972.
17. P. Fletcher, C. Lu, S. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005, Aug. 2004.

18. J. Gallier. Logarithms and Square Roots of Real Matrices. *arXiv:0805.0245 [math]*, May 2008. arXiv: 0805.0245.
19. D. Groisser. Newton’s method, zeroes of vector fields, and the Riemannian center of mass. *Adv. in Applied Math.*, 33:95–135, 2004.
20. M. Hadj-Hamou, M. Lorenzi, N. Ayache, and X. Pennec. Longitudinal Analysis of Image Time Series with Diffeomorphic Deformations: A Computational Framework Based on Stationary Velocity Fields. *Frontiers in Neuroscience*, 10(236):18, June 2016.
21. E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration : structure preserving algorithm for ordinary differential equations*, volume 31 of *Springer series in computational mathematics*. Springer, 2002.
22. M. Hernandez, M. N. Bossa, and S. Olmos. Registration of Anatomical Images Using Paths of Diffeomorphisms Parameterized with Stationary Vector Field Flows. *International Journal of Computer Vision*, 85(3):291–306, Dec. 2009.
23. N. J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM Journal on Matrix Analysis and Applications*, 26(4):1179–1193, Jan. 2005.
24. S. Hun Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub. Approximating the logarithm of a matrix to specified accuracy. *SIAM J. Matrix Anal. Appl.*, 22(4):1112–1125, 2001.
25. A. Iserles, H. Z. Munthe-Kaas, S. P. Norsett, and A. Zanna. Lie-group methods. *Acta numerica*, 9:215–365, 2000.
26. S. Jia, N. Duchateau, P. Mocerì, M. Sermesant, and X. Pennec. Parallel Transport of Surface Deformations from Pole Ladder to Symmetrical Extension. In *ShapeMI MICCAI 2018: Workshop on Shape in Medical Imaging*, Granada, Spain, Sept. 2018.
27. C. S. Kenney and A. J. Laub. Condition estimates for matrix functions. *SIAM J. Matrix Anal. Appl.*, 10:191–209, 1989.
28. B. A. Khesin and R. Wendt. *The Geometry of Infinite Dimensional Lie groups*, volume 51 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics*. Springer Verlag, 2009.
29. M. Lorenzi, N. Ayache, G. B. Frisoni, and X. Pennec. LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. *NeuroImage*, 81(1):470–483, 2013.
30. M. Lorenzi, N. Ayache, and X. Pennec. Schild’s ladder for the parallel transport of deformations in time series of images. In G. Székely and H. Hahn, editors, *IPMI - 22nd International Conference on Information Processing in Medical Images- 2011*, volume 6801, pages 463–474, Kloster Irsee, Germany, July 2011. Springer. Honorable Mention (runner-up) for the Erbsmann Award.
31. M. Lorenzi, N. Ayache, and X. Pennec. Regional flux analysis for discovering and quantifying anatomical changes: An application to the brain morphometry in Alzheimer’s disease. *NeuroImage*, 115:224–234, July 2015.
32. M. Lorenzi, G. B. Frisoni, N. Ayache, and X. Pennec. Mapping the effects of $A\beta_{1-42}$ levels on the longitudinal changes in healthy aging: hierarchical modeling based on stationary velocity fields. In G. Fichtinger, A. Martel, and T. Peters, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, volume 6892 of *LNCIS*, pages 663–670. Springer, Heidelberg, Sep 2011.
33. M. Lorenzi and X. Pennec. Efficient Parallel Transport of Deformations in Time Series of Images: from Schild’s to Pole Ladder. *Journal of Mathematical Imaging and Vision*, 50(1-2):5–17, Oct. 2013.
34. M. Lorenzi and X. Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International Journal of Computer Vision*, 105(2):111–127, Nov. 2013.
35. M. Lorenzi, X. Pennec, G. B. Frisoni, and N. Ayache. Disentangling normal aging from Alzheimer’s disease in structural MR images. *Neurobiology of Aging*, 36:S42–S52, Jan. 2015.

36. T. Mansi, X. Pennec, M. Sermesant, H. Delingette, and N. Ayache. iLogDemons: A demons-based registration algorithm for tracking incompressible elastic biological tissues. *International Journal of Computer Vision*, 92(1):92–111, 2011.
37. K. Mcleod, M. Sermesant, P. Beerbaum, and X. Pennec. Descriptive and Intuitive Population-Based Cardiac Motion Analysis via Sparsity Constrained Tensor Decomposition. In *Medical Image Computing and Computer Assisted Intervention (MICCAI 2015)*, volume 9351 of *Lecture notes in computer science (LNCS)*, pages 419–426, Munich, Germany, Oct. 2015.
38. K. Mcleod, M. Sermesant, P. Beerbaum, and X. Pennec. Spatio-Temporal Tensor Decomposition of a Polyaffine Motion Model for a Better Analysis of Pathological Left Ventricular Dynamics. *IEEE Transactions on Medical Imaging*, 34(7):1562–1675, July 2015.
39. K. Mcleod, K. Tøndel, L. Calvet, M. Sermesant, and X. Pennec. Cardiac Motion Evolution Model for Analysis of Functional Changes Using Tensor Decomposition and Cross-Sectional Data. *IEEE Transactions on Biomedical Engineering*, 65(12):2769 – 2780, Mar. 2018.
40. N. Miolane, A. Le Brigant, J. Mathe, B. Hou, N. Guigui, Y. Thanwerdas, S. Heyder, O. Peltre, N. Koep, H. Zaatiti, H. Hajri, Y. Cabanes, T. Gerald, P. Chauchat, C. Shewmake, B. Kainz, C. Donnat, S. Holmes, and X. Pennec. Geomstats: A Python Package for Riemannian Geometry in Machine Learning. working paper or preprint, Apr. 2020.
41. M. Modat, G. R. Ridgway, P. Daga, M. J. Cardoso, D. J. Hawkes, J. Ashburner, and S. Ourselin. Log-euclidean free-form deformation. In *Proc. of SPIE Medical Imaging 2011*. SPIE, 2011.
42. M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin. Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine*, 98(3):278–284, June 2010.
43. X. Pennec. *L’incertitude dans les problèmes de reconnaissance et de recalage – Applications en imagerie médicale et biologie moléculaire*. Phd thesis, Ecole Polytechnique, Dec. 1996.
44. X. Pennec. Probabilities and Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. In Cetin, A.E., Akarun, L., Ertuzun, A., Gurcan, M.N., Yardimci, and Y., editors, *Proc. of Nonlinear Signal and Image Processing (NSIP’99)*, volume 1, pages 194–198, June 20–23, Antalya, Turkey, Turkey, 1999. IEEE-EURASIP.
45. X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.
46. X. Pennec. Barycentric Subspace Analysis on Manifolds. *Annals of Statistics*, 46(6A):2711–2746, July 2018.
47. X. Pennec. Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces. working paper or preprint, May 2018.
48. X. Pennec and V. Arsigny. Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups. In F. Barbaresco, A. Mishra, and F. Nielsen, editors, *Matrix Information Geometry*, pages 123–168. Springer, May 2012.
49. X. Pennec and N. Ayache. Uniform distribution, distance and expectation problems for geometric features processing. *Journal of Mathematical Imaging and Vision*, 9(1):49–67, 1998.
50. X. Pennec, P. Fillard, and N. Ayache. A Riemannian Framework for Tensor Computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.
51. X. Pennec and M. Lorenzi. Beyond Riemannian: the affine connection setting for transformation groups. In X. Pennec, S. Sommer, and T. Fletcher, editors, *Riemannian Geometric Statistics in Medical Image Analysis*, number Chap. 5, pages 169–229. Academic Press, Jan. 2020.
52. X. Pennec, S. Sommer, and P. T. Fletcher. *Riemannian Geometric Statistics in Medical Image Analysis*. Elsevier, 2020.

53. J.-M. Peyrat, H. Delingette, M. Sermesant, and X. Pennec. Registration of 4D time-series of cardiac images with multichannel diffeomorphic demons. In D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008*, volume 5242 of *LNCS*, pages 972–979. Springer, Heidelberg, Sep 2008.
54. M.-M. Rohé, N. Duchateau, M. Sermesant, and X. Pennec. Combination of Polyaffine Transformations and Supervised Learning for the Automatic Diagnosis of LV Infarct. In *Statistical Atlases and Computational Modeling of the Heart (STACOM 2015)*, Munich, Germany, 2015.
55. C. Seiler, X. Pennec, and M. Reyes. Geometry-aware multiscale image registration via OBBTree-based polyaffine log-demons. In G. Fichtinger, A. Martel, and T. Peters, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, volume 6893 of *LNCS*, pages 631–638. Springer, Heidelberg, Sep 2011.
56. C. Seiler, X. Pennec, and M. Reyes. Capturing the Multiscale Anatomical Shape Variability with Polyaffine Transformation Trees. *Medical Image Analysis*, 16(7):1371–1384, 2012.
57. A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.
58. T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Non-parametric diffeomorphic image registration with the Demons algorithm. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*, pages 319–326, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
59. T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Symmetric Log-domain diffeomorphic registration: A Demons-based approach. In D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, editors, *Proc. of Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008*, volume 5241 of *Lecture Notes in Computer Science - LNCS*, pages 754–761. Springer, Heidelberg, Sep 2008.
60. T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage*, 45(1):S61–S72, Mar. 2009.
61. L. Younes. *Shapes and diffeomorphisms*. Number v. 171 in Applied mathematical sciences. Springer, Heidelberg ; New York, 2010.