

# SHAPE SPACES: FROM GEOMETRY TO BIOLOGICAL PLAUSIBILITY

NICOLAS CHARON AND LAURENT YOUNES

**ABSTRACT.** This chapter reviews several Riemannian metrics and evolution equations in the context of diffeomorphic shape analysis. After a short review of various approaches at building Riemannian spaces of shapes, with a special focus on the foundations of the large deformation diffeomorphic metric mapping algorithm, the attention is turned to elastic metrics, and to growth models that can be derived from it. In the latter context, a new class of metrics, involving the optimization of a growth tensor, is introduced and some of its properties are studied.

## 1. INTRODUCTION: SHAPE SPACES

Shape has long been an object of scientific study, especially in life sciences where it provided a primary element in the differentiation between species. It was—in complement to behavioral patterns—a central factor of the early justification of evolutionary theory, and was of course the main subject of D’Arcy Thompson seminal work “On Growth and Forms” [63].

The construction of mathematical models of shape spaces, however, was more recent, and started with David Kendall’s landmark paper introducing a shape space as a particular Riemannian manifold [45], a construction motivated by the need to provide a formal mathematical framework for statistical analyses of shape datasets. In Kendall’s model, shapes are represented as ordered collections of distinct points with fixed cardinality. The manifold structure is obtained as a quotient space through the action of rotations, translations and scaling and the metric as the projection of the Euclidean metric to this quotient space. Kendall’s shape space has since been used in a large variety of applications, with increasing numbers of available shape datasets and relevant associated statistical questions (see the recent edition of Dryden and Mardia [22] for additional details and references).

Kendall’s shape space is however limited by the need to provide a consistent ordering (or labeling) of the points constituting the shape, and by the requirement that they form a finite set. Shape datasets are typically formed by unlabelled geometric objects, and using Kendall’s shape space requires defining and indexing (often manually) collections of landmarks for each shape, resulting in an intensive and sometimes imperfectly specified problem. Defining shape spaces whose elements are curves or surfaces requires however more advanced mathematical tools, notably from global analysis [60], and a recent description of various formulations of shape spaces in this general context can be found in Bauer et al. [9]. In spite of the additional mathematical technicality, the construction of these shape spaces follows the general principles leading to Kendall’s space: first define a simple space of geometric objects as an open subset of a normed (or Fréchet) space, where the finite-dimensional space of ordered distinct points is replaced, e.g., by a space of immersions (or embeddings) from a fixed manifold  $M$  (the parameter space) to  $\mathbb{R}^d$ , the ambient space. This space (and its norm) is then quotiented by group actions to which shapes must be invariant, bringing in,

---

Nicolas Charon is partially supported by NSF 1945224 and NSF 1953267.

Laurent Younes is partially supported by NIH U19AG033655, R01NS102670 and R01AG055121.

in addition to previous actions of translations, rotations and scaling, the infinite-dimensional group of reparametrizations, provided by diffeomorphisms of the parameter space. Another modification to the finite-dimensional framework is that the Euclidean metric, as the base norm, which was a natural choice when working with finite sets of points, now needs to be replaced with some invariant Hilbert metric (if one wants a Riemannian structure at the end) on the space of immersions, for which there are many choices, including the whole family of invariant Sobolev norms. The well-posedness of various concepts in the resulting shape space, such as the non-degeneracy of the metric or the existence of geodesics, indeed depends on this choice. A striking example is the fact that the Riemannian distance between any pair of shapes may trivialize to zero for certain metrics, as initially discovered in Michor and Mumford [50] in the case of curves and then extended to other shapes spaces [11].

From the whole variety of shape spaces that can be built following this construction, a small number actually leads to practical algorithms and numerical implementations, which is an essential requirement when the goal is to analyze shape datasets. For curves, an important example is associated with a class of first-order Sobolev metrics on the space of immersions. One can indeed show that, after quotienting out rotations, translations and/or scalings, the resulting Riemannian manifold is isomorphic to standard manifolds (such as the infinite dimensional sphere, and Stiefel or Grassman manifolds) on which geodesic and geodesic distances can be explicitly computed. For curves, the additional cost of adding reparametrization invariance remains manageable, using, e.g., dynamic programming methods. A first example of such metrics was provided in Younes [69, 70] with further developments in Younes et al. [74]. A second example was then provided in Klassen et al. [46] (see Srivastava and Klassen [61]), and the approach was later extended to a one-parameter family including these two examples in Needham and Kurtek [58] and Younes [73], chapter 12 (see also Bauer et al. [8]).

Shape spaces have also been built using a different angle, leveraging the action of the diffeomorphism group of  $\mathbb{R}^d$  on a shape space. Diffeomorphisms of their ambient space indeed act transitively on most shapes of interest assuming that one fixes their topology (taking an example, diffeomorphisms of  $\mathbb{R}^2$  can be used to transform any  $C^1$  Jordan plane curve to any other). Using a metric on the diffeomorphism group with suitable properties, one can, given two shapes, compute the diffeomorphism closest to the identity that transforms the first shape into the other, and the distance between the identity and this optimal diffeomorphism also provides a distance between the considered shapes. (This construction will be described in detail in Section 2.) Formally, the considered shape space is provided by all diffeomorphic transformation of a given template. This approach can be seen as an application of Grenander's metric pattern theory [30, 29] and as a mathematical formulation of D'Arcy Thompson models [63]. It was introduced for shape spaces of images synchronously in Dupuis et al. [23] (with a precursor in Christensen et al. [20]) and Trouvé [64, 65], and for collections of labelled points in Miller et al. [53]. This formulation, very flexible, has later been applied to various shape spaces, such as unlabelled point sets [25], curves and surfaces [67, 26], vector or tensor fields [17, 18]. The reader may also refer to the recent survey in Bauer et al. [10] that describes in details the two previous approaches in the case of curves and surfaces.

Note that the previous discussion does not include the many methods that provide shape features, i.e., finite or infinite-dimensional descriptors that can be attached to a given shape, without necessarily providing them with a clear mathematical structure (such as that of a Riemannian manifold) which is one of the main concerns of the construction of shape spaces. Such methods were introduced in computer vision, medical imaging, biology and are too numerous to cite exhaustively in this chapter. Among the most important ones (a subjective statement), one can cite approaches using

complex analysis and the (quasi-) conformal maps to represent surfaces [33, 34, 76, 77, 48], isometry-invariant descriptors based on distance maps or Laplace-Beltrami eigenvectors [15, 16, 57, 59, 49], or the shape context [12]. The rest of this chapter will however remain focused on shape space approaches.

The construction of shape spaces as described above is based on purely geometric aspects. No physical law, or biological mechanism is used to define the various components that constitute the shape space. This non-committal approach is indeed justified, as shapes spaces are designed as containers for families of shapes that are not related to each other by a natural process (e.g., there is no physical process by which a finch’s beak can transform into the shape of another one). This fact provides the technical advantage that the construction of shape metrics is not constrained by the laws of nature and can therefore be selected so that they guarantee the existence, say, of geodesics, provide nicely behaved gradient flows, etc... This will be illustrated in Section 2.

On the other hand, biological processes provide many examples in which shapes change with time, in a process that is constrained by well specified laws. The goal of this chapter is to describe a few among recent attempts at representing such processes as trajectories in the shape spaces above, which, after small modifications or regularization, will be associated with evolution dynamics that behave well enough to allow for long time analysis and optimal control formulations.

This chapter is organized as follows. Section 2 provides a summary of the construction of shape spaces through diffeomorphic action. Section 3 focuses on variations of this construction with metrics that are inspired from elastic materials. Section 4 introduces a few examples of growth models in the context of shape spaces. For an extensive introduction to mathematical models of growth, the reader should refer to Goriely [28], which provides a splendid reference on the topic, and in particular on “morphoelasticity.” The representation of shape growth described in Section 4 will, however, deviate to some extent from that described in this reference, and more generally from the large literature exploring morphoelasticity, as models will be designed in the form of control systems, with a control equation interpreted as a differential equation in shape space and growth or atrophy directly associated with the control.

## 2. SHAPE SPACES UNDER DIFFEOMORPHIC ACTION

This section provides a summary of the construction of shape spaces based on the principles of D’Arcy-Thompson’s theory of transformations [63] and Grenander’s metric pattern theory [29]. The fundamental principles of the construction were laid in Dupuis et al. [23], Trouvé [64], Grenander and Miller [31] and the reader may refer to Younes [73], Miller et al. [54], Bauer et al. [10] for more recent accounts of the theory.

Shapes are modeled as embeddings from a fixed Riemannian manifold  $M$  into  $\mathbb{R}^d$ , and therefore have a with fixed topology (in practice,  $d = 2$  or  $3$ ). Typically,  $M$  is a unit circle or sphere, or a template shape of which one is computing deformations. Denote by  $Emb^p(M)$  the set of such  $C^p$  embeddings, or simply  $Emb$  when  $p$  and  $M$  are fixed. Each element  $m \in Emb$  provides a shape equipped with a parametrization. Objects of interest are shapes modulo parametrization (also called “unparametrized shapes”) in which one identifies embeddings  $m$  and  $\tilde{m}$  when they are related with each other through a change of parametrization, i.e.,  $\tilde{m} = m \circ \rho$  where  $\rho$  is a diffeomorphism of  $M$ . In other terms, the shape space is defined as the quotient space of  $Emb$  through the right action of the diffeomorphism group of  $M$ , and will be denoted as  $\mathcal{S}$ . Elements of  $\mathcal{S}$  will be denoted as  $[m]$ , for the equivalence class of  $m \in Emb$ .

Comparisons between shapes rely on the group of transformations acting on  $Emb$  or  $\mathcal{S}$ , which are modeled as diffeomorphisms of  $\mathbb{R}^d$ . Denote by  $Diff^p(\mathbb{R}^d)$ , or simply  $Diff^p$ , the group of  $C^p$

diffeomorphisms of  $\mathbb{R}^d$  and by  $Diff_0^p(\mathbb{R}^d)$ , or simply  $Diff_0^p$ , the subgroup of diffeomorphisms that converge to the identity map, denoted  $id_{\mathbb{R}^d}$ , at infinity (convergence being understood in the  $C^p$  sense). If  $\varphi \in Diff^p$  and  $m \in Emb$ ,  $\varphi \cdot m$  is simply  $\varphi \circ m$ , and this action commutes with reparametrization, so that one can define  $\varphi \cdot [m] = [\varphi \cdot m]$  without ambiguity.

To compare two embeddings (or their associated shapes)  $m$  and  $m'$ , one considers the transformations  $\varphi \in Diff_0^p$  that relate them, i.e., such that  $m' = \varphi \cdot m$ . One considers that  $m$  and  $m'$  are similar if one can find some  $\varphi$  relating them that is close to  $id_{\mathbb{R}^d}$ . This closeness is itself evaluated using a metric on  $Diff_0^p$ , with a construction described below.

To provide a Riemannian metric, one needs an inner-product norm that evaluates the velocity of time-dependent diffeomorphisms, or “diffeomorphic motions,” taking the form  $(x \mapsto \partial_t \varphi(t, x))$  where  $\varphi$  is a function of time and space such that  $(x \mapsto \varphi(t, x))$  is at all times an element of  $Diff_0^p$ . For a given time  $t$ ,  $(x \mapsto \partial_t \varphi(t, x))$  is a  $C^p$  vector field on  $\mathbb{R}^d$  and one therefore needs to provide a norm over such vector fields. The norm of the velocity at time  $t$  should in principle depend on the diffeomorphism at the same time,  $(x \mapsto \varphi(t, x))$ , but, for reasons seen below, it will be desirable for this norm to satisfy the invariance property that, when writing

$$\varphi(t + \delta t, x) = \varphi(t, x) + \delta \varphi(t, x) = (id_{\mathbb{R}^d} + \delta \varphi(t, \cdot) \circ \varphi^{-1}(t, \cdot)) \circ \varphi(t, x),$$

the cost associated with  $\delta \varphi$  is a fixed function of the deformation increment  $\delta \varphi \circ \varphi^{-1}$ . Passing to the limit, this means that the Riemannian norm of  $(x \mapsto \partial_t \varphi(t, x))$  at  $(x \mapsto \varphi(t, x))$  is equal to the norm of  $(x \mapsto \partial_t \varphi(t, x) \circ \varphi^{-1}(t, x))$  at  $id_{\mathbb{R}^d}$ . The vector field  $v(t, x) = \partial_t \varphi(t, x) \circ \varphi^{-1}(t, x)$  is called the Eulerian velocity of the diffeomorphic motion  $x \mapsto \varphi(t, x)$ , and the diffeomorphic motion is recovered from the Eulerian velocity by solving the ordinary differential equation

$$(1) \quad \partial_t \varphi(t, x) = v(t, \varphi(t, x)).$$

To define our Riemannian metric on  $Diff_0^p$ , it therefore suffices to specify a Hilbert norm on vector fields. For this purpose, let  $V$  denote a Hilbert of vector fields on  $\mathbb{R}^d$  that will be assumed, in order to recover elements of  $Diff_0^p$  after solving Eq. (1), to be continuously included in the space  $C_0^p(\mathbb{R}^d, \mathbb{R}^d)$  of  $C^p$  vector fields that vanish at infinity. This means that  $V \subset C_0^p(\mathbb{R}^d, \mathbb{R}^d)$  and that, for some constant  $c$ , one has (letting  $\|\cdot\|_\infty$  denote the supremum norm)

$$\sum_{k=0}^p \|d^k v\|_\infty \leq c \|v\|_V.$$

To satisfy this assumption,  $V$  can be built as a Hilbert Sobolev space of high enough order. In addition, since the continuous inclusion implies that  $V$  is a reproducing kernel Hilbert space (RKHS) of vector fields, RKHS theory can be used to build a large variety of Hilbert spaces of interest that satisfy the inclusion property [7, 42, 73]. One can then define the action functional of a diffeomorphic motion  $((t, x) \in [0, 1] \times \mathbb{R}^d \mapsto \varphi(t, x) \in \mathbb{R}^d)$  as

$$\int_0^1 \|v(t, \cdot)\|_V^2 dt$$

with  $\partial_t \varphi(t, x) = v(t, \varphi(t, x))$ . A geodesic diffeomorphic motion is an extremal of this action functional and a minimizing geodesic motion minimizes the functional subject to fixed boundary conditions at  $t = 0$  and  $t = 1$ . In particular, the geodesic distance between two diffeomorphisms  $\varphi_0$  and

$\varphi_1$  is defined as

$$(2) \quad d_V(\varphi_0, \varphi_1) = \inf \left\{ \left( \int_0^1 \|v(t, \cdot)\|_V^2 dt \right)^{1/2} : \partial_t \varphi(t, x) = v(t, \varphi(t, x)), \varphi(0, \cdot) = \varphi_0, \varphi(1, \cdot) = \varphi_1 \right\}.$$

Note that the set over which the infimum is computed may be empty, in which case the distance is infinite. If this set is not empty, then one says that  $\varphi_1$  is attainable from  $\varphi_0$ . Diffeomorphisms that are attainable from the identity form a subgroup of  $Diff_0^p$ , denoted  $Diff_V$ , and this subgroup is complete for the geodesic distance [64, 73]. (Because not every diffeomorphism in  $Diff_0^p$  is attainable from the identity, one is actually building a sub-Riemannian metric on this space. See Arguillère et al. [3], Younes et al. [75].)

By construction, the distance is right-invariant, i.e.,

$$d_V(\varphi_0, \varphi_1) = d_V(\text{id}_{\mathbb{R}^d}, \varphi_1 \circ \varphi_0^{-1}),$$

and this implies that it can be used to define a distance on  $\mathcal{S}$  via

$$\begin{aligned} d_{\mathcal{S}}([m_0], [m_1]) &= \inf \{d_V(\text{id}_{\mathbb{R}^d}, \varphi) : [\varphi \cdot m_0] = [m_1]\} \\ &= \inf \{d_V(\text{id}_{\mathbb{R}^d}, \varphi) : \varphi \cdot m_0 \in [m_1]\}. \end{aligned}$$

The distance on  $\mathcal{S}$  can itself be defined directly as

$$(3) \quad d_V([m_0], [\varphi_1]) = \inf \left\{ \left( \int_0^1 \|v(t, \cdot)\|_V^2 dt \right)^{1/2} : \partial_t m(t, \cdot) = v(t, m(t, \cdot)), m(0, \cdot) = m_0, m(1, \cdot) \in [m_1] \right\}.$$

This provides an optimal control problem in  $\mathcal{S}$  where the control is the time-dependent vector field  $v$ , and the state equation the ODE  $\partial_t m(t, \cdot) = v(t, m(t, \cdot))$ . The optimal trajectory transforms the initial  $m_0$  into an embedding that is a reparametrization of  $m_1$  and provides a minimizing geodesic in  $\mathcal{S}$ . If the Sobolev inclusion discussed above holds for  $p \geq 1$  at least, the variational problems described in Eq. (2) and Eq. (3) are well defined. The condition that  $\int_0^1 \|v(t, \cdot)\|_V^2 dt < \infty$  implies that solutions to the state equations ( $\partial_t \varphi = v \circ \varphi$  or  $\partial_t m = v \circ m$ ) exist and are unique (given initial conditions) over the full unit time interval, ensuring that the optimal control problem is well specified. Moreover, as soon as  $\varphi_1$  (resp.  $[m_1]$ ) is attainable from  $\varphi_0$  (resp.  $[m_0]$ ), an optimal solution to the considered problem always exists. Finally, under very mild assumptions on initial conditions, solutions of the geodesic equations exist and are uniquely specified by their initial position and velocity, i.e.,  $m(0, \cdot)$  and  $\partial_t m(0, \cdot)$  for  $d_{\mathcal{S}}$ . The geodesic equation is the Euler-Lagrange equation associated with the variational problem, satisfied by stationary points of Eq. (2) or Eq. (3) (equivalently, they are the equations provided by Pontryagin's maximum principle). In the case considered here, they are special instances of the geodesic equations for right-invariant Riemannian metrics on Lie groups, as described in Arnold [4, 5] and are often referred to as Euler-Arnold equations [6] or Euler-Poincaré equations [24, 35].

In practice, one does not solve this problem exactly, but relaxes the endpoint condition  $m(1, \cdot) \in [m_1]$  by adding a penalty term, therefore minimizing

$$(4) \quad \int_0^1 \|v(t, \cdot)\|_V^2 dt + U([m(1, \cdot)], [m_1])$$

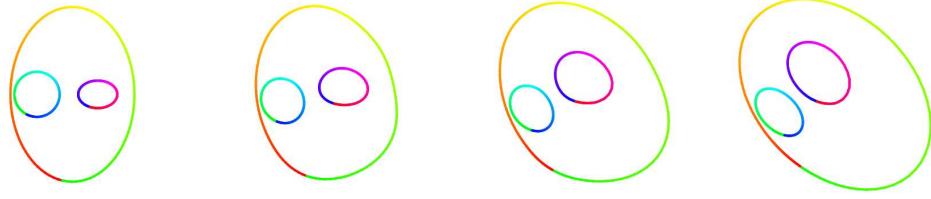


FIGURE 1. Four time points of a geodesic evolution in shape space. Note that shapes in this example have multiple components. Contour coloring match across time points and track the evolution of the curve initial parametrization.

subject to  $\partial_t m(t, \cdot) = v(t, m(t, \cdot))$ . In many of the applications, the function  $U$  takes the form

$$U([m_0], [m_1]) = \|\mathcal{J}_{[m_0]} - \mathcal{J}_{[m_1]}\|_H^2$$

where  $[m] \mapsto \mathcal{J}_{[m]}$  is a mapping from  $\mathcal{S}$  into a (much larger) Hilbert space  $H$ . These “chordal metrics” use representations of embedded curves or surfaces as measures, currents or varifold. For simplicity the presentation below will ignore this relaxation step (which is however necessary to make the computation numerically feasible) and work as if the endpoint conditions are exact. The reader is referred to Bauer et al. [10] or Charon et al. [19], and to the references within, for more information on chordal metrics.

A two-dimensional example of geodesic is presented in Fig. 1. These geodesics provide the non-linear equivalent of a linear interpolation in Euclidean space.

### 3. HYBRID MODELS

**3.1. Description.** The previous framework can be slightly extended to allow the norm used in the shape space to depend on the shape itself, replacing the control cost in Eq. (3) by

$$\int_0^1 \|v(t, \cdot)\|_{[m(t, \cdot)]}^2 dt,$$

so that the cost depends on both control and state. This still provides a sub-Riemannian distance in shape space, and the problem remains well specified as soon as one ensures that the shape-dependent norms still control the norm on  $V$ , so that an inequality ensuring

$$\|v\|_V \leq C\|v\|_{[m]}$$

holds for all  $m \in \mathcal{S}$  and  $v \in V$  (where the upper-bound may be infinite). Typical applications of this construction use a “weak norm”  $v \mapsto \llbracket v \rrbracket_{[m]}$  (which, by itself would not guarantee the existence of solutions to the state equation), possibly motivated by material or biological constraints, “regularized” by the norm on  $V$ , therefore taking

$$(5) \quad \|v\|_{[m]}^2 = \kappa\|v\|_V^2 + \llbracket v \rrbracket_{[m]}^2$$

for some  $\kappa > 0$ .

The following section discusses several possible choice for  $\llbracket v \rrbracket_{[m]}$  in Eq. (5), in which the shape is considered as an elastic material and the norm corresponds to the elastic energy associated with an infinitesimal displacement along  $v$  (the reader may refer to, e.g., Ciarlet [21], Gonzalez and Stuart [27] for more details on elasticity concepts that are used below). The concept of “hybrid” metrics in Eq. (5) was suggested in Younes [72]. A similar approach for spaces of images (combined with

a “metamorphosis” metric [52, 66]) was introduced in Berkels et al. [13], and metrics formed as discrete iterations of small elastic deformations were also studied in Wirth et al. [68].

### 3.2. Elastic metrics.

3.2.1. *Three-dimensional case.* The energy of a hyper-elastic material  $\Omega$  subject to a deformation  $\varphi$  takes the form (letting  $\text{Id}_{\mathbb{R}^d}$  denote the identity matrix in  $\mathbb{R}^d$ )

$$E = \int_{\varphi(\Omega)} G(x, \varphi(x)) dx$$

where

$$G(x, \varphi) = W(x, d\varphi^T d\varphi - \text{Id}_{\mathbb{R}^3})$$

for a function  $W : \Omega \times \text{Sym}^+ \rightarrow [0, +\infty)$  (where  $\text{Sym}^+$  is the set of 3 by 3 positive semi-definite matrices) such that  $W(x, S) = 0$  if and only if  $S = 0$ . The matrix  $C = d\varphi^T d\varphi$  is the Cauchy-Green strain tensor, which is such that  $u^T C u = |d\varphi u|^2$ , and  $W$  measures the deviation of this tensor from the identity matrix.

A second-order expansion of  $G$ , near  $\varphi = \text{id}_{\mathbb{R}^3}$  takes the form (using the fact that  $\partial_2 W(x, 0) = 0$ )

$$(6) \quad G(x, \varphi) \simeq \frac{1}{2} \partial_2^2 W(x, 0)(dv + dv^T, dv + dv^T)$$

where  $v = \varphi - \text{id}_{\mathbb{R}^3}$ . Here  $\partial_2 W(x, 0)$  and  $\partial_2^2 W(x, 0)$  are the first and second derivative with respect to the second variable of  $W$ , therefore a positive semi-definite symmetric bilinear form on  $\text{Sym}$  (the space of 3 by 3 symmetric matrices).

This can be used to define an elastic metric on 3D vector fields. Here,  $\Omega$  is considered as an “unparametrized shape,” taking the role of  $[m]$  in the previous sections. Using the previous notation, this corresponds to taking the manifold  $M$  to be an open subset of  $\mathbb{R}^3$  (e.g., an open ball),  $m$  an embedding of  $M$  into  $\mathbb{R}^3$  and identifying  $\Omega = m(M)$  to  $[m]$ . The hybrid norm will therefore be denoted

$$\|v\|_{\Omega}^2 = \kappa \|v\|_V^2 + \llbracket v \rrbracket_{\Omega}^2$$

and the rest of the discussion focuses on  $\llbracket v \rrbracket_{\Omega}$ . Based on Eq. (6), one is led to define a 3D elastic metric on vector fields as any norm taking the form

$$\llbracket v \rrbracket_{\Omega}^2 = \int_{\Omega} B(x, \varepsilon(x)) dx$$

where  $\varepsilon(x) = (dv(x) + dv(x)^T)/2$  is known as the infinitesimal strain tensor of the deformation and  $B(x, \cdot)$  is a positive semi-definite quadratic form on  $\text{Sym}$ , typically referred to as the elastic tensor. Generically,  $B(x, \cdot)$  can be represented as a  $6 \times 6$  symmetric positive semi-definite matrix, that is with 21 parameters in total at each  $x$ . In a majority of applications however, model symmetry assumptions significantly reduce the complexity of this elasticity tensor. In particular, in the case of a uniform and isotropic material,  $B(x, \cdot)$  is independent of the position and takes the specific form

$$(7) \quad B(x, \varepsilon) = B(\varepsilon) = \frac{\lambda}{2} \text{trace}(\varepsilon)^2 + \mu \text{trace}(\varepsilon^2)$$

which is the linearization of the energy of a Saint Venant-Kirchhoff material. In that case, the elasticity tensor is only described by the two parameters  $\lambda$  and  $\mu$  which are called the Lamé coefficients of the material.

To provide another example, consider the case of a partially isotropic and laminar model, introduced in Hsieh et al. [37, 38, 39] under the assumption that  $\Omega$  can be parametrized by a foliation.

More precisely, assume that there exist two surfaces  $\mathcal{M}_{\text{bottom}}$  and  $\mathcal{M}_{\text{top}}$  (bottom and top layers) included in  $\partial\Omega$  and a diffeomorphism  $\Phi : [0, 1] \times \mathcal{M}_{\text{bottom}} \rightarrow \Omega$  such that  $\Phi(\{0\} \times \mathcal{M}_{\text{bottom}}) = \mathcal{M}_{\text{bottom}}$  and  $\Phi(\{1\} \times \mathcal{M}_{\text{bottom}}) = \mathcal{M}_{\text{top}}$ . Let  $\mathcal{M}_s = \Phi(\{s\} \times \mathcal{M}_{\text{bottom}})$ ,  $s \in [0, 1]$ , denote “the layer at level  $s$ ,”  $S$  the transverse vector field  $S = \partial_s \Phi$  and  $N$  a unit vector field normal to all  $\mathcal{M}_s$ . One then introduces the following elasticity tensor:

$$(8) \quad B(x, \varepsilon) = \lambda_{\text{tan}} (\text{trace}(\varepsilon) - N^T \varepsilon N)^2 + \mu_{\text{tan}} (\text{trace}(\varepsilon^2) - 2 N^T \varepsilon^2 N + (N^T \varepsilon N)^2) + \mu_{\text{tsv}} (S^T \varepsilon S)^2 + 2 \mu_{\text{ang}} (S^T \varepsilon^2 S - (N^T \varepsilon S)^2),$$

The first two terms in this expression define an isotropic model on each layer. The third term measures a transversal string, evaluated along  $S$ . The last term measures an angular strain, that vanishes when  $S$  is normal to the layers. Here, the coefficients  $\lambda_{\text{tan}}, \mu_{\text{tan}}, \mu_{\text{tsv}}, \mu_{\text{ang}}$  must be constant on each layer  $m_s$  (they may depend on  $s$ ). Note that, if  $\tau_1, \tau_2$  are two orthonormal vectors fields that are tangent to the layers so that  $(\tau_1, \tau_2, N)$  forms at all points an orthonormal frame, then

$$\text{trace}(\varepsilon) - N^T \varepsilon N = \tau_1^T \varepsilon \tau_1 + \tau_2^T \varepsilon \tau_2$$

and

$$\text{trace}(\varepsilon^2) - 2 N^T \varepsilon^2 N + (N^T \varepsilon N)^2 = (\tau_1^T \varepsilon \tau_1)^2 + (\tau_2^T \varepsilon \tau_2)^2 + 2(\tau_1^T \varepsilon \tau_2)^2$$

so that the first two terms only involve deformations tangent to the layers.

Importantly, the space of such “layered structures” is stable by diffeomorphic action. Indeed, given  $\Omega$  and  $\Phi$  as above, and  $\varphi$  a diffeomorphism of  $\mathbb{R}^3$ , one defines the transformed structure by:

$$\varphi \cdot (\Omega, \Phi) = (\varphi(\Omega), \varphi \circ \Phi \circ \varphi^{-1}).$$

In particular,  $S$  transforms through  $\varphi$  as  $\varphi \cdot S = (d\varphi S) \circ \varphi^{-1}$ .

Returning to the general case, one must emphasize the fact that the action functional

$$\int_0^1 \int_{\Omega(t)} B(x, \varepsilon(x)) dx dt$$

with  $\partial_t \varphi(t, x) = v(t, \varphi(t, x))$  and  $\Omega(t) = \varphi(t, \cdot)(\Omega_0)$  is not the energy of a deforming elastic material, in the sense given to it in elasticity theory. In contrast, it may be understood as a sum of infinitesimal elastic energies, for a volume that slowly deforms, and at each time step, remodels its structure to reach an equilibrium state *without—up to reorientation—changing its elasticity properties*.

**3.2.2. Elastic metrics on surfaces.** The definition of elastic metrics on surfaces can be inferred using a pattern similar to the 3D derivation. Let  $\mathcal{M}$  be a surface in  $\mathbb{R}^3$  and consider a one-to-one immersion  $\varphi : \mathcal{M} \rightarrow \mathbb{R}^3$  (one can, in this discussion, think of  $\varphi$  as the restriction to  $\mathcal{M}$  of a diffeomorphism of  $\mathbb{R}^3$ ). To define a hyperelastic energy, assume (restricting  $\mathcal{M}$  if needed and introducing partitions of unity) that two vector fields  $\tau_1, \tau_2$  are chosen on  $\mathcal{M}$  such that they form at each point an orthonormal frame, and let  $\nu = \tau_1 \times \tau_2$ . Let  $F(x)$  denote the  $3 \times 2$  matrix  $[d\varphi \tau_1, d\varphi \tau_2](x)$ , where the 3D columns are expressed in the canonical basis of  $\mathbb{R}^3$ , and consider energies of the form

$$\int_M W(x, F(x)) d\text{vol}_M(x).$$

Material independence requires that  $W$  is invariant when  $F$  is multiplied on the right by a 2D rotation matrix, and this implies that  $W$  only depends on  $FF^T$ . To obtain the expression of

the metric, we let  $\varphi(x) = x + v(x)$  and make a first order expansion in  $v$  of  $FF^T$  with  $F = [\tau_1 + dv\tau_1, \tau_2 + dv\tau_2]$  yielding

$$FF^T \simeq \pi_{\mathcal{M}} + \pi_{\mathcal{M}}dv^T + dv\pi_{\mathcal{M}}$$

where  $\pi_{\mathcal{M}} = \tau_1\tau_1^T + \tau_2\tau_2^T$  is the orthogonal projection on the tangent plane to  $\mathcal{M}$  at  $x$ . The Riemannian elastic metric should therefore be taken as a quadratic form of  $\eta_{\mathcal{M}} := (\pi_{\mathcal{M}}dv^T + dv\pi_{\mathcal{M}})/2$ . Expressing this operator in the basis  $(\tau_1, \tau_2, \nu)$ , one sees that it depends on the five quantities  $a_{11} = \tau_1^Tdv\tau_1$ ,  $a_{22} = \tau_2^Tdv\tau_2$ ,  $a_{12} + a_{21} = \tau_1^Tdv\tau_2 + \tau_2^Tdv\tau_1$ ,  $a_{13} = \nu^Tdv\tau_1$  and  $a_{23} = \nu^Tdv\tau_2$ , yielding 15 free parameters for the “elastic norm”  $\|v\|_{\mathcal{M}}$ . (Like in the previous section, an identification is made between the unparametrized surface  $\mathcal{M} = m(M)$  and the equivalence class  $[m]$ .)

The norm is isotropic if it satisfies  $\|Rv\|_{\mathcal{M}} = \|v\|_{\mathcal{M}}$  for any 3D rotation that leaves  $\nu$  invariant. This implies that the matrix  $\mathbf{a} = \begin{pmatrix} a_{11} & (a_{12} + a_{21})/2 \\ (a_{12} + a_{21})/2 & a_{22} \end{pmatrix}$  is transformed by a 2D rotation as  $\mathbf{a} \mapsto R^T\mathbf{a}R$  and the vector  $\mathbf{b} = \begin{pmatrix} \nu^Tdv\tau_1 \\ \nu^Tdv\tau_2 \end{pmatrix}$  as  $\mathbf{b} \mapsto \mathbf{b}R$ . Using usual invariance arguments, this requires that the squared norm must be a (quadratic) function of  $\text{trace}(\mathbf{a})$ ,  $\text{trace}(\mathbf{a}^2)$  and  $|\mathbf{b}|^2$ , yielding

$$(9) \quad \|v\|_{\mathcal{M}}^2 = \int_{\mathcal{M}} \beta(x, \eta_{\mathcal{M}}) d\text{vol}(x)$$

with

$$(10) \quad \beta(x, \eta_{\mathcal{M}}) = \lambda_{\tan} \text{trace}(\mathbf{a})^2 + \mu_{\tan} \text{trace}(\mathbf{a}^2) + \mu_{\text{tsv}} |\mathbf{b}|^2.$$

The three different terms of this metric can be also interpreted as penalties on the changes of local area, metric tensor and normal vector respectively, as pointed out in Jermyn et al. [41] (see also their intrinsic expressions derived in Appendix A). Jermyn et al. [41] focuses on the special case  $\lambda_{\tan} = 1/16$ ,  $\mu_{\tan} = 0$ ,  $\mu_{\text{tsv}} = 1$ , which can be shown to be isometric to a Euclidean metric under a “square root normal transform.”

To consider another example, let  $\lambda_{\tan} = 0$  and  $\mu_{\tan} = \mu_{\text{tsv}} = 1$ . Then

$$(11) \quad \begin{aligned} \beta(x, \eta_{\mathcal{M}}) &= a_{11}^2 + a_{22}^2 + \frac{1}{2}(a_{12} + a_{21})^2 + a_{13}^2 + a_{23}^2 \\ &= a_{11}^2 + a_{22}^2 + a_{12}^2 + a_{21}^2 + a_{13}^2 + a_{23}^2 - \frac{1}{2}(a_{12} - a_{21})^2 \\ &= \text{trace}(dv dv^T) - \frac{1}{2}(a_{12} - a_{21})^2 \end{aligned}$$

The first term,  $\text{trace}(dv dv^T)$ , corresponds to the  $H^1$  metric on  $\mathcal{M}$ , used e.g., in Younes [72]. This metric, without the correction term  $\frac{1}{2}(a_{12} - a_{21})^2$  is not an elastic metric. It belongs however to a larger class of metrics, studied in Su et al. [62], where the correction term is added to Eq. (10) with a fourth parameter.

Such elastic metrics can be used in combination with the LDDMM metric through the hybrid setup described above. As an illustration, Fig. 2 provides a comparison of the geodesic trajectories between two surfaces, obtained with the pure LDDMM model and a hybrid model using the elastic term given by Eq. (11).

The norm in Eq. (9) can also be obtained as a limit of the laminar elastic model of the previous section and the energy in Eq. (8), which is shown in Appendix A by also providing an intrinsic

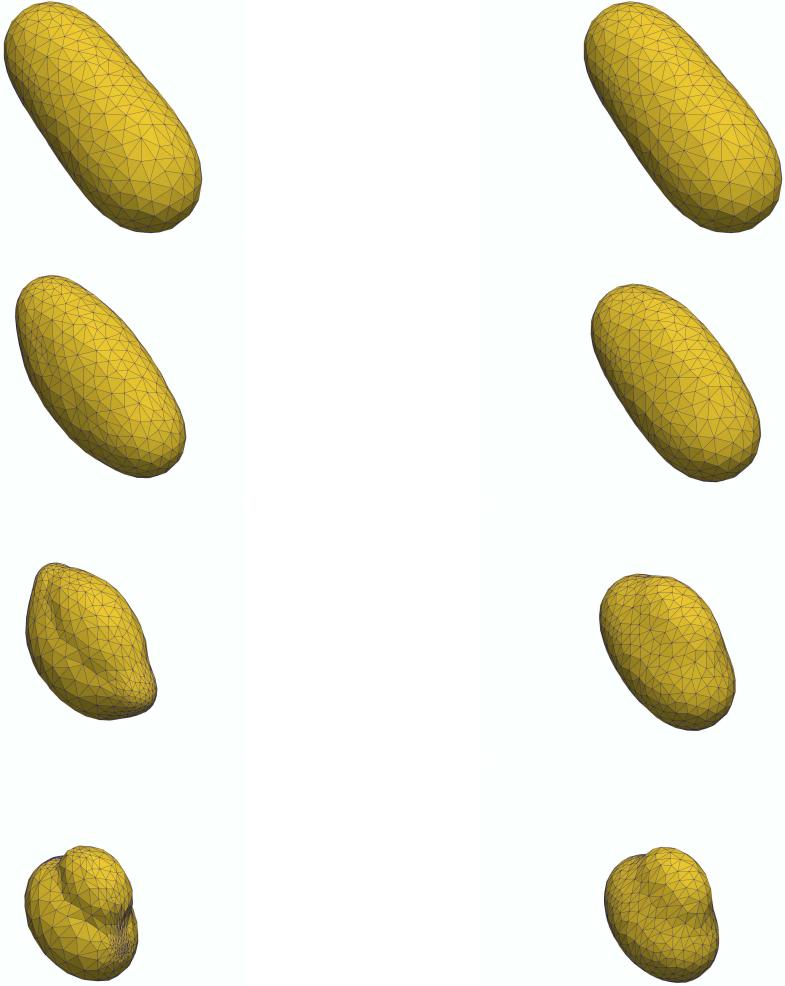


FIGURE 2. Comparison between geodesics between surfaces using a pure LDDMM and a hybrid LDDMM/elastic metric. First column: Four time points of an LDDMM geodesic ( $t = 0$ ,  $t = 0.3$ ,  $t = 0.7$  and  $t = 1$ ). Second column: same time points for the hybrid geodesic. One can note a difference in the intermediate shapes, and (as indicated by the triangulation) higher local contraction associated with the LDDMM metric. The hybrid metric uses the expression provided in Eq. (11).

expression of the elastic norm. This in part justifies the terminology of elastic metrics given to this framework in the related literature.

**3.2.3. Elastic metrics on curves.** If  $\mathcal{M}$  is a 3D curve, the same analysis shows that elastic metrics should depend on the products  $\tau^T d\tau$ ,  $\nu_1^T d\tau$  and  $\nu_2^T d\tau$ , where  $\tau$  is a unit tangent on  $M$  and  $(\tau, \nu_1, \nu_2)$  is a continuous positively oriented frame defined along the curve. Denote  $\partial_s v = d\tau$  for

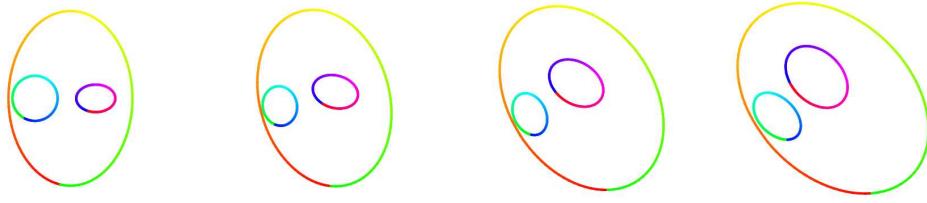


FIGURE 3. Four time points of a geodesic evolution in shape space for a hybrid metric. The initial and final shapes are the same as those in Fig. 1, but one can note, in particular, that the elliptical shapes are better conserved during the motion.

the derivative with respect to arc length, as introduced, e.g., in Michor and Mumford [51]. The metric must also be invariant to rotations of the normal frame  $(\nu_1, \nu_2)$  and changes of orientation on  $\mathcal{M}$ , which requires the metric to take the form

$$(12) \quad \|v\|_{\mathcal{M}}^2 = \int_{\mathcal{M}} \beta(x, \partial_s v) d\text{vol}(x)$$

with

$$\beta(x, \partial_s v) = \mu_{\tan}(\tau^T \partial_s v)^2 + \mu_{\text{tsv}}((\nu_1^T \partial_s v)^2 + (\nu_2^T \partial_s v)^2).$$

The special case of planar curves has been extensively discussed. In this case, letting  $\nu$  denote the unit normal, the metric has two parameters, with

$$\beta(x, \partial_s v) = \mu_{\tan}(\tau^T \partial_s v)^2 + \mu_{\text{tsv}}(\nu^T \partial_s v)^2.$$

When  $\mu_{\tan} = \mu_{\text{tsv}} = 1$ , one gets  $\beta(x, \partial_s v) = |\partial_s v|^2$ . The resulting metric was introduced in Younes [69, 70] and shown to be isometric to a flat metric using a square root transform. This metric was called “ $H_0^1$ ” in Mumford and Michor [56] and further studied in Younes et al. [74]. The case  $\mu_{\tan} = 1, \mu_{\text{tsv}} = 1/4$  was considered in Mio et al. [55], Srivastava and Klassen [61], and a similar square root transform was seen to provide an isometry with a flat space in this case also. This isometry was extended to the general case in Younes [71, 73] and in Needham and Kurtek [58] (another isometry was also introduced in Bauer et al. [8]). The reader is referred to the cited references for more details on the exact expression of the isometry. Figure 3 provides an example of geodesic evolution for a hybrid metric, to be compared with Fig. 1.

#### 4. GROWTH MODELS

**4.1. Introduction.** The previous section described various metrics in shape space that are built as a regularized linearized elastic energy. Optimal paths (i.e., geodesics) associated with these metrics prefer different trajectories from those associated with the “standard” spaces discussed in Section 2, and tend to inherit some of the properties suggested by the elastic intuition. However, not all trajectories of interest need to be geodesics for some metric or satisfy a least-action principle. In particular, including external actions, with in particular possible mechanisms describing growth<sup>1</sup>, will provide shape analysis methods with additional capability of modeling transformations typically observed in biology or medicine.

<sup>1</sup>Following common terminology, we consider growth as a general shape change mechanism, also including atrophy, as a “negative growth.”

A leading model for shape change in the framework of elasticity theory introduces the notion of *morpho-elasticity* in which shapes are subject to the action of a “growth tensor,” which partly accounts for the derivative of the deformation,  $d\varphi$  (see Goriely [28] for an extensive introduction to the subject and for references). Letting  $G$  denote the growth tensor, one writes  $d\varphi = AG$ , where  $A$  completes the growth tensor to provide a valid differential  $d\varphi$ , in a way that would minimize the elastic energy (so that one applies the elastic cost to  $A^T A - \text{Id}_{\mathbb{R}^d}$  rather than to  $d\varphi^T d\varphi - \text{Id}_{\mathbb{R}^d}$ ). This approach does not necessarily lead to the trivial solution  $A = \text{Id}_{\mathbb{R}^d}$  because the growth tensor  $G$  is not necessarily “compatible”, i.e., there may not always exist a transformation  $\varphi$  such that  $d\varphi^T d\varphi = G^T G$ .

Considering small deformations, i.e., linearizing  $d\varphi = AG$  for  $\varphi$  and  $G$  close to the identity, and writing  $\varphi = \text{id}_{\mathbb{R}^d} + v$ ,  $A = \text{Id}_{\mathbb{R}^d} + a$  and  $G = \text{Id}_{\mathbb{R}^d} + g$ , one gets, simply,  $dv = a + g$ . So, for a given tensor  $g$ , the vector field  $v$  must minimize an expression of the form

$$(13) \quad \int_{\Omega} B(x, (dv + dv^T - g - g^T)/2) dx$$

where  $B$  was discussed in Section 3.2.1. There is no loss of generality in assuming that  $g$  is symmetric, which will be done in the following. The minimum of Eq. (13) is not always zero, i.e., the equation  $\frac{dv + dv^T}{2} = g$  does not always have a solution. A necessary condition (which is sufficient when  $\Omega$  is simply connected) is that  $\nabla \times (g \times \nabla g) = 0$  (row-wise curl application, followed by column-wise; see, e.g., Gonzalez and Stuart [27]).

**4.2. Riemannian viewpoint.** Returning to the Riemannian situation discussed in shape spaces, the metric was defined as  $\|v\|_{\Omega}^2 = \kappa \|v\|_V^2 + \llbracket v \rrbracket_{\Omega}^2$  with  $\llbracket v \rrbracket_{\Omega}^2$  given by the right-hand side of Eq. (13) with  $g = 0$ . One can apply the same approach here, letting

$$\llbracket v \rrbracket_{\Omega}^2 = \inf_g \int_{\Omega} B(x, (dv + dv^T)/2 - g) dx.$$

Obviously, this definition has little interest unless one restricts the space of growth tensors under consideration (otherwise,  $\llbracket v \rrbracket_{\Omega}^2 = 0$  since one can take  $g = (dv + dv^T)/2$ ). Letting  $\mathcal{G}(\Omega)$  denote a set of tensor fields ( $x \mapsto g(x) \in \text{Sym}(\mathbb{R}^d)$ ), one can define

$$\llbracket v \rrbracket_{\Omega}^2 = \inf_{g \in \mathcal{G}(\Omega)} \int_{\Omega} B(x, (dv + dv^T)/2 - g) dx$$

which is not trivial in general. If  $\mathcal{G}(\Omega)$  forms a vector space, then  $\llbracket v \rrbracket_{[\Omega]}$  is a semi-norm on  $V$ .

Note that one can also switch the focus to the growth tensor and define, for  $g \in \mathcal{G}(\Omega)$ ,

$$\|g\|_{\Omega}^2 = \min_{v \in V} \left( \kappa \|v\|_V^2 + \int_{\Omega} B(x, (dv + dv^T)/2 - g) dx \right),$$

which defines a norm on growth tensors. The introduction of the regularization by the  $V$  norm ensures that the minimum is attained at a unique  $v \in V$ , that one can denote  $v_{g,\Omega}$ , which depends linearly on  $g$  and is such that  $\kappa \|v_{g,\Omega}\|_V^2 \leq \|g\|_{\Omega}^2$ . One can therefore consider evolution equations in the form

$$\begin{cases} \partial_t \varphi(t, x) = v_{g(t), \Omega(t)}(\varphi(t, x)) \\ \Omega(t) = \varphi(t, \Omega(0)) \end{cases}$$

which are well posed (starting with  $\varphi(0, \cdot) = \text{id}_{\mathbb{R}^3}$ ) as long as

$$\int_0^1 \|g(t)\|_{\Omega(t)}^2 dt < \infty.$$

This framework therefore provides two formally equivalent optimal control problems. In the first one, one minimizes, with respect to  $v(\cdot)$

$$(14) \quad \int_0^1 \|v(t)\|_{\Omega(t)}^2 dt,$$

subject to  $\varphi(1, \Omega_0) = \Omega_1$ ,  $\varphi(0, \cdot) = \text{id}_{\mathbb{R}^3}$ ,  $\partial_t \varphi(t, \cdot) = v(t, \varphi(t, \cdot))$ ,  $\Omega(t) = \varphi(t, \Omega_0)$ . In the second one, one minimizes, with respect to  $g(\cdot)$

$$(15) \quad \int_0^1 \|g(t)\|_{\Omega(t)}^2 dt,$$

subject to  $\varphi(1, \Omega_0) = \Omega_1$ ,  $\varphi(0, \cdot) = \text{id}_{\mathbb{R}^3}$ ,  $\partial_t \varphi(t, \cdot) = v_{g(t)}(\varphi(t, \cdot))$ ,  $g(t) \in \mathcal{G}(\Omega(t))$ ,  $\Omega(t) = \varphi(t, \Omega_0)$ . Both problems are, in addition, equivalent to minimizing, with respect to both  $v(\cdot)$  and  $g(\cdot)$ ,

$$(16) \quad \kappa \int_0^1 \|v(t)\|_V^2 dt + \int_0^1 \int_{\Omega(t)} B(x, (dv(t, x) + dv(t, x)^T)/2 - g(t, x)) dx$$

subject to  $\varphi(1, \Omega_0) = \Omega_1$ ,  $\varphi(0, \cdot) = \text{id}_{\mathbb{R}^3}$ ,  $\partial_t \varphi(t, \cdot) = v(t, \varphi(t, \cdot))$ ,  $g(t) \in \mathcal{G}(\Omega(t))$ ,  $\Omega(t) = \varphi(t, \Omega_0)$ .

When  $\mathcal{G}(\Omega)$  is a vector space, the minimum value of these optimal control problems with given  $\Omega_0$  and  $\Omega_1$  is symmetric in  $\Omega_0$  and  $\Omega_1$  and its square root satisfies the triangular inequality. This minimum is always larger to that obtained with  $B = 0$  and therefore cannot be zero unless  $\Omega_0 = \Omega_1$ . (Note that the minimum can be infinite if the problem is unfeasible.) Under suitable assumptions, solutions of this optimal control problem always exist. A precise statement of this result and a sketch of its proof are provided in the appendix.

**4.3. Growth as an internal force.** Some additional notation is needed here. Denote the topological dual of a Hilbert space  $H$ , with inner product  $\langle \cdot, \cdot \rangle_H$ , by  $H^*$  and if  $\mu \in H^*$  is a linear form and if  $h \in H$ , denote their pairing by  $(\mu | h)$  (i.e.,  $\mu(h)$ ). Riesz's representation theorem gives an isometric correspondence between  $H$  and  $H^*$  with, denoting by  $K_H : H^* \rightarrow H$  the operator that associates to a linear form  $\mu$  the unique vector  $h \in H$  such that  $(\mu | \tilde{h}) = \langle h, \tilde{h} \rangle_H$  for all  $\tilde{h} \in H$ ,  $\|h\|_H^2 = (K_H^{-1}h | h)$ . This construction will be applied to  $H = V$ .

Introduce the (finite-dimensional) linear operator  $\beta(x)$  operating on symmetric  $3 \times 3$  matrices such that  $B(x, S) = \langle S, \beta(x)S \rangle$  (with  $\langle S, S' \rangle = \text{trace}(SS')$ ), and define, for a tensor field  $x \mapsto S(x)$ ,

$$\boldsymbol{\beta}_\Omega(S) = \int_\Omega \beta(x)S(x) dx.$$

Defining  $\mathbf{d}v = (dv + dv^T)/2$ , one has

$$\int_\Omega B(x, (dv + dv^T)/2 - g) dx = (\mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d}v | v) - 2(\mathbf{d}^* \boldsymbol{\beta}_\Omega g | v) + (\boldsymbol{\beta}_\Omega g | g).$$

Letting  $\mathbf{j}_{g,\Omega} = \mathbf{d}^* \boldsymbol{\beta}_\Omega g$ , one has

$$(17) \quad v_{g,\Omega} = (\kappa K_V^{-1} + \mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d})^{-1} \mathbf{j}_{g,\Omega}.$$

This relation provide an alternative way of modeling the growth process. One can indeed, following Hsieh et al. [39], directly define a “yank” (derivative of a force)  $\mathbf{j}$  as a control, with  $v_{\mathbf{j}} = (\kappa K_V^{-1} + \mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d})^{-1} \mathbf{j}$  and use the running cost

$$\int_0^1 (\mathbf{j}(t) | v_{\mathbf{j}(t)}) dt$$

with  $\partial_t \varphi_j(t, x) = v_{j(t)}(\varphi_j(t, x))$ . One can then show that the finiteness of the cost implies that the ODE has solutions over all time interval. One can also prove that optimal control  $j$  always exist in this case.

Note that this problem is different from the one described in equations Eqs. (14) to (16). In that case, one has

$$\|g\|_\Omega^2 = (\beta_\Omega g | g) - (j_g | v_g),$$

showing that the geodesics for the  $\|\cdot\|_\Omega$  metric (which remain to be explored) are likely to behave differently than those studied in Hsieh et al. [39].

**4.4. A simple example.** Assume that the growth tensor is scalar, i.e.,  $g(x) = \rho(x)\text{Id}_{\mathbb{R}^3}$  and that  $g(x) = 0$  on  $\partial\Omega$ , to avoid keeping track of boundary terms. Also assume that the elastic energy on  $\Omega$  is homogeneous and isotropic (Eq. (7)), which implies that  $B(x, g(x))$  is proportional to  $\rho(x)^2$ , the proportionality constant being, using the Lamé coefficients, equal to  $3(3\lambda/2 + \mu)$ . Letting  $\xi = 3\lambda/2 + \mu$  and using the bilinearity of  $B(x, \cdot)$  and the fact that  $\text{trace}(dv) = \text{trace}(dv^T) = \nabla \cdot v$ , a direct computation yields:

$$B(x, (dv + dv^T)/2 - g) = B(x, (dv + dv^T)/2) - 2\xi\rho(x)\nabla \cdot v(x) + 3\xi\rho(x)^2$$

Integrating by parts, one has

$$\int_\Omega \rho(x)\nabla \cdot v(x)dx = - \int_\Omega \nabla\rho(x)^T v(x)dx$$

so that, using the previous notation,

$$j_{g,\Omega} = -\xi\nabla\rho.$$

One therefore finds that

$$\|\rho\text{Id}_{\mathbb{R}^3}\|_\Omega^2 = 3 \int_\Omega \rho(x)^2 dx - \int_\Omega \nabla\rho(x)^T (\kappa K_V^{-1} + d^* \beta_\Omega d)^{-1} \nabla\rho(x) dx.$$

Similarly, the minimum in  $\rho$  of  $B(x, (dv(x) + dv(x)^T)/2 - \rho(x)\text{Id}_{\mathbb{R}^3})$  is attained at  $\rho = \nabla \cdot v/3$  and

$$\|v\|_\Omega^2 = \kappa\|v\|_V^2 + \int_\Omega B(x, (dv(x) + dv(x)^T)/2 - \nabla \cdot v(x)/3) dx$$

**4.5. Growth due to external action.** Shape variations resulting from a growth tensor as described above may be caused by external effects (e.g., impact of a disease) and do not need to follow a least-action principle such as described in the previous paragraph. More likely, the growth tensor will follow its own course, according to a process influenced by elements that are independent of the material properties of the deforming shape. The growth tensor evolution cannot be completely independent of the shape, however, since it must be supported by the time dependent domain  $\Omega(t)$ . It is also possible that changes in the geometry of the shape impact how growth behaves.

All this results in evolution systems with coupled evolution equations, typically involving moving domains. In Bressan and Lewicka [14], a scalar growth is assumed, with the relationship  $\nabla \cdot v = \rho$ , consistent with Section 4.4. The growth function depends on another function,  $u$ , representing the “concentration of morphogen”, so that  $\rho = \alpha \circ u$  for a fixed function  $\alpha$ . This morphogen concentration follows a partial differential equation (PDE), namely  $\Delta u = w - u$ , with Neumann’s boundary conditions, where  $w$  itself is a density advected by the motion, i.e., satisfying  $\partial_t w + \nabla \cdot (vw) = 0$ , which provides the coupling between growth and shape change. Initial conditions are the initial domain  $\Omega_0$  and the initial value of  $w$ ,  $w_0$ . One can then show that, when starting with

a domain  $\Omega_0$  with smooth enough boundary and with a smooth enough density  $w_0$ , a solution to the growth system exists over some finite interval  $[0, T]$  for some (small enough)  $T$ .

In Hsieh et al. [39] and Hsieh et al. [38], the additional regularization term  $\|v\|_V$  described in this chapter is added, using the formulation in Eq. (17)

$$(\kappa K_V^{-1} + \mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d})v = \mathbf{j},$$

where  $\mathbf{j}$  is the modeled control (as seen in our simple example of Section 4.4,  $\mathbf{j}$  has an interpretation similar to that of  $-\nabla\rho$ ). Hsieh et al. [39] models  $\mathbf{j}$  as a function  $\mathbf{j}(\varphi, \theta)$ , for some time-independent parameter  $\theta$ , providing coupled equations

$$\begin{cases} \partial_t \varphi(t, x) = v(t, \varphi(t, x)) \\ (\kappa K_V^{-1} + \mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d})v(t, \cdot) = \mathbf{j}(\varphi(t, \cdot), \theta) \end{cases}$$

The system is shown to have a unique solution  $t \mapsto \varphi(t, \cdot)$  over arbitrary large time intervals, for any fixed  $\theta$ , provided that  $\mathbf{j}(\varphi, \theta)$  is Lipschitz in  $\varphi$  for the  $(1, \infty)$  norm. Denoting this solution by  $\varphi(t, \cdot; \theta)$ , this property allows for the specification of optimization problems over the parameter  $\theta$  involving the transformation  $\varphi(1, \cdot; \theta)$ .

A more complex system is introduced in Hsieh et al. [38] in which  $\mathbf{j}$  is itself modeled based on a solution of a “reaction-diffusion-convection” equation on the moving domain  $\Omega$ . Ignoring a few technicalities,  $\mathbf{j}$  is given by  $\mathbf{j} = \nabla(Q(p))$  where  $Q$  is a fixed function and  $p$  satisfies

$$\partial_t p = \nabla \cdot (S_\varphi \nabla p - p v) + R(p)$$

where  $R$ , the reaction function, is fixed, and  $S_\varphi$ , the diffusion matrix, is allowed to evolve with the transformation  $\varphi$ . One can then formulate suitable conditions under which the system

$$(18) \quad \begin{cases} \partial_t \varphi(t, x) = v(t, \varphi(t, x)) \\ (\kappa K_V^{-1} + \mathbf{d}^* \boldsymbol{\beta}_\Omega \mathbf{d})v(t, \cdot) = \nabla(Q(p)) \\ \partial_t p = \nabla \cdot (S_\varphi \nabla p - p v) + R(p) \end{cases}$$

has solutions over arbitrary time intervals for a given initialization  $p_0 = p(0, \dots)$ . The determination of this initial condition for an optimal behavior at time 1 is tackled in Hsieh [36], where the existence of solutions of the optimization problem is shown. Figure 4 provides an example of growth process obtained as solution of this system.

**4.6. Constraints, deformation modules and other growth models.** Specific behavior can be enforced in a deformation process by constraining the values of the vector field at given locations in the shape. Theoretical bases for constrained and sub-Riemannian versions of LDDMM were introduced in Arguillère et al. [3], Arguillère et al. [2], Arguillère and Trélat [1], and a survey of such methods is provided in Younes et al. [75]. Among such approaches, deformation modules [32, 47] offer a generic framework in which various types of behaviors can be defined by combining suitable constraints in a modular manner. Referring to the publications above for more details, the example of “implicit elastic modules” is closely related to this chapter’s discussion. For such modules, the vector field  $v$  is obtained as a minimizer of

$$v \mapsto \lambda \|v\|_V^2 + \sum_{k=1}^m |\varepsilon_v(x_k) - S_k(h_k)|^2$$

where  $\varepsilon_v = (dv + dv^T)/2$ ,  $x_1, \dots, x_N$  are control points that are attached to (and move together with) the evolving shape, and  $h \mapsto S_k(h)$  are symmetric matrices, parametrized by a control  $h$ ,

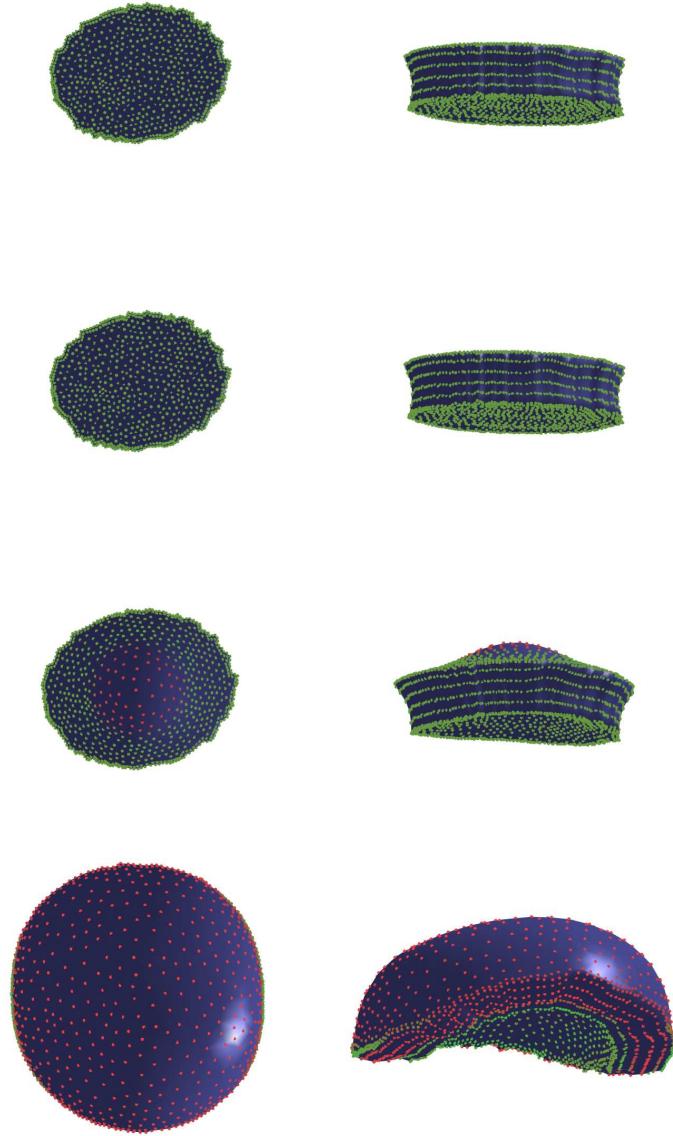


FIGURE 4. Growth model from Hsieh [36] applied to a 3D volume. Dots are colored proportionally to the magnitude of  $p$  in Eq. (18). Rows 1 to 4 provide two views of the evolving shapes at times  $t = 0$ ,  $t = 0.33$ ,  $t = 0.67$  and  $t = 1.0$ . (Images generated from code developed by Dai-Ni Hsieh.)

inducing a desired behavior (e.g., dilation) near the control points. This norm therefore introduces a finite set of (soft) constraints on the strain tensor.

A different approach at modeling growth can be found in Kaltenmark [44], Kaltenmark and Trouv   [43]. In this work, a growing shape at a given time  $t$  is defined as a transformation  $q_t$  of a co-dimension-one foliation  $X$ , which encodes the full growth process. During the evolution, only the restriction of  $q_t$  to the set  $X_t$  formed by leaves at time  $s \leq t$  of the foliation is relevant to describe the growing shape. The value of  $q_t(x)$  remains constant until  $t$  reaches the foliation index of  $x$ , so that the function  $q_0$  encodes all future initializations of the growth process. This process can be constructed through an evolution equation in the form  $\partial_t q_t = v(t, q_t)$ , and an example is developed in Kaltenmark and Trouv   [43] to model animal horn growth.

## 5. CONCLUSION

Starting from the notion of shape spaces built along the principles of Grenander's metric pattern theory and the action of diffeomorphism groups, this chapter surveyed a few recent efforts to incorporate physical constraints in the modelling of trajectories in such spaces. It first discussed the class of hybrid models that consist in combining the original shape space metric induced by the deformation group with other more physically-informed metrics, in particular those derived from linear elasticity theory. A second general approach is to further constrain shape evolution via the introduction of a growth model underlying the morphological transformation.

One of the main motivation behind all of these works is to advance the ability of shape space frameworks to model physical or biological processes, while still preserving the advantages of the geometric shape space metric setting. Indeed, this enables the formulation of the dynamics of those processes as control systems and provides adequate regularization norms to ensure existence and smoothness of solutions in many cases. Furthermore, by considering the associated optimal control problems, those same models can often lead to natural and well-posed approaches for tackling the inverse problem of e.g. determining the causes/sources of morphological changes based on some observed shape evolution. The ideas provided by the present chapter are examples of emerging efforts toward cross-fertilization between the fields of shape analysis, mathematical biology, biomedical engineering and material science.

## REFERENCES

- [1] Sylvain Arguill  re and Emmanuel Tr  lat. Sub-Riemannian structures on groups of diffeomorphisms. *Journal of the Institute of Mathematics of Jussieu*, 16(4):745–785, 2017. Publisher: Cambridge University Press.
- [2] Sylvain Arguill  re, Emmanuel Tr  lat, Alain Trouv  , and Laurent Younes. Shape deformation analysis from the optimal control viewpoint. *Journal de math  matiques pures et appliqu  es*, 104(1):139–178, 2015. Publisher: Elsevier Masson.
- [3] Sylvain Arguill  re, Emmanuel Tr  lat, Alain Trouv  , and Laurent Younes. Shape deformation and optimal control. *ESAIM: Proceedings and Surveys*, 45:300–307, 2014. Publisher: EDP Sciences.
- [4] V. I. Arnold. Sur un Principe Variationnel pour les Ecoulements Stationnaires des Liquides Parfaits et ses Applications aux Probl  mes de Stabilit   non lin  aires. *J. M  canique*, 5:29–43, 1966.
- [5] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1978.
- [6] Vladimir I Arnold and Boris A Khesin. *Topological methods in hydrodynamics*, volume 125. Springer Nature, 2021.

- [7] N. Aronszajn. Theory of Reproducing Kernels. *Trans. Am. Math. Soc.*, 68:337–404, 1950.
- [8] Martin Bauer, Martins Bruveris, Stephen Marsland, and Peter W. Michor. Constructing reparameterization invariant metrics on spaces of plane curves. *Differential Geometry and its Applications*, 34:139–165, 2014. Publisher: Elsevier.
- [9] Martin Bauer, Martins Bruveris, and Peter W Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision*, 50(1-2):60–97, 2014. Publisher: Springer.
- [10] Martin Bauer, Nicolas Charon, and Laurent Younes. Metric registration of curves and surfaces using optimal control. In *Handbook of Numerical Analysis*, volume 20, pages 613–646. Elsevier, 2019.
- [11] Martin Bauer, Philipp Harms, and Stephen C Preston. Vanishing distance phenomena and the geometric approach to SQG. *Archive for Rational Mechanics and Analysis*, 235(3):1445–1466, 2020. Publisher: Springer.
- [12] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape Matching and Object Recognition Using Shape Contexts. *IEEE Trans. PAMI*, 24(24):509–522, 2002.
- [13] B. Berkels, A. Effland, and M. Rumpf. Time Discrete Geodesic Paths in the Space of Images. *SIAM Journal on Imaging Sciences*, 8(3):1457–1488, 2015. doi: 10.1137/140970719.
- [14] Alberto Bressan and Marta Lewicka. A Model of Controlled Growth. *Archive for Rational Mechanics and Analysis*, 227(3):1223–1266, March 2018. ISSN 1432-0673.
- [15] Alexander Bronstein, Michael Bronstein, Alfred Bruckstein, and Ron Kimmel. Analysis of Two-Dimensional Non-Rigid Shapes. *International Journal of Computer Vision*, 78(1):67 – 88, 2008. ISSN 09205691.
- [16] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008.
- [17] Yan Cao, Michael I Miller, Raimond L Winslow, and Laurent Younes. Large deformation diffeomorphic metric mapping of vector fields. *IEEE transactions on medical imaging*, 24(9):1216–1230, 2005. Publisher: IEEE.
- [18] Yan Cao, Michael I Miller, Susumu Mori, Raimond L Winslow, and Laurent Younes. Diffeomorphic matching of diffusion tensor images. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pages 67–67. IEEE, 2006.
- [19] Nicolas Charon, Benjamin Charlier, Joan Glaunès, Pietro Gori, and Pierre Roussillon. Fidelity metrics between curves and surfaces: currents, varifolds, and normal cycles. In *Riemannian geometric statistics in medical image analysis*, pages 441–477. Elsevier, 2020.
- [20] Gary E. Christensen, Richard D. Rabbitt, and Michael I. Miller. Deformable templates using large deformation kinematics. *IEEE Trans. Image Proc.*, 1996.
- [21] Philippe G Ciarlet. *Three-dimensional elasticity*, volume 20. Elsevier, 1988.
- [22] Ian L Dryden and Kanti V Mardia. *Statistical shape analysis: with applications in R*, volume 995. John Wiley & Sons, 2016.
- [23] Paul Dupuis, Ulf Grenander, and Michael I. Miller. Variational Problems on Flows of Diffeomorphisms for Image Matching. *Quarterly of Applied Mathematics*, LVI(4):587–600, 1998.
- [24] D. G. Ebin and J. E. Marsden. Groups of Diffeomorphisms and the Motion of an Incompressible Fluid. *Ann. of Math*, 92:102–163, 1970.
- [25] Joan Glaunès, Alain Trouvé, and Laurent Younes. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 2, pages II–II. IEEE, 2004.

- [26] Joan Glaunès, Anqi Qiu, Michael I. Miller, and Laurent Younes. Large Deformation Diffeomorphic Metric Curve Matching. *International Journal of Computer Vision*, 80(3):317–336, 2008.
- [27] Oscar Gonzalez and Andrew M Stuart. *A first course in continuum mechanics*, volume 42. Cambridge University Press, 2008.
- [28] Alain Goriely. *The mathematics and mechanics of biological growth*, volume 45. Springer, 2017.
- [29] Ulf Grenander. *General Pattern Theory*. Oxford Science Publications, 1993.
- [30] Ulf Grenander and Daniel M. Keenan. On the shape of plane images. *Siam J. Appl. Math.*, 53(4):1072–1094, 1991.
- [31] Ulf Grenander and Michael I Miller. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics*, 56(4):617–694, 1998.
- [32] Barbara Gris, Stanley Durrleman, and Alain Trouvé. A Sub-Riemannian Modular Framework for Diffeomorphism-Based Analysis of Shape Ensembles. *SIAM Journal on Imaging Sciences*, 11(1):802–833, January 2018. Publisher: Society for Industrial and Applied Mathematics.
- [33] Xianfeng Gu, Yalin Wang, Tony F Chan, Paul M Thompson, and Shing-Tung Yau. Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE transactions on medical imaging*, 23(8):949–958, 2004.
- [34] Xianfeng David Gu and Shing-Tung Yau. *Computational conformal geometry*, volume 1. International Press Somerville, MA, 2008.
- [35] Darryl D Holm, Jerrold E Marsden, and Tudor S Ratiu. The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81, 1998.
- [36] Dai-Ni Hsieh. *On model-based diffeomorphic shape evolution and diffeomorphic shape registration*. PhD thesis, Johns Hopkins University, 2021.
- [37] Dai-Ni Hsieh, Sylvain Arguillère, Nicolas Charon, Michael I Miller, and Laurent Younes. A model for elastic evolution on foliated shapes. In *International Conference on Information Processing in Medical Imaging*, pages 644–655. Springer, Cham, 2019.
- [38] Dai-Ni Hsieh, Sylvain Arguillère, Nicolas Charon, and Laurent Younes. Diffeomorphic shape evolution coupled with a reaction-diffusion PDE on a growth potential. *Quarterly of Applied Mathematics*, 2021. ISSN 0033-569X, 1552-4485. doi: 10.1090/qam/1600.
- [39] Dai-Ni Hsieh, Sylvain Arguillère, Nicolas Charon, and Laurent Younes. Mechanistic Modeling of Longitudinal Shape Changes: equations of motion and inverse problems. *SIAM Journal on Applied Dynamical Systems*, 21(1):80–101, 2022. Publisher: SIAM.
- [40] Tuomas Hytönen, Jan Van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces*, volume 12. Springer, 2016.
- [41] Ian H Jermyn, Sebastian Kurtek, Eric Klassen, and Anuj Srivastava. Elastic shape matching of parameterized surfaces using square root normal fields. In *European conference on computer vision*, pages 804–817. Springer, 2012.
- [42] Hachem Kadri, Emmanuel Duflos, Philippe Preux, Stéphane Canu, Alain Rakotomamonjy, and Julien Audiffren. Operator-valued kernels for learning from functional response data. *Journal of Machine Learning Research*, 17(20):1–54, 2016.
- [43] Irène Kaltenmark and Alain Trouvé. Estimation of a growth development with partial diffeomorphic mappings. *Quarterly of Applied Mathematics*, 77(2):227–267, 2019.
- [44] Irène Kaltenmark. *Geometrical Growth Models for Computational Anatomy*. PhD thesis, Université Paris-Saclay (ComUE), 2016.

- [45] David G. Kendall. Shape manifolds, Procrustean metrics and complex projective spaces. *Bull. London Math. Soc.*, 16:81–121, 1984.
- [46] Eric P. Klassen, Anuj Srivastava, Washington Mio, and Shantanu H. Joshi. Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(3):372–383, 2004. ISSN 0162-8828.
- [47] Leander Lacroix, Benjamin Charlier, Alain Trouvé, and Barbara Gris. IMODAL: creating learnable user-defined deformation models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12905–12913, 2021.
- [48] Lok Ming Lui, Wei Zeng, Shing-Tung Yau, and Xianfeng Gu. Shape analysis of planar multiply-connected objects using conformal welding. *IEEE transactions on pattern analysis and machine intelligence*, 36(7):1384–1401, 2014. Publisher: IEEE.
- [49] Facundo Mémoli. Gromov–wasserstein distances and the metric approach to object matching. *Foundations of computational mathematics*, 11(4):417–487, 2011.
- [50] Peter W. Michor and David Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Mathematica*, 10:217–245, 2005.
- [51] Peter W Michor and David Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.
- [52] Michael I. Miller and Laurent Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1-2):61–84, 2001. Publisher: Kluwer Academic Publishers.
- [53] Michael I. Miller, Sarang C. Joshi, and Gary E. Christensen. Large deformation fluid diffeomorphisms for landmark and image matching. In A. Toga, editor, *Brain Warping*, pages 115–131. Academic Press, 1999.
- [54] Michael I. Miller, . Trouvé, and Laurent Younes. Hamiltonian systems and optimal control in computational anatomy: 100 years since D’Arcy Thompson. *Annual review of biomedical engineering*, 17:447–509, 2015. Publisher: Annual Reviews.
- [55] Washington Mio, Anuj Srivastava, and Shantanu Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73(3):307–324, 2007. Publisher: Springer.
- [56] David B Mumford and Peter W Michor. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8(1):1–48, 2006.
- [57] Facundo Mémoli. Gromov-Hausdorff distances in Euclidean spaces. In *CVPR workshop on nonrigid shape analysis*, 2008.
- [58] Tom Needham and Sebastian Kurtek. Simplifying transforms for general elastic metrics on the space of plane curves. *SIAM Journal on Imaging Sciences*, 13(1):445–473, 2020.
- [59] Maks Ovsjanikov, Quentin Mérigot, Facundo Mémoli, and Leonidas Guibas. One point isometric matching with the heat kernel. In *Computer Graphics Forum*, volume 29-5, pages 1555–1564. Wiley Online Library, 2010.
- [60] Richard S Palais. *Foundations of global non-linear analysis*. Benjamin New York, 1968.
- [61] Anuj Srivastava and Eric P. Klassen. *Functional and shape data analysis*. Springer, 2016.
- [62] Zhe Su, Martin Bauer, Stephen C Preston, Hamid Laga, and Eric Klassen. Shape analysis of surfaces using general elastic metrics. *Journal of Mathematical Imaging and Vision*, 62(8):1087–1106, 2020.
- [63] D’Arcy Wendworth Thompson. *On Growth and Form*. Dover Publications, 1917.
- [64] Alain Trouvé. Action de groupe de dimension infinie et reconnaissance de formes. *Comptes Rendus de l’Académie des Sciences. Série I. Mathématique*, 321(8):1031–1034, 1995. ISSN 0764-4442.

0764-4442.

[65] Alain Trouv . Diffeomorphism groups and pattern matching in image analysis. *Int. J. of Comp. Vis.*, 28(3):213–221, 1998.

[66] Alain Trouv  and Laurent Younes. Metamorphoses through lie group action. *Foundations of Computational Mathematics*, 5(2):173–198, 2005. Publisher: Springer-Verlag.

[67] Marc Vaillant and Joan Glaun s. Surface Matching via Currents. In G. E. Christensen and M. Sonka, editors, *Proceedings of Information Processing in Medical Imaging (IPMI 2005)*, Lecture Notes in Computer Science. Springer, 2005. Issue: 3565.

[68] Benedikt Wirth, Leah Bar, Martin Rumpf, and Guillermo Sapiro. A Continuum Mechanical Approach to Geodesics in Shape Space. *International Journal of Computer Vision*, 93(3):293–318, July 2011. ISSN 1573-1405. doi: 10.1007/s11263-010-0416-9.

[69] Laurent Younes. A distance for elastic matching in object recognition. *Comptes rendus de l'Ac ademie des sciences. S  rie 1, Math  matique*, 322(2):197–202, 1996.

[70] Laurent Younes. Computable elastic distances between shapes. *SIAM Journal on Applied Mathematics*, 58(2):565–586, 1998. Publisher: Society for Industrial and Applied Mathematics.

[71] Laurent Younes. Elastic distance between curves under the metamorphosis viewpoint. *arXiv preprint arXiv:1804.10155*, 2018.

[72] Laurent Younes. Hybrid riemannian metrics for diffeomorphic shape registration. *Annals of Mathematical Sciences and Applications*, 3(1):189–210, 2018.

[73] Laurent Younes. *Shapes and Diffeomorphisms*. Applied Mathematical Sciences. Springer-Verlag, Berlin Heidelberg, 2 edition, 2019. ISBN 978-3-662-58495-8. doi: 10.1007/978-3-662-58496-5.

[74] Laurent Younes, Peter W. Michor, Jayant Shah, and David Mumford. A metric on shape space with explicit geodesics. *Rend. Lincei Math. Appl.*, 19:25–57, 2008.

[75] Laurent Younes, Barbara Gris, and Alain Trouv . Sub-Riemannian Methods in Shape Analysis. In *Handbook of Variational Methods for Nonlinear Geometric Data*, pages 463–495. Springer, Cham, 2020.

[76] Wei Zeng and Xianfeng David Gu. Registration for 3D surfaces with large deformations using quasi-conformal curvature flow. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2457–2464. IEEE, 2011.

[77] Wei Zeng, Lok Ming Lui, Feng Luo, Tony Fan-Cheong Chan, Shing-Tung Yau, and David Xianfeng Gu. Computing quasiconformal maps using an auxiliary metric and discrete curvature flow. *Numerische Mathematik*, 121(4):671–703, 2012. Publisher: Springer.

#### APPENDIX A. ELASTIC SURFACE METRIC AS THE LIMIT OF THE LAMINAR MODEL (SECTION 3.2.2)

Given an oriented surface  $\mathcal{M}_0$  in  $\mathbb{R}^3$ , and its unit normal vector field denoted  $\nu_0$ , one can generate a foliated 3D volume as the set of points  $\Phi(s, x_0) = x_0 + s\delta\nu_0(x_0)$ ,  $x_0 \in \mathcal{M}_0$ ,  $s \in [0, 1]$ , and  $\Phi$  is a diffeomorphism for small enough  $\delta > 0$ . In this case, the unit normal  $N$  to the layer  $\mathcal{M}_s = \Phi(\{s\} \times \mathcal{M}_0)$  at the point  $x = \Phi(x_0, s) \in \Omega$  is also  $N(x) = \nu_0(x_0)$ . It coincides, up to a factor  $\delta$ , with  $S = \partial_s \Phi$  and satisfies  $dN = 0$ . Let  $v_0 : \mathcal{M}_0 \rightarrow \mathbb{R}^3$  be a vector field on  $\mathcal{M}_0$ , and define its extension  $v$  to  $\Omega$  by  $v(\Phi(s, x_0)) = v_0(x_0)$ , so that  $v$  satisfies  $dvN = 0$ . See Fig. 5 for an illustration. Let  $\sigma_0 = d\nu_0$  denote the shape operator on the surface  $\mathcal{M}_0$  and similarly  $\sigma_s$  the shape operator of layer  $\mathcal{M}_s$  i.e. the restriction of  $dN$  to the tangent space of  $\mathcal{M}_s$ . Recall that the shape operator on a surface is a symmetric operator.

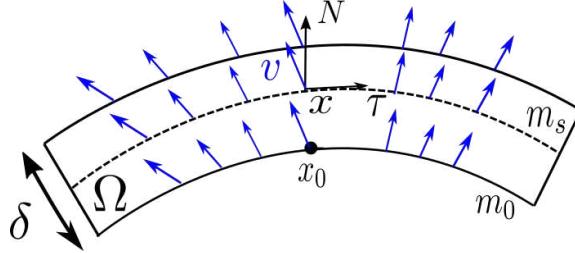


FIGURE 5. Cross-sectional schematic representation of the thin-shell layered elastic domain with the deformation field  $v$  in blue.

Write  $v = v_T + v_N N$  where  $v_T \in \mathbb{R}^3$  is tangent to the layers and  $v_N$  is scalar. If  $\tau, \tilde{\tau}$  are vectors tangent to the layers, we have

$$(19) \quad \tilde{\tau}^T dv\tau = \tilde{\tau}^T dv_T\tau + (\nabla v_N^T \tau)(\tilde{\tau}^T N) + v_N \tilde{\tau}^T dN\tau = \tilde{\tau}^T dv_T\tau + v_N \tilde{\tau}^T dN\tau$$

In particular, letting  $\varepsilon_T = (dv_T + dv_T^T)/2$ , and for  $\{\tau_1, \tau_2\}$  an orthonormal basis of the tangent plane to  $\mathcal{M}_s$  at  $x$ , one has:

$$\tau_1^T \varepsilon \tau_1 + \tau_2^T \varepsilon \tau_2 = \tau_1^T dv_T \tau_1 + \tau_2^T dv_T \tau_2 + v_N [\tau_1^T dN \tau_1 + \tau_2^T dN \tau_2].$$

With the sum of the first two terms, one recognizes the divergence of  $v_T$  on the surface  $\mathcal{M}_s$  which will be denoted by  $\nabla_{\mathcal{M}_s} \cdot v_T$ . Similarly, the term within brackets is the divergence of the shape operator on  $\mathcal{M}_s$  which equals  $-2H_{\mathcal{M}_s}$  where  $H_{\mathcal{M}_s}$  is the mean curvature of  $\mathcal{M}_s$ . Therefore, one deduces that, on  $\mathcal{M}_s$ :

$$\tau_1^T \varepsilon \tau_1 + \tau_2^T \varepsilon \tau_2 = \nabla_{\mathcal{M}_s} \cdot v_T - 2v_N H_{\mathcal{M}_s}.$$

Moreover, as  $dvN = 0$ , it follows that  $N^T \varepsilon N = 0$  and thus, on  $\mathcal{M}_s$ :

$$\text{trace}(\varepsilon) = \tau_1^T \varepsilon \tau_1 + \tau_2^T \varepsilon \tau_2 + N^T \varepsilon N = \nabla_{\mathcal{M}_s} \cdot v_T - 2v_N H_{\mathcal{M}_s}.$$

Similarly, looking at the second term in Eq. (8) and using Eq. (19), one has:

$$\begin{aligned} & (\tau_1^T \varepsilon \tau_1)^2 + (\tau_2^T \varepsilon \tau_2)^2 + 2(\tau_1^T \varepsilon \tau_2)^2 \\ &= (\tau_1^T dv_T \tau_1 + v_N \tau_1^T \sigma_s \tau_1)^2 + (\tau_2^T dv_T \tau_2 + v_N \tau_2^T \sigma_s \tau_2)^2 + 2(\tau_1^T \varepsilon_T \tau_2 + v_N \tau_1^T \sigma_s \tau_2)^2 \\ &= (\tau_1^T dv_T \tau_1)^2 + (\tau_2^T dv_T \tau_2)^2 + 2(\tau_1^T \varepsilon_T \tau_2)^2 \\ & \quad + v_N^2 [(\tau_1^T \sigma_s \tau_1)^2 + (\tau_2^T \sigma_s \tau_2)^2 + 2(\tau_1^T \sigma_s \tau_2)^2] \\ & \quad + 2v_N [(\tau_1^T dv_T \tau_1)(\tau_1^T \sigma_s \tau_1) + (\tau_2^T dv_T \tau_2)(\tau_2^T \sigma_s \tau_2) + 2(\tau_1^T \varepsilon_T \tau_2)(\tau_1^T \sigma_s \tau_2)]. \end{aligned}$$

In this computation, one uses the fact that the operator  $dN$  restricted to the tangent space to  $\mathcal{M}_s$  at  $x$  (i.e. the space spanned by  $\tau_1$  and  $\tau_2$ ) coincides with  $\sigma_s$ . Now, by symmetry, one has  $\tau_1^T dv_T \tau_1 = \tau_1^T \varepsilon_T \tau_1$  and  $\tau_2^T dv_T \tau_2 = \tau_2^T \varepsilon_T \tau_2$ . Moreover, recalling that for any  $2 \times 2$  symmetric tensors  $\omega$  and  $\tilde{\omega}$ , one has  $\text{trace}(\omega \tilde{\omega}) = \omega_{1,1} \tilde{\omega}_{1,1} + \omega_{2,2} \tilde{\omega}_{2,2} + 2\omega_{1,2} \tilde{\omega}_{1,2}$ , one gets:

$$\begin{aligned} & (\tau_1^T \varepsilon \tau_1)^2 + (\tau_2^T \varepsilon \tau_2)^2 + 2(\tau_1^T \varepsilon \tau_2)^2 = \text{trace}(\varepsilon_T^2) + v_N^2 \text{trace}(\sigma_s^2) + 2v_N \text{trace}(\varepsilon_T \sigma_s) \\ &= \text{trace}((\varepsilon_T + v_N \sigma_s)^2). \end{aligned}$$

Now, using the symmetry of  $\varepsilon$  and the fact that  $N^T \varepsilon N = 0$ :

$$(20) \quad N^T \varepsilon^2 N = |\varepsilon N|^2 = (\tau_1^T \varepsilon N)^2 + (\tau_2^T \varepsilon N)^2.$$

If  $\tau$  is tangent to the layers, one has  $\tau^T dvN = 0$  and

$$(21) \quad \begin{aligned} \tau^T dv^T N &= N^T dv\tau = N^T dv_T \tau + (\nabla v_N^T \tau)(N^T N) + v_N N^T dN\tau \\ &= N^T dv_T \tau + \nabla v_N^T \tau. \end{aligned}$$

Moreover, since  $N^T v_T = 0$ , it follows that  $N^T dv_T \tau = -v_T^T dN\tau$ . Using this together with Eq. (20), Eq. (21), and with the fact that  $dN$  is symmetric, one deduces that:

$$\begin{aligned} N^T \varepsilon^2 N - (N^T \varepsilon N)^2 &= ((-dN^T v_T + \nabla v_N)^T \tau_1)^2 + ((-dN^T v_T + \nabla v_N)^T \tau_2)^2 \\ &= | -\sigma_s v_T + \nabla_{\mathcal{M}_s} v_N |^2 \end{aligned}$$

where  $\nabla_{\mathcal{M}_s}$  is the gradient operator on  $\mathcal{M}_s$ .

Based on all the above expressions, one can finally rewrite Eq. (8) at  $x = \Phi(x_0, s)$  as

$$(22) \quad \begin{aligned} B(x, \varepsilon) &= \lambda_{\tan} (\nabla_{\mathcal{M}_s} \cdot v_T - 2H_{\mathcal{M}_s} v_N)^2 + \mu_{\tan} \text{trace}((\varepsilon_T + v_N \sigma_s)^2) \\ &\quad + 2\mu_{\text{ang}} | -\sigma_s v_T + \nabla_{\mathcal{M}_s} v_N |^2, \end{aligned}$$

and using by a change of variables in the integral expression of the energy, one further has:

$$\frac{1}{\delta} \int_{\Omega} B(x, \varepsilon) dx = \frac{1}{\delta} \int_0^1 \int_{\mathcal{M}_0} B(x_0 + s\delta\nu_0, \varepsilon) |J_{\Phi}(s, x_0)| d\text{vol}_{m_0}(x_0) ds$$

where  $|J_{\Phi}(s, x_0)|$  denotes the Jacobian determinant of  $\Phi$  at  $(s, x_0)$ . As  $\partial_s \Phi(s, x_0) = \delta\nu_0(x_0)$  and  $d_{x_0} \Phi(s, x_0) = \text{Id} + s\delta d\nu_0(x_0)$ , one gets  $d_{x_0} \Phi(0, x_0) = \text{Id}$  where  $\text{Id}$  denotes here the identity on the tangent space to  $m_0$  at  $x_0$ . Therefore,  $|J_{\Phi}(0, x_0)| = \delta$  for all  $x_0 \in m_0$ . Consequently, taking the limit  $\delta \rightarrow 0$  in the above and using the continuity of  $B$  and  $J_{\Phi}$  leads to the following expression of the elastic metric on the surface  $\mathcal{M}_0$ :

$$\llbracket v \rrbracket_{\mathcal{M}_0}^2 = \int_{m_0} B(x_0, \varepsilon) d\text{vol}_{\mathcal{M}_0}(x_0)$$

with  $B$  given by Eq. (22) (with  $s = 0$ ). Furthermore, it can be easily checked, based on their expressions in the frame  $(\tau_1, \tau_2, N)$ , that the three terms in  $B(x_0, \varepsilon)$  correspond precisely, up to multiplicative constants, to the ones of Eq. (10) thus showing that the elastic metric in Eq. (9) can be also recovered as the thin shell limit of the 3D laminar model introduced in Section 3.2.1.

## APPENDIX B. EXISTENCE OF OPTIMAL PATHS (SECTION 4.2)

Considering the minimization problem introduced in Eqs. (14) to (16), this section proves that, under suitable assumptions, optimal solutions exist. These assumptions are as follows.

(1) Let  $p \geq 1$ . The Hilbert space  $V$  is continuously embedded in the Banach space  $C_0^p(\mathbb{R}^3, \mathbb{R}^3)$  of  $p$  times continuously differentiable vector fields that vanish (with their first  $p$  derivatives) at infinity, with the norm

$$\|v\|_{p, \infty} = \sum_{k=0}^p \max\{|d^k v(x)| : x \in \mathbb{R}^3\}.$$

(2)  $V$  is also continuously embedded in  $H^1(\mathbb{R}^3, \mathbb{R}^3)$ , the Sobolev space of square-integrable functions with square-integrable first derivatives.

(3) The mapping  $x \mapsto B(x, \cdot)$  from  $\mathbb{R}^3$  to the set of positive semi-definite quadratic forms is continuous in  $x$ . In particular,  $|B(x, \cdot)|$  is bounded on compact subsets of  $\mathbb{R}^3$ .

(4) There exists a constant  $c$  such that  $B(x, S) \geq c|S|^2$  for all  $S \in \text{Sym}$  and all  $x \in \mathbb{R}^3$ .

(5) The sets  $\mathcal{G}(\Omega)$ , defined over compact subsets of  $\Omega \subset \mathbb{R}^3$ , satisfy the following conditions.

(5)-i If  $\Omega \subset \tilde{\Omega}$ , then  $\mathcal{G}(\Omega) \subset \mathcal{G}(\tilde{\Omega})$ .

(5)-ii Define, for  $\delta > 0$ ,  $\Omega^\delta = \{x : \text{dist}(x, \Omega) \leq \delta\}$ . Then  $\bigcap_{\delta > 0} \mathcal{G}(\Omega^\delta) = \mathcal{G}(\Omega)$ .

(5)-iii  $\mathcal{G}(\Omega)$  is a strongly closed convex subset of  $H_{Sym} := L^2(\mathbb{R}^3, Sym(\mathbb{R}^3))$ .

For example the sets  $\mathcal{G}(\Omega) = \{g \text{Id}_{\mathbb{R}^3} : g \in L^2(\Omega)\}$  satisfy condition (5).

Making these assumptions, let  $v_n(\cdot) \in L^2([0, 1], V)$  and  $g_n \in L^2([0, 1], H_{Sym})$  be minimizing sequences for the considered problem. To shorten notation, let  $\varepsilon_n = (dv_n + dv_n^T)/2$ . Because  $v_n$  is bounded in  $L^2([0, 1], V)$ , one can replace it by a subsequence that converges weakly to some  $v$  in that space, and using arguments developed in Dupuis et al. [23], Trouvé [64], Younes [73], the flows  $\varphi_n$  associated with  $v_n$  converge uniformly in time and uniformly on compact sets in space to the flow  $\varphi$  associated with  $v$ . From weak convergence and weak lower semicontinuity of the norm, one has

$$\int_0^1 \|v\|_V^2 dt \leq \liminf \int_0^1 \|v_n\|_V^2 dt$$

and from the convergence of the flows, one has  $\varphi(1, \Omega_0) = \Omega_1$  because this holds for each  $\varphi_n$ .

Based on the assumptions made on  $B$ , one has, for all  $x \in \mathbb{R}^3$  and  $t \in [0, 1]$ :

$$\begin{aligned} c|g_n(t, x)|^2 &\leq B(x, g_n(t, x)) \leq \left( B(x, \varepsilon_n(t, x) - g_n(t, x))^{1/2} + B(x, \varepsilon_n(t, x))^{1/2} \right)^2 \\ &\leq 2(B(x, \varepsilon_n(t, x) - g_n(t, x)) + B(x, \varepsilon_n(t, x))) \end{aligned}$$

Because of the convergence of  $\varphi_n$ , there exists a compact set  $\bar{\Omega} \subset \mathbb{R}^d$  that contains all the  $\Omega_n(t)$ ,  $n \in \mathbb{N}$ ,  $t \in [0, 1]$ . This implies that there exist constants  $C, C'$  such that, for all  $n \in \mathbb{N}$  (using the boundedness of  $B(x, \cdot)$  on compact sets):

$$\int_0^1 \int_{\Omega_n(t)} B(x, \varepsilon_n(t, x)) dx dt \leq C \int_0^1 \int_{\Omega_n(t)} |\varepsilon_n(t, x)|^2 dx dt \leq C' \int_0^1 \|v_n(t)\|_{H^1}^2 dt.$$

By the continuous embedding of  $V$  into  $H^1$ ,  $\|v_n(t)\|_{H^1}$  is bounded up to a multiplicative constant by  $\|v_n(t)\|_V$ , which implies that the above term is bounded independently of  $n$ . The same holds for:

$$\int_0^1 \int_{\Omega_n(t)} B(x, \varepsilon_n - g_n) dx dt = \frac{1}{4} \int_0^1 \int_{\Omega_n(t)} B(x, dv_n + dv_n^T - 2g_n) dx dt$$

as  $(v_n, g_n)$  is a minimizing sequence for the functional in Eq. (16). This implies that the sequence  $\int_0^1 \|g_n\|_{H_{Sym}}^2 dt$  is bounded and that one can assume, using a subsequence if needed, that  $g_n \rightharpoonup g$  in  $L^2([0, 1], H_{Sym})$ .

It remains to prove that  $g(t) \in \mathcal{G}(\Omega(t))$  to show that  $(v, g)$  provides a solution of the minimization problem. Fixing  $\delta > 0$ , one can restrict the minimizing sequence to those large enough  $n$  for which  $\max\{|\varphi_n(t, x) - \varphi(t, x)|, t \in [0, 1], x \in \bar{\Omega}\} < \delta$ , so that  $\Omega_n(t) \subset \Omega^\delta(t)$  for all  $n$  and  $t$ .

Let

$$\Gamma(\Omega(\cdot), \delta) = \{\tilde{g}(\cdot) : \tilde{g}(t) \in \mathcal{G}(\Omega^\delta(t)), \text{ for a.e } t \in [0, 1]\},$$

so that  $g_n \in \Gamma(\Omega(\cdot), \delta)$ . This is a convex set, which follows directly from our hypotheses on the sets  $\mathcal{G}(\Omega)$ , and it is closed in  $L^2([0, 1], H_{Sym})$ . Indeed, if  $\tilde{g}_n \in \Gamma(\Omega(\cdot), \delta)$  converges to  $\tilde{g} \in L^2([0, 1], H_{Sym})$ , then a subsequence converges for almost all  $t \in [0, 1]$  and since each  $\mathcal{G}(\Omega^\delta(t))$  is closed in  $H_{Sym}$ , it results that  $\tilde{g}(t) \in \mathcal{G}(\Omega^\delta(t))$  for almost all  $t$ . Now, as strongly closed convex sets are also weakly closed in  $L^2([0, 1], H_{Sym})$  (see, Hytönen et al. [40]), one deduces from  $g_n \rightharpoonup g$  that  $g \in \Gamma(\Omega(\cdot), \delta)$ . Since this is true for all  $\delta > 0$ , one has, taking a sequence  $\delta_n \rightarrow 0$ , that  $g(t) \in \mathcal{G}(\Omega(t))$  for almost all  $t \in [0, 1]$ .

This concludes the proof that  $(v, g)$  is a minimizer of Eq. (16).

N. CHARON. CENTER FOR IMAGING SCIENCE, JOHNS HOPKINS UNIVERSITY  
*Email address:* `ncharon1@jhu.edu`

L. YOUNES. CENTER FOR IMAGING SCIENCE, JOHNS HOPKINS UNIVERSITY  
*Email address:* `laurent.younes@jhu.edu`