

Constructing an atlas for the diffeomorphism group of a compact manifold with boundary, with application to the analysis of image registrations

Carole J. Twining^a, Stephen Marsland^{b,*}

^a *Imaging Science and Biomedical Engineering (ISBE), University of Manchester, Manchester, UK*

^b *Institute of Information Sciences, Massey University, Palmerston North, New Zealand*

Received 11 September 2006; received in revised form 16 May 2007

Abstract

This paper considers the problem of defining a parameterization (chart) on the group of diffeomorphisms with compact support, motivated primarily by a problem in image registration, where diffeomorphic warps are used to align images. Constructing a chart on the diffeomorphism group will enable the quantitative analysis of these warps to discover the normal and abnormal variation of structures in a population.

We construct a chart for particular choices of boundary conditions on the space on which the diffeomorphism acts, and for a particular class of metrics on the diffeomorphism group, which define a class of diffeomorphic interpolating splines. The geodesic equation is computed for this class of metrics, and we show how it can be solved in the spline representation. Furthermore, we demonstrate that the spline representation generates submanifolds of the diffeomorphism group, and we study this mapping. Explicit computational examples are included, showing how this chart can be constructed in practice, and that the use of the geodesic distance allows better classification of variation than those obtained using just a Euclidean metric on the space of warps.

© 2007 Elsevier B.V. All rights reserved.

MSC: 58D05; 53C22; 62H35

Keywords: Diffeomorphism group; Geodesics; Chart; Image registration

1. Introduction and overview

There has been much interest in the group of volume-preserving diffeomorphisms since Arnold's celebrated discovery [2] that the Euler fluid equations describe geodesics on the group of volume-preserving (i.e. incompressible) diffeomorphisms. This paper considers the construction of local charts on the group of all compactly-supported diffeomorphisms and the chart extension to form a coordinate system or atlas. We will show that the Euler equations for diffeomorphisms arise naturally from the construction of a right-invariant Riemannian metric on the diffeomorphism group. These Euler equations appear in several different fields, coinciding with the Camassa-Holm

* Corresponding address: Institute of Information Sciences & Technology, Massey University, Private Bag 11-222, Palmerston North, New Zealand. Tel.: +64 6 350 5799; fax: +64 6 350 2259.

E-mail addresses: carole.twining@manchester.ac.uk (C.J. Twining), S.R.Marsland@massey.ac.nz (S. Marsland).

wave equation in fluid dynamics for the H^1 metric [13] and being used in two and three dimensions in the field of image registration.

It is this second application that provides the principal motivation for our work. We provide computational examples of how our methods can be used in Section 6, where we show that elements of the diffeomorphism group can be approximated to arbitrary accuracy on pixelated images based on a relatively small number of knotpoints, and Section 7 where we demonstrate that the relevant metric enables legal and illegal examples of 2D images of ventricles to be reliably classified.

Image warping is concerned with applying non-linear warps to an image in order to make structures in the image line up with corresponding structures in another (reference) image. There are a variety of applications, but the primary one is in medical imaging, where it is hoped that aligning images will enable the automated diagnosis of disease through analysis of the resultant deformation field. Under the assumption that registration between medical images should define a bijective, continuous, and invertible mapping between all points in the images, suitable warps come from some diffeomorphism group. Furthermore, in medical imaging the structures being warped are discrete, bounded entities. This suggests that the image warps should also be discrete and bounded, and so we restrict ourselves to some group of diffeomorphisms with compact support, $\mathcal{G} = \text{Diff}_0(\mathcal{M})$. While there are places where such modelling choices are unsuitable (for example, where additional structure such as a tumour is seen in just one of the images) there are many other problems where they appear to be valid, such as degenerative brain diseases like Alzheimer's. For a general review of registration methods, and other applications, see [31], and for an overview of medical image registration, see [28].

Much of the most relevant work to this paper comes under the rubric of Computational Anatomy (see [23] for an overview), where the problem of inexact matching for landmarks, shapes and images is viewed as constructing orbits from a template under diffeomorphic transformations. Following [9,29], the problem of finding the minimum-distance image deformation under inexact matching is shown to have the same solution as the Euler–Lagrange equations, with methods such as geodesic shooting [22] used to construct an approximation to the true diffeomorphism. This differs from our approach where we consider only exact matching. In terms of our target application of analysing image deformation fields, this means that we analyse the exact field, not some approximation to it. While it may be that in the presence of noise the inexact matching provides a smoother deformation field, the derivation of the theory is far clearer when exact matching is considered. We also derive and consider the metric on the whole of $\text{Diff}_0(\mathcal{M})$, rather than the metric on sets of landmarks (which is induced by the full metric) – see Section 5 for further details.

Further, this paper considers solving for geodesics on the diffeomorphism group within a spline representation, which is subtly but significantly different to the template matching approach of solving for geodesics on the space of knotpoint parameters using the induced metric [30,11,23].

The secondary purpose of our paper is to introduce the mathematics of the diffeomorphism group to those whose expertise is in the *application* of the methods in fields such as image registration. We therefore assume in the reader only knowledge of some standard differential geometry and functional calculus.

1.1. Problem statement

We consider the diffeomorphism group $\mathcal{G} = \text{Diff}_0^s(\mathcal{M})$ of compactly-supported Sobolev H_0^s -mappings,¹ with identity element e . This group is a smooth (C^∞) Hilbert space, but it is *not* a Lie group [25,10]. In general, while these spaces are not Lie groups (group operations are not necessarily smooth between the same spaces except in the limit as s tends to infinity), the diffeomorphism groups do have structures corresponding to the Lie algebras and group exponential maps of Lie groups (see Section 2). This group exponential map is only continuous; there is no guarantee that any diffeomorphism in a neighbourhood of the identity can be embedded into a flow of vector fields [15,10].

The group acts on a space \mathcal{M} , of dimension $n < 2s$, which we will take to be a compact space with boundary $\partial\mathcal{M}$. Points on the boundary $\partial\mathcal{M}$ do not move, since the diffeomorphisms have compact support. We take \mathcal{M} to be the closed unit ball \bar{B} in \mathbb{R}^n . This simplifies some of the computations considerably, and fits well with the target application – the study of diffeomorphic warps of images – as the image plane/volume can be scaled to lie wholly within the unit ball.

¹ $H_0^s(\mathcal{M})$ is the space of functions f which vanish on the boundary $\partial\mathcal{M}$ (the meaning of the 0 subscript), and whose derivatives up to order s are square-integrable functions on \mathcal{M} .

The metric in the image plane/volume will be taken to be Euclidean, with global Cartesian coordinates $\{x_\mu : \mu = 1, \dots, n\}$. The action of $g \in \text{Diff}_0^s(\mathcal{M})$ is given by $g : x \mapsto g(x) \in \mathcal{M}$, with coordinates $\{g_\mu(x)\}$. The composition² \circ of two elements g, h is given by $(g \circ h)(x) = g(h(x))$. The relevant vector space (of Eulerian velocity fields) on \mathcal{M} is the (product of) Sobolev space(s) $\mathfrak{X}_0^s(\mathcal{M})$, where the vector fields are s -times differentiable, and vanish on the boundary $\partial\mathcal{M} = \mathbb{S}^{n-1}$.

1.2. Overview of the paper

As one of the purposes of this paper is to survey the relevant results for non-specialists, we first run through some standard results (Sections 2, 3 and 4 up to Section 4.1). Useful reviews of the area for those wishing to know more are [26] and chapter 9 of [17].

The key result for the solution of geodesics on the diffeomorphism group within the spline representation is given in Section 4.2. We follow the usual approach to constructing a chart, beginning by considering the tangent space at the identity $T_e\mathcal{G}$. An exponential map from this tangent space to the group is then constructed (see Section 2), such that it is locally surjective from some neighbourhood of zero in the tangent space to some neighbourhood of the identity in the group. Inverting this via the corresponding log map gives the chart for this neighbourhood. Translating this construction across the group in some appropriate manner then provides a chart construction based on any point in the group, and hence an atlas, where the parameter space is the space of vector fields $\mathfrak{X}_0^s(\mathcal{M})$.

Because $\text{Diff}_0^s(\mathcal{M})$ is not a Lie group, the mapping from the tangent space into the diffeomorphism group is the Riemannian exponential map, not the group exponential map. A suitable right-invariant metric is therefore developed in Section 3, and the geodesic equation is computed (Section 4). Following this, in Section 4.2 we introduce a representation of $\text{Diff}_0^s(\mathcal{M})$ based on a spline basis (geodesic interpolating splines) and show that for this representation the geodesic equations for the velocity field reduce to the Euler–Lagrange equation of the spline action.

We then consider the construction of elements of $\text{Diff}_0^s(\mathcal{M})$ using the spline basis in Section 5. We demonstrate that from the space of knotpoint positions and motions that define the spline, the construction generates a mapping to submanifolds of $\text{Diff}_0^s(\mathcal{M})$. This is followed in Section 5.2 by the explicit computation of the chart at an arbitrary point p . We describe one particular family of Green’s functions in Section 5.3, the clamped-plate equation of Boggio [4]. This has both Dirichlet and von-Neumann boundary condition on the unit ball in \mathbb{R}^n . The resultant geodesic interpolating clamped-plate spline produces warps that are guaranteed to belong to $\text{Diff}_0^s(\mathcal{M})$.

2. The group exponential map

For a finite-dimensional Lie Group, the group exponential map is obtained by exponentiating elements of the Lie algebra \mathfrak{g} , and corresponds to the flow of left-invariant vector fields on the group. The Lie bracket is obtained by considering the Lie derivative of a pair of such left-invariant vector fields.

Moving to the infinite-dimensional case, left- and right-invariant vector fields are defined as follows. Consider a vector $X \in T_e\mathcal{G}$ that is the tangent vector to some smooth curve $\lambda : \mathbb{R} \mapsto \lambda(s)$ at the point $\lambda(0) = e$. Left and right multiplication by some $g \in \text{Diff}_0^s(\mathcal{M})$ then gives the associated mappings of this curve:

$$L_g : \lambda(s) \mapsto g \circ \lambda(s) = L_g(\lambda(s)), \quad R_g : \lambda(t) \mapsto \lambda(t) \circ g = R_g(\lambda(t)). \tag{1}$$

The first derivative (tangent mapping) of these mappings gives us the left and right extensions of X , L_X and R_X on \mathcal{G} , where:

$$L_X(g) = \left. \frac{d}{ds} \right|_{s=0} L_g(\lambda(s)), \quad L_X(g)(x)|_\mu = \frac{\partial g_\mu(x)}{\partial x_\nu} X_\nu(x), \tag{2}$$

$$R_X(g) = \left. \frac{d}{dt} \right|_{t=0} R_g(\lambda(t)) = X \circ g = R_g(X), \tag{3}$$

$$R_X(g)(x)|_\mu = X_\mu(g(x)) \implies R_X(g \circ f) = R_X(g) \circ f \forall f \in \mathcal{G}. \tag{4}$$

²Note that the definition of composition of group elements differs between authors (e.g. see [6,29]), so that others ‘left’ (in the sense of (2) and (4)) may mean our ‘right’ and vice versa.

As before, the indices μ, ν, \dots refer to components in the global Cartesian coordinate basis on \mathcal{M} . For the case of right multiplication, we see that $TR_g = R_g$, hence right group multiplication is C^∞ for any $g \in \text{Diff}_0^s(\mathcal{M})$. Whereas for left multiplication, we have that $TL_g = L_{Tg}$. If g is H_0^s , constructing $L_X(g)$ “uses up” a derivative (2), so that the result is only H_0^{s-1} . It then follows that left multiplication viewed as the mapping $\text{Diff}_0^s(\mathcal{M}) \times \text{Diff}_0^s(\mathcal{M}) \mapsto \text{Diff}_0^s(\mathcal{M})$, is only continuous (see [26], section 3.3.1. for further details).

The tangent space algebra³ \mathfrak{g} of $\mathcal{G} = \text{Diff}_0^s(\mathcal{M})$ is generated by considering right-invariant vector fields. Following [1], this algebra is given by considering the Lie derivative of R_Y along R_X :

$$(\mathfrak{L}_{R_X} R_Y)(e) = [R_X, R_Y](e) = -[X, Y]_{\mathfrak{g}} = [X, Y]_{\mathcal{M}}, \tag{5}$$

where $[\cdot, \cdot]_{\mathfrak{g}}$ is the commutator bracket defined by the Lie derivative of *left*-invariant vector fields on $\text{Diff}_0^s(\mathcal{M})$, hence the minus sign, and $[\cdot, \cdot]_{\mathcal{M}}$ is the usual commutator bracket for vector fields on \mathcal{M} :

$$[X, Y]_{\mathcal{M}}(x)|_{\mu} = \sum_{\nu} \left(X_{\nu}(x) \frac{\partial Y_{\mu}(x)}{\partial x_{\nu}} - Y_{\nu}(x) \frac{\partial X_{\mu}(x)}{\partial x_{\nu}} \right). \tag{6}$$

This tangent space algebra is not closed – again we “lose” derivatives, so that if $X, Y \in \mathfrak{X}_0^{s+1}(\mathcal{M})$, then $[X, Y]_{\mathfrak{g}}$ is only in $\mathfrak{X}_0^s(\mathcal{M})$.

The group exponential map $\exp_{\mathfrak{g}} : T_e \text{Diff}_0^s(\mathcal{M}) = \mathfrak{X}_0^s(\mathcal{M}) \mapsto \text{Diff}_0^s(\mathcal{M})$ is then defined by considering the flow of a right-invariant vector field R_X as follows. Consider $X \in T_e \text{Diff}_0^s(\mathcal{M})$ and $R_X(g)$, as defined above, and let $t \mapsto \phi(t)$ be the C^1 curve in $\text{Diff}_0^s(\mathcal{M})$ that is an integral curve of R_X :

$$\phi(t)(x) = \phi(t, x), \quad \frac{\partial}{\partial t} \phi(t, x) = \dot{\phi}(t, x) = R_X(\phi(t))(x) = X(\phi(t, x)). \tag{7}$$

This corresponds to the Eulerian description of flow on \mathcal{M} . A particle that starts at $x_0 \in \mathcal{M}$ follows a path $x(t)$ under the flow, where:

$$\phi(0, x_0) = x(0) = x_0, \quad x(t) = \phi(t, x_0), \quad \frac{dx_{\mu}(t)}{dt} = \dot{\phi}_{\mu}(t, x_0) = X_{\mu}(x(t)). \tag{8}$$

The group exponential map is then given by

$$\exp_{\mathfrak{g}} : \mathfrak{X}_0^s(\mathcal{M}) \mapsto \text{Diff}_0^s(\mathcal{M}), \quad \exp_{\mathfrak{g}}(tX)(x_0) = x(t) = \phi(t, x_0). \tag{9}$$

The problem is that there is no neighbourhood of the identity in $\text{Diff}_0^s(\mathcal{M})$ for which this group exponential mapping is surjective. In fact, there are elements of the group infinitesimally close to the identity that cannot be reached by the group exponential, as is discussed in [17], page 456, and references therein.

For a given curve $\phi(t)$, we can also consider the associated Eulerian spatial velocity field $v^{\phi} \in T\mathcal{M}$, which is defined as the right logarithmic derivative:

$$v^{\phi}(t) = \dot{\phi}(t) \circ \phi^{-1}(t), \quad v_{\mu}^{\phi}(t, x) = \dot{\phi}_{\mu}(t, \phi^{-1}(t, x)). \tag{10}$$

It can be seen that this spatial velocity field is explicitly right-invariant.

If $\phi(t) \Rightarrow \phi(t) \circ g$ then:

$$v^{\phi \circ g} = (\dot{\phi}(t) \circ g) \circ (g^{-1} \circ \phi^{-1}(t)) = \dot{\phi}(t) \circ \phi^{-1}(t) = v^{\phi} \quad \forall g \in \text{Diff}_0^s(\mathcal{M}). \tag{11}$$

The group exponential map $\exp_{\mathfrak{g}}(tX)$ therefore corresponds to the Eulerian flow of a *static* vector field:

$$v(t, x) = X(x). \tag{12}$$

It is possible to use such *static* vector fields to parameterise diffeomorphisms [3], which when integrated yield one-parameter subgroups of diffeomorphisms. However, the corresponding logarithm mapping is only well defined for

³ We use the phrase ‘tangent space algebra’ rather than Lie algebra to emphasize the fact that this diffeomorphism group is not a Lie group.

transformations close enough to the identity, although this may not be problematic for some image registration applications.

The alternative and more general approach follows from the observation that the above analysis suggests that to find an exponential map that is locally surjective, we should instead consider the flow of time-varying Eulerian velocity fields. One way to do this is by constructing a Riemannian metric on the diffeomorphism group. Then, as in standard finite-dimensional Riemannian geometry, the Riemannian exponential map of $X \in T_e \text{Diff}_0^s(\mathcal{M})$ is given by constructing a geodesic with X as the initial tangent vector.

3. Constructing Riemannian metrics on $\text{Diff}_0^s(\mathcal{M})$

We will take the standard construction of right-invariant Riemannian metrics on a diffeomorphism group, starting with a bilinear scalar product $\langle \cdot, \cdot \rangle_x$ on the Hilbert space of vector fields $\mathfrak{X}_0^s(\mathcal{M}) \times \mathfrak{X}_0^s(\mathcal{M})$ that is isomorphic to a scalar product on $T_e \text{Diff}_0^s(\mathcal{M}) \times T_e \text{Diff}_0^s(\mathcal{M})$, which we will denote by $\langle \cdot, \cdot \rangle_e$.

A Riemannian metric $\langle \cdot, \cdot \rangle_x$ on \mathcal{M} induces such a scalar product by integration:

$$\langle X, Y \rangle_e = \langle X, Y \rangle_x = \int_{\mathcal{M}} d\mu(X(x), Y(x))_x, \tag{13}$$

where $d\mu$ is the corresponding volume element on \mathcal{M} .

Given a scalar product on the space of velocity fields, we then define a scalar product at any point in $\text{Diff}_0^s(\mathcal{M})$ by right-translation. So, for $X_g, Y_g \in T_g \text{Diff}_0^s(\mathcal{M})$:

$$\langle X_g, Y_g \rangle_g = \langle X_g \circ g^{-1}, Y_g \circ g^{-1} \rangle_e. \tag{14}$$

For any curve $\phi(t) \in \text{Diff}_0^s(\mathcal{M})$, we have the associated energy/action functional:

$$E[\phi] = \int_0^1 dt \langle \dot{\phi}(t), \dot{\phi}(t) \rangle_{\phi(t)} = \int_0^1 dt \langle \dot{\phi}(t) \circ \phi^{-1}(t), \dot{\phi}(t) \circ \phi^{-1}(t) \rangle_e \tag{15}$$

$$= \int_0^1 dt \langle v^\phi(t), v^\phi(t) \rangle_x = \mathcal{L}[v^\phi], \tag{16}$$

where \mathcal{L} is the associated Lagrangian on the space of Eulerian velocity fields. We see from (11) that this action, and hence the associated Riemannian metric, is explicitly right invariant.

3.1. Metrics based on differential operators

In previous work on diffeomorphism groups and their physical applications (e.g. [10,26,21]), authors have considered Riemannian metrics on the group of volume-preserving diffeomorphisms, with the metric induced by a metric on \mathcal{M} . However, this is not the only possible choice; we can instead construct a scalar product on the Hilbert space of velocity fields that is not derivable from a metric on \mathcal{M} . In particular, we will consider scalar products of the form:

$$\begin{aligned} (u, v)_L &= \int_{\mathcal{M}} dx (Lv(x)) \cdot (Lu(x)) \\ &= \int_{\mathcal{M}} dx v(x) \cdot (L^\dagger L)u(x) \quad \forall u, v \in \mathfrak{X}_0^s(\mathcal{M}), \end{aligned} \tag{17}$$

where dx is the Euclidean flat-space volume element on \mathcal{M} , and \cdot is the flat-space vector scalar product. L (with Lagrange dual L^\dagger) is taken to be some scalar differential operator of appropriate non-zero order. Such a scalar product (the H^1 scalar product) has been considered previously by Shkoller [27] for the case of volume-preserving diffeomorphisms, by Kouranbaeva [16], for the case of the diffeomorphism group of either the circle or the line, where it generates the Camassa-Holm equation, and for the higher-dimensional Camassa-Holm equations in [12].

Note that we have so far considered curves in the space of Eulerian spatial velocity fields $\mathfrak{X}_0^s(\mathcal{M})$ derived from some curve $\phi(t)$ in the corresponding diffeomorphism group. The corresponding inverse link between the space of

velocity fields and the diffeomorphism group was established by Dupuis et al. [9], who showed that any curve $v(t)$ in the space of Eulerian spatial velocity fields $\mathfrak{X}_0^s(\mathcal{M})$ for which $\int_0^\tau dt (v(t), v(t))_L < \infty$ generates a corresponding curve on the diffeomorphism group $\text{Diff}_0^s(\mathcal{M})$. We can therefore, without loss of generality, work in the space of Eulerian spatial velocity fields $\mathfrak{X}_0^s(\mathcal{M})$.

4. The Geodesic equation on the diffeomorphism group

The geodesic equation on the diffeomorphism group for the right-invariant Riemannian metric can be derived from the scalar product defined in (17). This is a simplified version of the variational calculus techniques involved — for more details, readers should consult more specialist works such as [12]. An alternative form of the equation that makes the link to the Euler equations more specific is given in Section 4.1.

Theorem 1. *Given a scalar product of the form (for L a scalar differential operator)*

$$(u, v)_L = \int_{\mathcal{M}} dx v(x) \cdot (L^\dagger L)u(x) \quad \forall u, v \in \mathfrak{X}_0^s(\mathcal{M}), \tag{18}$$

that is induced by a right-invariant Riemannian metric on the diffeomorphism group $\text{Diff}_0^s(\mathcal{M})$, then the geodesic equation can be written as:

$$(L^\dagger L)\dot{v}_\mu + \sum_\alpha \left(\frac{\partial}{\partial x_\alpha} (v_\alpha (L^\dagger L)v_\mu) + \frac{\partial v_\alpha}{\partial x_\mu} (L^\dagger L)v_\alpha \right) = 0, \tag{19}$$

where $\{x_\alpha\}$ are global cartesian coordinates on \mathcal{M} .

Proof. Consider a curve $\phi(t) \in \text{Diff}_0^s(\mathcal{M})$, with the associated action (from (15)):

$$E[\phi] = \sum_\mu \int_0^1 dt \int_{\mathcal{M}} dx (Lv_\mu(t, x))(Lv_\mu(t, x)), \tag{20}$$

where $v_\mu(t, x) = \dot{\phi}_\mu(t, \phi^{-1}(t, x))$. A geodesic between the points $\phi(0)$ and $\phi(1)$ is, by definition, an extremum of this action. Varying the curve $\phi(t)$ whilst keeping the endpoints $\phi(0)$ and $\phi(1)$ fixed, gives an infinitesimal variation:

$$\phi(t, y) \Rightarrow \phi(t, y) + \epsilon(t, y), \tag{21}$$

where $\epsilon(t, \partial\mathcal{M}) = 0 \forall t \in [0, 1]$ and $\epsilon(0, y) = \epsilon(1, y) = 0 \forall y \in \mathcal{M}$.

We need to find the corresponding variation in the velocity field $v_\mu(t, x)$ induced by the variation in ϕ for fixed x and t . Define $y = \phi^{-1}(t, x)$. Then:

$$x_\mu = \phi_\mu(t, y) = \phi_\mu(t, y + \Delta y) + \epsilon_\mu(t, y) + \text{higher order terms} \tag{22}$$

$$\Rightarrow 0 = \sum_v \Delta y_v \left. \frac{\partial}{\partial y_v} \right|_t \phi_\mu(t, y) + \epsilon_\mu(t, y) \tag{23}$$

$$\Rightarrow \Delta y_\mu = - \sum_v M_{\mu v}(t, y) \epsilon_v(t, y), \tag{24}$$

where $(M^{-1})_{\mu v} = \left. \frac{\partial}{\partial y_v} \right|_t \phi_\mu(t, y)$. The variation in the Eulerian velocity is:

$$\Delta v_\mu(t, x) = \dot{\epsilon}_\mu(t, y) + \sum_\alpha \left(\frac{\partial v_\mu(t, x)}{\partial x_\alpha} \right)_t \epsilon_\alpha(t, y), \tag{25}$$

where $\dot{\epsilon}(t, y) = \left. \frac{\partial}{\partial t} \right|_y \epsilon(t, y)$ and $y = \phi^{-1}(t, x)$. Setting

$$\eta_\mu(t, x) = \epsilon_\mu(t, y) = \epsilon_\mu(t, \phi^{-1}(t, x)), \tag{26}$$

gives the variation of the velocity field as:

$$\Delta v_\mu(t, x) = \dot{\eta}_\mu(t, x) + \sum_\alpha \left(\frac{\partial \eta_\mu}{\partial x_\alpha} v_\alpha(t, x) - \frac{\partial v_\mu}{\partial x_\alpha} \eta_\alpha(t, x) \right), \tag{27}$$

$$\text{or equivalently, } \Delta v = \dot{\eta} + [\eta, v]_{\mathfrak{g}}, \tag{28}$$

where the latter form is just the Lin constraints [24] on the allowed variation of a Eulerian spatial velocity field. The boundary conditions on the redefined variation $\eta(t, x)$ are the same as those on the original variation $\epsilon(t, y)$ (21).

Inserting this into the expression for the action (20) provides the first-order variation of the action. Note that all functions now depend on x and t , so that the differential operator L contains only derivatives w.r.t. x , and $\eta(t, x)$ has appropriate boundary conditions, so that both spatial and temporal derivatives can be passed across as required.

$$\begin{aligned} \frac{1}{2} \Delta E &= - \sum_\mu \int_0^1 dt \int_{\mathcal{M}} dx \eta_\mu (L^\dagger L) \dot{v}_\mu - \sum_{\mu, \alpha} \int_0^1 dt \int_{\mathcal{M}} dx \eta_\mu \frac{\partial}{\partial x_\alpha} (v_\alpha (L^\dagger L) v_\mu) \\ &\quad - \sum_{\mu, \alpha} \int_0^1 dt \int_{\mathcal{M}} dx \eta_\alpha \frac{\partial v_\mu}{\partial x_\alpha} (L^\dagger L) v_\mu. \end{aligned} \tag{29}$$

$$\implies -\frac{1}{2} \Delta E = \int_0^1 dt (\eta(t), \dot{v}(t))_L + ([v(t), \eta(t)]_{\mathfrak{g}}, v(t))_L. \tag{30}$$

Applying the fundamental lemma of variational calculus to (29) leads directly to (19). \square

Introducing the standard momentum field $m_\mu \doteq (L^\dagger L)v_\mu$, into (19) leads to the Euler equations in the form known as EPDiff by Holm et al. [11]. In contrast to their approach, we choose to retain the velocity field as our functional variable. By considering an explicit representation of the space of velocity fields, we are able to solve the geodesic equation for the velocity field whilst remaining in the full diffeomorphism group. We choose to use a spline representation in Section 4.2, but we first rederive the geodesic equation using the adjoint representation, generating the Euler–Poincare equation of (16).

4.1. The adjoint representation of the tangent-space algebra

Corollary 2. *If we assume that the adjoint of the adjoint exists, which is not guaranteed in infinite dimensions [20], then we can then rewrite (30) as:*

$$-\frac{1}{2} \Delta E = \int_0^1 dt (\eta(t), \dot{v}(t))_L + (\eta(t), ad(v(t))^* v(t))_L, \tag{31}$$

and by applying the fundamental lemma of variational calculus, obtain the geodesic equation in the form

$$-\dot{v}(t) = ad(v(t))^* v(t), \tag{32}$$

which is the Euler–Poincare equation of the associated Lagrangian (16) for right-invariant, infinite-dimensional systems as used by other authors (e.g. [18,21,12]).

Proof. As in the case of finite-dimensional Lie Groups, the adjoint representation of the tangent-space algebra is defined as

$$ad(u)(w) = [u, w]_{\mathfrak{g}} \quad \forall u, w \in \mathfrak{X}_0^s(\mathcal{M}), \tag{33}$$

where $[\cdot, \cdot]_{\mathfrak{g}}$ is defined in (5) and (6). $ad(u)$ can then be considered as a mapping of the space of vector fields onto itself. If it exists, the adjoint $*$ of this operator with respect to the scalar product $(\cdot, \cdot)_L$ is given by:

$$(u, ad(v)(w))_L = (ad(v)^* u, w)_L \quad \forall u, v, w \in \mathfrak{X}_0^s(\mathcal{M}). \quad \square \tag{34}$$

4.2. The spline representation

We introduce a spline representation (the Green’s function of the operator $L^\dagger L$) and demonstrate that in this representation the geodesic equations (19) derived in Section 4 reduce to the Euler–Lagrange equation of the knotpoint action (52). This means that paths in the infinite-dimensional group $\text{Diff}_0^s(\mathcal{M})$ can be parameterized by the finite-dimensional space of knotpoints positions, with a Riemannian metric (55), resulting in the usual affinely parameterized geodesic. Alternative spline representations based on continuous objects, e.g. curves, can be constructed, but we consider the simplest case of splines based on a finite set of points.

Theorem 3. Consider the Green’s function $G(x, x')$ defined by:

$$(L^\dagger L)G(x, x') = \delta(x - x'). \tag{35}$$

(Note that the above definition of $G(x, x')$, and the geodesic equation (19) involve just the combination of operators $L^\dagger L$, which is now a self-dual differential operator. Hence what follows is valid for situations where an appropriate $L^\dagger L$ exists, but L does not — explicit examples of this are given in Section 5.3.)

The velocity fields can be written as

$$v_\mu(t, x) = \sum_i a_{i\mu}(t)G(x, b_i(t)), \quad a_i(t), b_i(t) \in \mathcal{M}, \tag{36}$$

where the set of functions $\{a_{i\mu}(t), b_{iv}(t)\}$ can be regarded as the parameters of the velocity field within this spline basis. The geodesic equations (19) for the velocity field reduce to the Euler–Lagrange equation of the action:

$$\dot{a}_{i\mu}(t) + \sum_{j,\alpha} a_{i\alpha}(t)a_{j\alpha}(t) \frac{\partial G(b_i, b_j)}{\partial b_{i\mu}} = 0. \tag{37}$$

Proof. The parameterized velocity field (36) will lie in the correct space $\mathfrak{X}_0^s(\mathcal{M})$ if the Green’s function has the corresponding appropriate boundary conditions. Substituting from (36) into (19) provides:

$$\begin{aligned} 0 &= \sum_i \dot{a}_{i\mu}(t)\delta(x - b_i(t)) + \sum_{i,v} a_{i\mu}(t)\dot{b}_{iv}(t) \frac{\partial}{\partial b_{iv}(t)} \delta(x - b_i(t)) \\ &+ \sum_{i,j,\alpha} a_{i\mu}(t)a_{j\alpha}(t) \left(\frac{\partial G(x, b_j(t))}{\partial x_\alpha} \delta(x - b_i(t)) + G(x, b_j(t)) \frac{\partial}{\partial x_\alpha} \delta(x - b_i(t)) \right) \\ &+ \sum_{i,j,\alpha} a_{i\alpha}(t)a_{j\alpha}(t) \frac{\partial G(x, b_j(t))}{\partial x_\mu} \delta(x - b_i(t)). \end{aligned} \tag{38}$$

Equating the coefficient of the $\frac{\partial}{\partial b_{iv}(t)} \delta(x - b_i(t))$ term to zero⁴ gives:

$$a_{i\mu}(t)\dot{b}_{iv}(t) - \sum_j a_{i\mu}(t)a_{jv}(t)G(b_i(t), b_j(t)) = 0 \quad (\text{no sum on } i). \tag{39}$$

So, either $a_{i\mu}(t) = 0$, or:

$$\dot{b}_{iv}(t) = \sum_j a_{jv}(t)G(b_i(t), b_j(t)) = \sum_j G_{ij}(t)a_{jv}(t) = v_v(t, b_i(t)), \tag{40}$$

⁴ This procedure is valid because of the particular form of the integral equation (30). From (38), the general form (with test function $f(x)$) is:

$$\int_{\mathcal{D}} dx f(x) \left(\phi(x)\delta(x - y) + A \frac{\partial}{\partial y} \delta(x - y) + \frac{\partial}{\partial x} [\psi(x)\delta(x - y)] \right) = 0, \quad f(\partial\mathcal{D}) = 0$$

where A does not depend on x . The solution is: $\phi(y) = 0$ and $\psi(y) = A$. An exactly equivalent form is:

$$\int_{\mathcal{D}} dx f(x) \left(\left[\phi(x) + \frac{\partial \psi(x)}{\partial x} \right] \delta(x - y) + [A - \psi(x)] \frac{\partial}{\partial y} \delta(x - y) \right) = 0.$$

Equating coefficients as in the text gives: $A = \psi(y) \Rightarrow \frac{\partial \psi(y)}{\partial y} = 0 \Rightarrow \phi(y) = 0$.

where we have defined the matrix $G_{ij}(t)$:

$$G_{ij}(b(t)) = G_{ij}(t) \doteq G(b_i(t), b_j(t)). \tag{41}$$

From Eq. (40), we see that the previously undefined parameters $\{b_i(t)\}$ are the coordinates of some set of particles that follow the flow. They also define the flow, in that their initial velocities define the initial velocity field. Following the usual usage in the spline literature, we will call them knotpoints, although it should be remembered that their real meaning is as particles that follow the flow.

The parameters $\{a_{i\mu}(t)\}$ can also be written in terms of the $\{b_i(t)\}$. Taking the inverse of the matrix $G_{ij}(t)$, we then have:

$$a_{i\mu}(t) = \sum_j G_{ij}^{-1}(t) \dot{b}_{j\mu}(t). \tag{42}$$

Equating the coefficient of the $\delta(x - b_i(t))$ term from (38) to zero gives:

$$\dot{a}_{i\mu}(t) + \sum_{j,\alpha} a_{i\mu}(t) a_{j\alpha}(t) \frac{\partial G(b_i, b_j)}{\partial b_{i\alpha}} + \sum_{j,\alpha} a_{i\alpha}(t) a_{j\mu}(t) \frac{\partial G(b_i, b_j)}{\partial b_{i\mu}} = 0. \tag{43}$$

Considering the second term on the right-hand side and observing that

$$\begin{aligned} \sum_j a_{j\alpha}(t) \frac{\partial G(b_i, b_j)}{\partial b_{i\alpha}} &= \frac{\partial}{\partial b_{i\alpha}} \sum_j a_{j\alpha}(t) G(b_i, b_j) \\ &= \frac{\partial}{\partial b_{i\alpha}} \dot{b}_{i\alpha}(t) \end{aligned} \tag{44}$$

$$= 0, \tag{45}$$

gives the result. \square

To see the meaning of this equation, consider the variation of

$$F = \sum_{i,j,\mu} \int_0^1 dt a_{i\mu}(t) G_{ij}(t) a_{j\mu}(t) \tag{46}$$

with respect to the knotpoint paths $\{b_{j\alpha}(t)\}$, where the definitions and relations in (40) hold. Then

$$\Delta a_{n\alpha} = \sum_{m,p,\beta} \frac{\partial G_{nm}^{-1}}{\partial b_{p\beta}} \dot{b}_{m\alpha} \Delta b_{p\beta} + \sum_m G_{nm}^{-1} \frac{d}{dt} \Delta b_{m\alpha}. \tag{47}$$

Writing $\frac{\partial G_{nm}^{-1}}{\partial b_{p\beta}} = -\sum_{r,s} G_{nr}^{-1} \frac{\partial G_{rs}}{\partial b_{p\beta}} G_{sm}^{-1}$ gives

$$\Delta a_{n\alpha} = \sum_m G_{nm}^{-1} \frac{d}{dt} \Delta b_{m\alpha} - \sum_{r,s,p,\beta} G_{nr}^{-1} \frac{\partial G_{rs}}{\partial b_{p\beta}} a_{s\alpha} \Delta b_{p\beta} \tag{48}$$

$$\Rightarrow \Delta F = \int_0^1 dt - 2 \sum_{i,\mu} \Delta b_{i\mu} \left(\dot{a}_{i\mu} + \frac{1}{2} \sum_{j,k,\alpha} a_{k\alpha} a_{j\alpha} \frac{\partial G_{jk}}{\partial b_{i\mu}} \right), \tag{49}$$

$\frac{\partial G_{jk}}{\partial b_{i\mu}} \equiv 0$ unless $i = j$ or $i = k$

$$\Rightarrow \sum_{j,k,\alpha} a_{j\alpha} a_{k\alpha} \frac{\partial G_{jk}}{\partial b_{i\mu}} = 2 \sum_{j,\alpha} a_{i\alpha} a_{j\alpha} \frac{\partial G_{ij}}{\partial b_{i\mu}}, \tag{50}$$

$$\therefore \Delta F = 0 \Rightarrow \dot{a}_{i\mu} + \sum_{j,\alpha} a_{i\alpha} a_{j\alpha} \frac{\partial G_{ij}}{\partial b_{i\mu}} = 0, \tag{51}$$

which is the same as the geodesic equation (37).

Corollary 4. *The energy/action of a path $y(t)$ in the finite-dimensional space of knotpoint positions (endowed with the Riemannian metric with components $\{g_{AB}(y)\}$) is the Euler–Lagrange equation of the action, which is the geodesic Eq. (19) for the velocity field in the spline representation. This can be written as*

$$F[y, \dot{y}] = \int_0^1 dt g_{AB}(y(t)) \dot{y}^A(t) \dot{y}^B(t), \quad (52)$$

using the multi-index tensor notation (with the Einstein summation convention for the multi-indices)

$$\text{If } \mathbf{A} \doteq (i\mu) \quad \text{and} \quad \mathbf{B} \doteq (j\nu) \\ y^{\mathbf{A}}(t) \doteq b_{i\mu}(t), \quad (53)$$

$$g_{AB}(y(t)) \doteq G_{ij}^{-1}(y(t)) \delta_{\mu\nu}, \quad (54)$$

$$g^{\mathbf{AB}}(y(t)) \doteq G_{ij}(y(t)) \delta_{\mu\nu}. \quad (55)$$

From (42)

$$a_{i\mu}(t) = \sum_j G_{ij}^{-1}(t) \dot{b}_{j\mu}(t) = g_{AB}(y(t)) \dot{y}^B(t). \quad (56)$$

Also, note that with the spline-interpolated velocity field as given in Eq. (36)

$$\begin{aligned} (v(t, \cdot), v(t, \cdot))_L &= \int_{\mathcal{M}} dx v(t, x) \cdot (L^\dagger L) v(t, x), \\ &= \sum_{i,\mu} a_{i\mu}(t) \dot{b}_{i\mu}(t) = g_{AB}(y(t)) \dot{y}^B(t) \dot{y}^A(t) = \left(\frac{ds}{dt} \right)^2, \end{aligned} \quad (57)$$

where ds^2 is the square of the infinitesimal metric distance. Hence metric distances in the space of knotpoint positions correspond with metric distances in $\text{Diff}_0^s(\mathcal{M})$.

So the geodesic equations are just the usual ones from finite-dimensional Riemannian geometry for an affinely parameterised geodesic⁵

$$\ddot{y}^D(t) + \Gamma_{AB}^D(y) \dot{y}^A(t) \dot{y}^B(t) = 0, \quad (58)$$

where Γ_{AB}^D are the usual Christoffel symbols defined by the metric

$$\Gamma_{AB}^D(y) = \frac{1}{2} g^{DC}(y) \left(\frac{\partial g_{AC}(y)}{\partial y^B} + \frac{\partial g_{BC}(y)}{\partial y^A} - \frac{\partial g_{AB}(y)}{\partial y^C} \right). \quad (59)$$

5. Geodesic interpolating splines (GIS) and sub-manifolds of $\text{Diff}_0^s(\mathcal{M})$

The spline representation that we introduced in the previous section is related to those introduced by Joshi and Miller [14] and Camion and Younes [6]. Both of these papers considered the problem of finding a diffeomorphic interpolant for a set of (non-coincident) knotpoints in some (compact or non-compact) space \mathcal{M} . Joshi and Miller imposed initial conditions on the initial positions and velocities of the knotpoints, $y(0)$ and $v(0)$, while Camion and Younes imposed initial and final positions $y(0)$, $y(1)$ on the points. In fact, as with finite dimensional Riemannian geometry, both are equivalent. To solve this interpolation problem, both papers introduce a scalar product on the space of Eulerian velocity fields generated by some differential operator, as in (17), and use the result of Dupuis et al. to ensure diffeomorphic mappings. However, the result of Dupuis et al. refers only to *compact* manifolds \mathcal{M} ; the

⁵ Note that we obtain an affinely-parameterized geodesic, with the path length a linear function of the time t , because we optimized the energy functional as opposed to the path length functional.

generalized thin-plate spline considered in both papers and the Gaussian interpolant considered by Camion and Younes do not generate mappings with compact support.

Our more general formulation, derived from the explicit Riemannian metric on the diffeomorphism group, highlights some other interesting features. Of particular interest is the fact that the knotpoint positions $\{b_{i\mu}(t)\}$, which were originally introduced in (36) as *arbitrary* parameters, move according to the flow (40) as a consequence of imposing the general geodesic equation on the diffeomorphism group (19), whereas in [6], the exact (or inexact) matching condition for the knotpoints was imposed *a priori*. This does not, however, imply that the quality of the image match is independent of the choice of knotpoints.

We next consider the GIS construction in the context of the diffeomorphism group $\text{Diff}_0^s(\mathcal{M})$ for the case of exact matching only. Starting from the space of knotpoint positions and motions, we show how the GIS construction generates a mapping to sub-manifolds of the diffeomorphism group, which is precisely the chart construction that we were originally seeking. By considering the tangent spaces to this submanifold, we show how this chart can be extended across the whole group to generate the final atlas.

5.1. The Space of knotpoint positions

Consider a set of N non-coincident⁶ knotpoints lying in the open unit ball B in \mathbb{R}^n , with the i th knotpoint having coordinates $y_i \doteq \{y_{i\mu} : \mu = 1, \dots, n\}$. A valid configuration of knotpoints maps to a point p with coordinates $\{y_i : i = 1, \dots, N\}$ in the space $\Omega^{(N)} = B^N/S$, where S is the set of coincident knotpoint positions (i.e. the set of all hyperplanes $y_i = y_j \forall i \neq j$.)

An element of the tangent space $T_p\Omega^{(N)}$ can be considered as a velocity vector v_i for each knotpoint. This maps to an initial velocity field $v(0, \cdot) \in \mathfrak{X}_0^s(\mathcal{M})$ on the whole of \mathcal{M} via the spline interpolant (36) and (42):

$$v_\mu(0, x) = \sum_{i,j} G(x, y_i) G_{ij}^{-1}(y) v_{j\mu}, \tag{60}$$

where $G_{ij}(y) = G(y_i, y_j)$ is the Green’s function (in matrix form, with matrix inverse $G_{ij}^{-1}(y)$). The velocity field given above is the unique interpolant of $\{v_i\}$ that minimizes the spline energy functional:

$$E[v] = \sum_\mu \int_{\mathcal{M}} dx v_\mu(x) L^\dagger L v_\mu(x) + \sum_{i=1}^N \lambda_i \|v_i - v(y_i)\|^2 \tag{61}$$

$$= (v, v)_L + \sum_{i=1}^N \lambda_i \|v_i - v(y_i)\|^2, \tag{62}$$

where the $\{\lambda_i\}$ are Lagrange multipliers.

We now map the point $p \in \Omega^{(N)}$ to the identity in $\text{Diff}_0^s(\mathcal{M})$. The interpolated velocity field $v(0, x) \in \mathfrak{X}_0^s(\mathcal{M})$ then becomes $v(0, x) \in T_e\text{Diff}_0^s(\mathcal{M})$, and we construct the geodesic in the diffeomorphism group corresponding to these initial conditions. For any velocity field v this defines the Riemannian exponential map:

$$\exp_R : \mathfrak{X}_0^s(\mathcal{M}) \mapsto \text{Diff}_0^s(\mathcal{M}), \quad \exp_R(tv) = \phi(t), \tag{63}$$

where $\phi(t)$ is the affinely-parameterized geodesic curve in $\text{Diff}_0^s(\mathcal{M})$ with the initial conditions $\phi(0) = e, \dot{\phi}(0, x) = v(0, x)$, which is in the spline representation of velocity fields. This representation is parameterized by a set of N knotpoints, a set of particles that follow the flow.

The number of knotpoints in the spline representation (N) is constant as we traverse the geodesic, as follows from the following observations:

- adding a knotpoint that does not follow the current velocity field increases the spline energy in (61);
- adding a new knotpoint that follows the current velocity field requires that the corresponding coefficient a_{N+1} is zero at time t (see (36)), and (40) means that a coefficient that is zero at time t is zero for all times after this;
- geodesics are reversible.

⁶ Since from (40) knotpoints are particles that follow and define the flow, coincident knotpoints must stay together or break the diffeomorphism constraint. Hence multiple coincident knotpoints are redundant.

In summary, a geodesic in $\text{Diff}_0^s(\mathcal{M})$ that starts from the identity, with the initial velocity described by N knotpoints in the spline representation, can be mapped to a unique continuous curve in $\Omega^{(N)}$, and this curve is actually a geodesic of the metric g_Ω (which is the Riemannian metric with components $g_{AB}(y)$: see (55)). This geodesic starts at the point $p = y_p$, where y_p corresponds to the positions of knotpoints required to describe the initial velocity of the geodesic in $\text{Diff}_0^s(\mathcal{M})$, and the tangent vector to the curve in $\Omega^{(N)}$ is just the velocities of the knotpoints as they follow the flow that corresponds to the geodesic in $\text{Diff}_0^s(\mathcal{M})$. So, any geodesic through the identity e in $\text{Diff}_0^s(\mathcal{M})$ maps to a geodesic in some space $\Omega^{(N)}$.

In the next section we start from $\Omega^{(N)}$, and show how we can construct charts on submanifolds of $\text{Diff}_0^s(\mathcal{M})$.

5.2. Chart based at p

Without loss of generality, we consider a specific fixed point $p \in \Omega^{(N)}$. By taking all possible values of the initial velocities for these knotpoints, we generate a family of geodesics in $\Omega^{(N)}$ with p as initial point, defined w.r.t. the metric g_Ω (55). This family of geodesics in $\Omega^{(N)}$ maps to a corresponding family of geodesics in $\text{Diff}_0^s(\mathcal{M})$, which all start from the identity, and span a sub-manifold in $\text{Diff}_0^s(\mathcal{M})$, meaning that any point $q \in \Omega^{(N)}$ can be reached (in unit time) by a geodesic from p . This geodesic maps to a corresponding geodesic in $\text{Diff}_0^s(\mathcal{M})$ that defines the mapping of the point $q \in \Omega^{(N)}$ to an element of $\text{Diff}_0^s(\mathcal{M})$ lying in the sub-manifold generated by p .

Similarly, any point $g \in \text{Diff}_0^s(\mathcal{M})$ lying on the submanifold generated by p can be reached, in unit time, from the identity e by some geodesic; this geodesic, by definition, lies wholly in the submanifold, and hence maps uniquely to a geodesic in $\Omega^{(N)}$. The components of the initial velocity at p of this geodesic in $\Omega^{(N)}$ provides the parameterization of g in this chart, whereas the coordinates at p define the particular chart.

Let $d_\Omega(\cdot, \cdot)$ denote the geodesic distance between points in $\Omega^{(N)}$ defined by the metric g_Ω . $d_G(\cdot, \cdot)$ is then the geodesic distance between points in $\text{Diff}_0^s(\mathcal{M})$, defined according to the metric on the diffeomorphism group. We will denote by \mathcal{N}_p the submanifold of $\text{Diff}_0^s(\mathcal{M})$ generated by p , with:

$$\text{Spl} : \Omega^{(N)} \mapsto \mathcal{N}_p, \quad \text{Spl}(p) \doteq e \tag{64}$$

being the mapping (with Spl being short for ‘spline’) between the space of knotpoints and the submanifold, as defined above. The results of the previous section can then be expressed as:

$$\forall q \in \Omega^{(N)}, \quad d_\Omega(p, q) = d_G(e, \text{Spl}(q)), \tag{65}$$

which is to say that geodesic distances from p are preserved under the mapping Spl. However, the metric g_Ω is defined on the whole of $\Omega^{(N)}$, so that for any two points $q, r \neq p$, we have the result

$$d_\Omega(q, r) = \inf_g \{d_G(e, g) | g(q) = r\}. \tag{66}$$

Any curve between q and r maps under Spl to a curve in \mathcal{N}_p between Spl(q) and Spl(r), but the geodesic in $\Omega^{(N)}$ between q and r does not necessarily correspond to the shortest such path joining Spl(q) and Spl(r) that lies wholly in the submanifold \mathcal{N}_p . Furthermore, the shortest geodesic linking Spl(q) and Spl(r) will not in general lie in the submanifold \mathcal{N}_p .

So the mapping Spl between $\Omega^{(N)}$ and $\text{Diff}_0^s(\mathcal{M})$ is far from straightforward. It uniquely maps points in $\Omega^{(N)}$ to points in \mathcal{N}_p , which is a submanifold of $\text{Diff}_0^s(\mathcal{M})$, and it maps closed curves to closed curves, but it only preserves distances for geodesics through p . However, the associated mapping

$$\widehat{\text{Spl}} : T_p \Omega^{(N)} \mapsto \Omega^{(N)} \mapsto \mathcal{N}_p \tag{67}$$

generated via geodesics of g_Ω that start at p does provide us with a chart for \mathcal{N}_p . It can be seen that it corresponds to the Riemannian map exp_R (63) when restricted to the subset of $\mathfrak{X}_0^s(\mathcal{M})$ that corresponds to the spline interpolant of elements of $T_p \Omega^{(N)}$. Furthermore, we can see how this chart can be extended; all we have to do is to add a new knotpoint to those that defined the point $p \in \Omega^{(N)}$. This then defines the point $p' \in \Omega^{(N+1)}$, the new submanifold $\mathcal{N}_{p'}$, where $\mathcal{N}_p \subset \mathcal{N}_{p'}$, and the new chart based at p' .

5.2.1. An alternative spline-based mapping

As an illustration of the fact that the choice of mapping between $\Omega^{(N)}$ and $\text{Diff}_0^s(\mathcal{M})$ generated by splines is far from straightforward, we here consider an alternative mapping.

Suppose we have a general, smooth curve $\lambda(t) \in \Omega^{(N)}$, with corresponding knotpoint coordinates $\{y_{i\mu}(t)\}$. At each point on the curve, we can construct the tangent vector:

$$\dot{\lambda}(t) = \{\dot{y}_{i\mu}(t)\}, \quad v_{i\mu}(t) = \dot{y}_{i\mu}(t),$$

which is just the velocities of the knotpoints when they are at position $\{y_{i\mu}(t)\}$. We can interpolate these velocities using the usual spline interpolant (60), to generate the Eulerian velocity field $v(t, x)$. The image of the curve $\lambda(t) \in \Omega^{(N)}$ under our new spline-based mapping is then defined to be any curve $\psi(t) \in \text{Diff}_0^s(\mathcal{M})$ such that the tangent vector to ψ at t is given by $v(t, \cdot)$. Distances along curves are preserved by this mapping, since the spline interpolant of velocity fields preserves the inner product of velocities upon which this distance is based (see Corollary 4, Eq. (57)).

This curve $\psi(t)$ is, however, not unique. Consider a general element $g \in \text{Diff}_0^s(\mathcal{M})$. The curves $\psi(t)$ and $\psi(t) \circ g$ have the same associated velocity fields (11). Hence this mapping, rather than mapping the curve $\lambda(t)$ to a single curve $\psi(t)$ instead maps to a family of curves that map into each other by right translation. This mapping does, however, preserve metric distances along all smooth curves, and hence maps all geodesics in $\Omega^{(N)}$ into geodesics in $\text{Diff}_0^s(\mathcal{M})$.

Further, suppose that we consider a piecewise smooth closed curve, for example one formed by three curves $\lambda_{pq}, \lambda_{qr}, \lambda_{rp}$ between three points $p, q, r \in \Omega^{(N)}$. If the mapping of the curve λ_{pq} ends at some point $\psi_{pq}(1) \in \text{Diff}_0^s(\mathcal{M})$. The mapping of the second curve λ_{qr} need not start at $\psi_{pq}(1)$, but instead at a point $\psi_{qr}(0) = \psi_{pq}(1) \circ f$, where f is any diffeomorphism that leaves the knotpoint positions at $q \in \Omega^{(N)}$ unchanged. Hence a composite closed curve in $\Omega^{(N)}$ does not in general map to families of closed curves in $\text{Diff}_0^s(\mathcal{M})$, since the corresponding curves in $\text{Diff}_0^s(\mathcal{M})$ need only be ‘closed’ with respect to the motions of the knotpoints, not all points $x \in \mathcal{M}$. This means that a point in $\Omega^{(N)}$ no longer maps to a single point in $\text{Diff}_0^s(\mathcal{M})$.

This demonstrates the importance of choosing the spline mapping correctly – $\widehat{\text{Spl}}$ (presented in (67)) preserves all closed curves, but only certain distances, whereas the one described in this section preserves all distances, but points and curves are no longer mapped into single points and curves. For our application of image matching, we consider the positions of landmarks, and so $\widehat{\text{Spl}}$ is sufficient, with the appropriate metric distance between patterns (i.e. configurations of knotpoints) being the metric d_Ω given in (66).

5.3. A family of Green’s functions with compact support

To produce practical applications of this theory, we need a Green’s function with compact support that generates a suitable spline. We take \mathcal{M} to be the closed unit ball \bar{B} in \mathbb{R}^n . As noted in [6], the scalar differential operator L appears in the formalism as the self-dual operator $L^\dagger L$, hence we can consider any such scalar self-dual differential operator $K = L^\dagger L$, even if, strictly speaking, L does not exist. To be specific, we consider operators of the form

$$K = (\nabla^2)^m = \Delta^m, \tag{68}$$

where, for the case of m odd, the corresponding operator L does not exist.

The Green’s functions for these operators with \mathcal{M} the unit ball in \mathbb{R}^n were given by Boggio [4] in 1905

$$G_{(m,n)}(p, q) \propto \|p - q\|^{2m-n} \int_1^{A(p,q)} d\tau \frac{(\tau^2 - 1)^{m-1}}{\tau^{n-1}}, \tag{69}$$

$$A(p, q) = \frac{[PQ]}{\|p - q\|}, \quad [PQ] = \sqrt{p^2 q^2 - 2p \cdot q + 1},$$

and have the boundary conditions that $G_{(m,n)}(p, q)$ and its derivatives of order 1 to $m - 1$ are zero on $\partial\mathcal{M}$.

In the context of splines in general, these Green’s functions generate the polyharmonic clamped-plate splines [19], the compactly-supported analogue of the polyharmonic thin-plate splines introduced by Duchon [8].

6. Constructing the log map

Thus far we have shown how to construct the Riemannian exponential map in $\text{Diff}_0^s(\mathcal{M})$ via the spline representation. However, in practice we actually want to find the velocity field (in some spline representation) that yields a given warp upon exponentiation; the parameters of the velocity field in the spline representation then provide the parameterization of the original warp. This mapping, the inverse of the exponential mapping, is called the log map.

Consider a general warp $g \in \text{Diff}_0^s(\mathcal{M})$, and suppose that it corresponds to the exponential mapping of some velocity field thus:

$$g = \exp_R(v). \tag{70}$$

Let us also take some spline basis, defined by a point $p \in \Omega^{(N)}$, which represents a set of non-coincident knotpoints $y = \{y_i \in \mathcal{M}, i = 1, \dots, N\}$. In the Hilbert space of velocity fields $\mathfrak{X}_0^s(\mathcal{M})$, any velocity field can be decomposed into a part lying in the spline subspace, and a part perpendicular to it, where ‘perpendicular’ is defined w.r.t. the scalar product $(\cdot, \cdot)_L$. That is:

$$v = v^{\parallel p} + v^{\perp p}, \quad (v^{\parallel p}, v^{\perp p})_L \equiv 0, \tag{71}$$

where $v^{\parallel p}(x) = \sum_{i,j} G(x, y_i) G_{ij}^{-1}(y) v_\mu(y_j)$. Then $v^{\parallel p}$ is just the spline interpolant (60) of the set of values of the original velocity field v taken at the knotpoint positions y . So, if we knew v (the log map applied to g), then we could iteratively construct some spline basis such that the Hilbert space norm of the perpendicular part, $\|v^{\perp p}\|_L^2 = (v^{\perp p}, v^{\perp p})_L$ tended towards zero.

However, this depends on knowing the log mapping of g , which is what we were originally trying to find. Instead, we consider decomposing the warp g rather than the velocity field:

$$g = g^\sharp \circ g^p, \tag{72}$$

where we define:

$$g^p = \arg \min\{d_G(e, f) \mid f(y_i) = g(y_i) \forall i = 1, \dots, N\}. \tag{73}$$

By definition, g^p is the warp given by the geodesic in $\Omega^{(N)}$ between $y = \{y_i\}$ and $\{g(y_i)\}$, and g^\sharp is the correcting warp, where:

$$(g^\sharp \circ g^p)(y_i) = g^\sharp(g(y_i)) = g(y_i) \quad \forall i = 1, \dots, N, \tag{74}$$

i.e., g^\sharp is the identity as far as the motion of the knotpoints is concerned. So, given a method of iteratively constructing the spline basis so that the correcting warp tends to the identity for the motion of all points $x \in \mathcal{M}$, the velocity field given by the log map of the spline warp $v^p = \log_R(g^p)$ will approach the log map of the original warp. We can construct the log map in the spline basis by finding the geodesic in $\Omega^{(N)}$ with endpoints $\{y_i\}$ and $\{g^p(y_i)\}$, and the log map of g^p is just the initial velocity of this geodesic.

We should note here that this log map in $\Omega^{(N)}$ is not necessarily single-valued. For example, in two dimensions, for the case where we have just two diametrically opposed knotpoints that change positions under the warp, we can see that there are two equivalent solutions, corresponding to clockwise and anticlockwise motion. Of course, in general, knotpoint initial and final positions will not have the symmetries present in the above case.

6.1. Implementation: Computing the log map

The authors have previously described [19] an iterative, greedy algorithm to construct the log map of an arbitrary warp as defined by the warp of an image. An unwarped image is described by a set of pixel/voxel values $I(x_0)$, where x_0 is the set of pixel positions, which usually lie on some regular grid. An image warp is then defined by the set of warped pixel positions $x_g = g(x_0)$, which generates the warped image I_g , where the set of pixel values $I_g(x_0)$ is obtained from $I_g(x_g) = I(x_0)$ by resampling.

For a given spline representation p and spline approximant g^p , we calculate the set of pixel discrepancies:

$$\|g^\sharp(g(x_0)) - g(x_0)\|_{\mathcal{M}}, \tag{75}$$

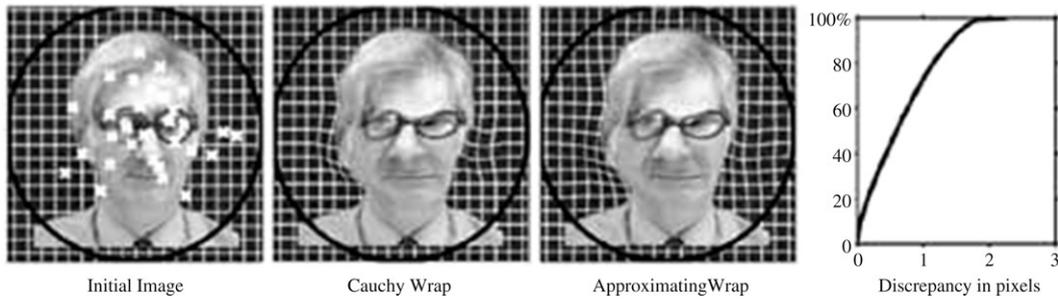


Fig. 1. From the left: the unwarped image, with initial knotpoint positions (white crosses) and the unit circle (black line), the Cauchy warped image, the approximating warp generated by the set of knotpoints, and the cumulative distribution of discrepancies of pixel positions within the unit circle, in units of the unwarped pixel size.

where $\|\cdot\|_{\mathcal{M}}$ is the Euclidean norm on \mathcal{M} . The pixel position for which this discrepancy is greatest is then added as a new knotpoint, enlarging the spline representation $p \rightarrow p'$, and the process is repeated until the discrepancies are sufficiently small.

Ideally, we would use the Hilbert norm $\|\cdot\|_L$ of the discrepancy field or the geodesic distance $d_{\mathcal{G}}(e, g^{p'})$ to define convergence. However, these are only calculable for fields/warps in a spline representation, and $g^{p'}$ is by definition not in the spline representation p , hence we use the Euclidean norm on \mathcal{M} .

An example is given in Fig. 1. The warp that we wish to approximate is generated by concatenating 20 random, localized Cauchy warps; each warp acts only within a defined ellipsoidal region, which is itself restricted to lie wholly within the unit circle. For the purposes of illustration, we here define the unit circle as the inscribed circle of the image. Each Cauchy warp is a parameterized, strictly diffeomorphic warp of any ellipsoidal region. The description of the entire Cauchy-based warp requires $20 \times 6 = 120$ degrees of freedom.

This warp is then applied to a 190×190 pixel image. The warped image is not resampled, but plotted as a coloured surface with deformed faces/pixels. We then approximate this warp using the greedy algorithm described above. The result for the case of 30 knotpoints (equal to $30 \times 2 \times 2 = 120$ degrees of freedom) is given. Note that the Cauchy-based representation is inherently local and mathematically unrelated to the GIS representation. However, for the same number of degrees of freedom, we see that the approximating warp is visually extremely close to the Cauchy warp. The cumulative distribution of discrepancies (75) for pixels inside the unit circle is given in the figure, with 50% of the pixels having a discrepancy of less than 0.6 pixel units, and the maximum discrepancy being 2.25 pixels.

7. Application: Classifying variation

In this section we provide a simple demonstration of the fact that the inclusion of the geodesic distance in the analysis of sets of warps produces better results than using just the naïve Euclidean metric based on the statistics of the set (that is, the Mahalanobis distance), which is the standard approach that is taken in medical image analysis. Our demonstration dataset consists of 20 annotated outlines of the anterior lateral ventricles taken from a larger set of annotated T1 MRI scans.⁷ Each pair of ventricles consisted of 40 landmarks selected by a radiologist. According to Bookstein [5] these points are ‘pseudo-landmarks’, since they lie on the image edges and need not correspond to genuine anatomical landmarks. The extent to which this effects the results reported here is unknown. The set of training examples was Procrustes aligned and then scaled to fit inside the unit circle.

We built a (linear) symmetrical Statistical Shape Model (SSM) [7] from this training set of points (and their mirror images). This model was then used to generate a set of 48 random examples; these example were classified by hand to be illegal (13 examples) if the outlines of the two ventricles intersected either themselves or each other, and legal otherwise — see Fig. 2 for examples. The training set were all defined to be legal.

Using the 40 landmarks on each image as knotpoints, we computed the GIS warp between each annotated outline and the mean shape from the SSM, using the biharmonic clamped-plate spline basis (69); for full implementation

⁷ An earlier, similar analysis using examples from the same larger dataset was presented in [19]. It is included here as a demonstration of the method.

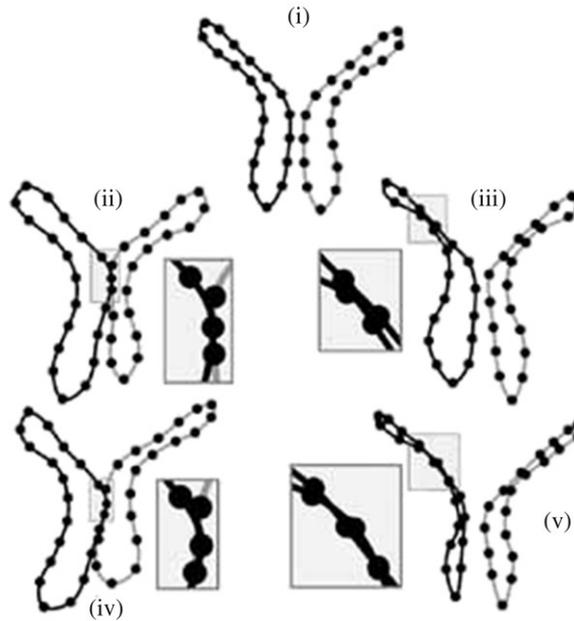


Fig. 2. (i) The mean shape from the training set, (ii) & (iii) Legal examples and, (iv) & (v) illegal examples from the test set. Note that the lines are for the purposes of illustration only, and the detail shows the points of near and actual intersection.

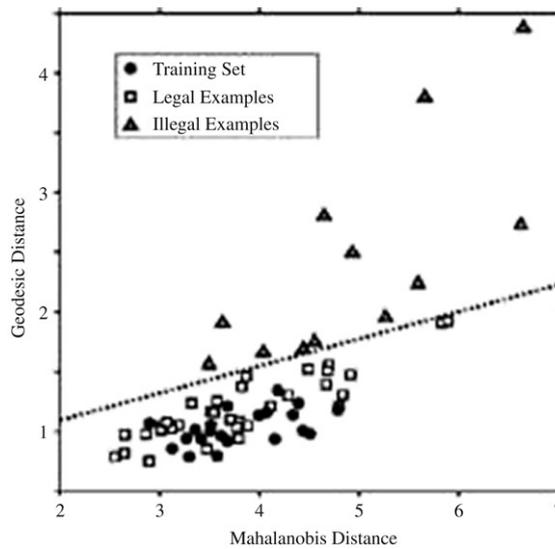


Fig. 3. Mahalanobis versus geodesic distances from the mean shape for Black circles: the training set, Grey triangles: illegal shapes generated by the SSM, White squares: legal shapes generated by the SSM.

details, see [19]. Note that the outlines in the Figure are for the purposes of illustration only, and that in all cases, the warps are actually diffeomorphisms, even when the added outlines then intersect.

Fig. 3 shows a plot of the geodesic (metric) distance against the Mahalanobis (statistical, Euclidean) distance for all of the examples. It can be seen that it is impossible to separate the legal from the illegal examples using the Mahalanobis distance alone, but that including the metric distance allows a linear classifier that perfectly separates the classes. This is possible because the metric distance has a sensitivity to changes in the ‘topology’ of the knotpoints that the naïve Euclidean metric lacks, not because the Euclidean metric can generate non-diffeomorphic warps.

8. Discussion and conclusions

In this paper we have considered the problem of constructing an atlas on the diffeomorphism group of a compact manifold with boundary. Our approach is based on a right-invariant metric on $\text{Diff}_0^s(\mathcal{M})$ that comes from a scalar differential operator. This gives a spline interpolant on the space of velocity fields, and a geodesic interpolating spline that is guaranteed to be diffeomorphic on the space of warps. The particular spline interpolant we have introduced is the polyharmonic clamped-plate spline, which is zero on and outside the unit ball, and vanishes smoothly at the boundary.

The spline generates a chart on the submanifold defined by the knotpoints of the spline, with the coordinates in the chart being the initial velocities of those knotpoints. The choice of a differential operator as the basis for the metric on the diffeomorphism group means that the spline smoothes the data. This means that the warp parameterization is usually of a reasonably low dimensionality, which has obvious benefits for computational applications. Another computational aspect is that it is significantly easier and more robust, given our intended application, to use the final knotpoint positions as the coordinates of the chart, rather than the initial velocities or momenta [30]. This is exactly equivalent, and leads to the coordinate space $\Omega^{(N)}$, as introduced in Section 5.1. However, as noted in Section 6, the log map on $\Omega^{(N)}$ is not necessarily single-valued. In practice however, we do not expect this to be a problem.

A considerable body of previous work on using the diffeomorphism group in image analysis has come from the work of Miller, Trounev, Younes, Joshi, and co-workers. This was mainly based on the idea of *patterns* or *templates* (for example, see [22,29,11,30]). The simplest pattern is defined as the configuration of some set of N landmark points; the landmarks can either be derived from an image, or correspond to actual physical markers (e.g. attached to the skull as in [5]). Then $\Omega^{(N)}$ is the space of configurations of the patterns.

Our approach differs from that previously mentioned in that we do not treat objects in images as patterns or templates. We explicitly deal with the entire image, not a shape/pattern derived from the image. We could treat the entire *pixelated* image as a large pattern, but this would be very inefficient – the landmark points would have to be the positions of each pixel/voxel in the image. Instead, we consider the image as a discrete sampling of some underlying continuous function, so that the appropriate configuration space is now the diffeomorphism group of the image plane/volume, and the appropriate metric is the metric on the diffeomorphism group, not the derived metric d_Ω .

The primary motivation for this work is the modelling of sets of image warps for the classification of normal and abnormal variation of biological structures, and hence the tracking of disease progression. We have shown experimentally that using the interpolating spline for chart construction generates low-dimensional representations of the image warps. The iterative greedy algorithm used to construct the spline-based representation of the warps can be generalized to produce a common, low-dimensional representation of a set of image warps. Computing statistics on this representation and building a generative model on the spline submanifold defined by such a set of warps are the obvious next steps, and are the subject of our future research.

Acknowledgements

Joint first authors. This research was supported by the MIAS IRC project, EPSRC grant number GR/N14248/01, UK Medical Research Council Grant No. D2025/31. S. Marsland is also supported by the Marsden Fund grant MAU0408, Royal Society of New Zealand.

References

- [1] M. Adams, T. Ratiu, R. Schmid, The Lie group structure of diffeomorphism groups and invertible Fourier integral operators with applications, in: V. Kac (Ed.), *Infinite Dimensional Groups with Applications*, in: *Mathematical Sciences Research Institute Publications*, vol. 4, Springer-Verlag, 1985.
- [2] V.I. Arnold, Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, *Annales de L'Institut Fourier (Grenoble)* 16 (1) (1966) 319–361.
- [3] V. Arsigny, O. Commowick, X. Pennec, N. Ayache, A log-euclidean framework for statistics on diffeomorphisms, in: R. Larsen, M. Nielsen, J. Sporring (Eds.), *Proceedings of 9th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) Part I*, in: *Lecture Notes in Computer Science*, vol. 4190, Springer-Verlag, 2006, pp. 924–931.
- [4] T. Boggio, Sulle funzioni di Green d'ordine m , *Rendiconti - Circolo Matematico di Palermo* 20 (1905) 97–135.
- [5] F.L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, Cambridge, UK, 1997.

- [6] V. Camion, L. Younes, Geodesic interpolating splines, in: M. Figueiredo, J. Zerubia, A.K. Jain (Eds.), Proceedings of EMMCVPR'01, in: Lecture Notes in Computer Science, vol. 2134, Springer-Verlag, 2001, pp. 513–527.
- [7] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models – their training and application, *Computer Vision and Image Understanding* 61 (1) (1995) 38–59.
- [8] J. Duchon, Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, *Revue Française d'Automatique, Informatique, Recherche Opérationnelle (RAIRO) Analyse Numerique* 10 (1976) 5–12.
- [9] P. Dupuis, U. Grenander, M.I. Miller, Variational problems on flows of diffeomorphisms for image matching, *Quarterly of Applied Mathematics* 56 (3) (1998) 587–600.
- [10] D.G. Ebin, J.E. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, *Annals of Mathematics*, 2nd Series 92 (1) (1970) 102–163.
- [11] D.D. Holm, J.T. Ratnanather, A. Trouvé, L. Younes, Soliton dynamics in computational anatomy, *NeuroImage* 23 (Suppl. 1) (2004) S170–S178.
- [12] D.D. Holm, J.E. Marsden, T.S. Ratiu, The Euler-Poincaré equations and semidirect products, with applications to continuum theories, *Advances in Mathematics* 137 (1) (1998) 1–81.
- [13] D.D. Holm, J.E. Marsden, Momentum maps and measure-valued solutions (peakons, filaments and sheets) for the EPDiff Equation, in: *Progressive Mathematics*, vol. 232, Birkhäuser, Boston, 2005, pp. 203–235.
- [14] S.C. Joshi, M.M. Miller, Landmark matching via large deformation diffeomorphisms, *IEEE Transactions on Image Processing* 9 (8) (2000) 1357–1370.
- [15] N. Kopell, Commuting Diffeomorphisms, in: *Proceedings of the Symposium of Pure Mathematics*, vol. 14, American Mathematical Society, 1970, pp. 165–184.
- [16] S. Kouranbaeva, The Camassa-Holm equation as a geodesic flow on the diffeomorphism group, *Journal of Mathematical Physics* 40 (2) (1999) 857–868.
- [17] A. Kriegl, P.W. Michor, *The Convenient Setting of Global Analysis*, in: *Mathematical Surveys and Monographs*, vol. 53, American Mathematical Society, Providence, RI, 1997.
- [18] J.E. Marsden, T.S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, 2nd edition, Springer, Berlin, 1999.
- [19] S. Marsland, C.J. Twining, Constructing diffeomorphic representations for the groupwise analysis of the nonrigid registrations of medical images, *IEEE Transactions on Medical Imaging* 23 (8) (2004) 1006–1020.
- [20] P.W. Michor, Some geometric evolution equations arising as geodesic equations on groups of diffeomorphism, including the Hamiltonian approach, in: A. Bove, F. Colombini, D. Del Santo (Eds.), *Phase Space Analysis of Partial Differential Equations*, in: *Progress in Non Linear Differential Equations and Their Applications*, vol. 69, Birkhauser Verlag, 2006, pp. 133–215.
- [21] P.W. Michor, T.S. Ratiu, On the geometry of the Virasoro-Bott group, *Journal of Lie Theory* 8 (2) (1998) 293–309.
- [22] M. Miller, A. Trouvé, L. Younes, Geodesic shooting for computational anatomy, *Journal of Mathematical Imaging and Vision* 24 (2) (2006) 209–228.
- [23] M.I. Miller, A. Trouvé, L. Younes, On the metrics and Euler-Lagrange equations of computational anatomy, *Annual Review of Biomedical Engineering* 4 (2002) 375–405.
- [24] W.A. Newcomb, Lagrangian and Hamiltonian methods in magnetohydrodynamics, *Nuclear Fusion: Suppl. Part 2* (1962) 451–463.
- [25] H. Omori, On the group of diffeomorphisms on a compact manifold, in: *Proceedings of the Symposium of Pure Mathematics*, vol 15, American Mathematical Society, 1970, pp. 167–184.
- [26] R. Schmid, Infinite dimensional Lie groups with applications to mathematical physics, *Journal of Geometry and Symmetry in Physics* 1 (2004) 54–120.
- [27] S. Shkoller, Geometry and curvature of diffeomorphism groups with H^1 metric and mean hydrodynamics, *Journal of Functional Analysis* 160 (1998) 337–365.
- [28] A.W. Toga, *Brain Warping*, Academic Press, San Diego, 1999.
- [29] A. Trouvé, Diffeomorphism groups and pattern matching in image analysis, *International Journal of Computer Vision* 28 (3) (1998) 213–221.
- [30] M. Vaillant, M. Miller, L. Younes, A. Trouvé, Statistics on diffeomorphisms via tangent space representations, *NeuroImage* 23 (Suppl. 1) (2004) S161–S169.
- [31] B. Zitová, J. Flusser, Image registration methods: A survey, *Image and Vision Computing* 21 (2003) 977–1000.