

A RELAXED APPROACH FOR CURVE MATCHING WITH ELASTIC METRICS

MARTIN BAUER^{1,*}, MARTINS BRUVERIS², NICOLAS CHARON³
AND JAKOB MØLLER-ANDERSEN¹

Abstract. In this paper, we study a class of Riemannian metrics on the space of unparametrized curves and develop a method to compute geodesics with given boundary conditions. It extends previous works on this topic in several important ways. The model and resulting matching algorithm integrate within one common setting both the family of H^2 -metrics with constant coefficients and scale-invariant H^2 -metrics on both open and closed immersed curves. These families include as particular cases the class of first-order elastic metrics. An essential difference with prior approaches is the way that boundary constraints are dealt with. By leveraging varifold-based similarity metrics we propose a relaxed variational formulation for the matching problem that avoids the necessity of optimizing over the reparametrization group. Furthermore, we show that we can also quotient out finite-dimensional similarity groups such as translation, rotation and scaling groups. The different properties and advantages are illustrated through numerical examples in which we also provide a comparison with related diffeomorphic methods used in shape registration.

Mathematics Subject Classification. 68Q25, 68R10, 68U05.

Received April 1, 2018. Accepted September 23, 2018.

1. INTRODUCTION

In this article, we study Riemannian metrics on the space of unparametrized, \mathbb{R}^d -valued curves. The interest in this topic is fueled by applications in medical imaging, computer animation, geometric morphometry and other fields [45]. The space of closed curves is important in shape analysis where it is used to study objects that can be represented by the shape of their boundary [4, 30, 33, 51]. At the same time open curves are relevant in applications such as the analysis of hurricane paths, bird migration patterns [47, 49], and human character motions [9, 24] or in character and speech recognition [48].

The analysis of shapes and their differences relies on the notion of a distance. To define such a distance, we will start with a Riemannian metric on the space of curves and use the induced geodesic distance to quantify differences between curves. Mathematically we model curves as smooth mappings from a parameter space to \mathbb{R}^d . The parameter space is S^1 for closed curves and $[0, 2\pi]$ for open curves. On the space of curves we can consider the action of the reparametrization group and we consider two curves equivalent if they only differ by

Keywords and phrases: Shape analysis, curve matching, intrinsic metrics, varifolds.

¹ Florida State University, Tallahassee, FL, USA.

² Brunel University London, London, UK.

³ Johns Hopkins University, Baltimore, MD, USA.

* Corresponding author: bauer@math.fsu.edu

a reparametrization, *i.e.*, $c_1 \sim c_2$ if $c_1 = c_2 \circ \varphi$ for some reparametrization φ . To define a Riemannian metric on the space of unparametrized curves we will start with a reparametrization invariant metric on the space of parametrized curves and consider the induced metric on the quotient space of unparametrized curves.¹

The simplest invariant metric on the space of parametrized curves is the L^2 -metric

$$G_c(h, k) = \int_{M^1} \langle h, k \rangle \, ds,$$

where c is the curve, h, k are tangent vectors to c , $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, $ds = |c'| \, d\theta$ is integration with respect to arc length and M^1 is either S^1 or $[0, 2\pi]$. It came as a big surprise when Michor and Mumford [35] found in that this metric induces vanishing geodesic distance on the space of unparametrized curves.² Here, vanishing geodesic distance means that given any two curves there exist paths of arbitrary short length connecting them and consequently the geodesic distance, which is defined as the infimum over all path lengths, is identically zero. This result renders the L^2 -metric impractical for applications in shape analysis and thus started the quest for stronger and more meaningful metrics: Michor and Mumford proposed curvature weighted versions of the L^2 -metric [35] and Shah [43] studied length weighted versions to successfully overcome the degeneracy of vanishing geodesic distance. A more promising approach, first investigated out of purely theoretical interest, is to include derivatives of the tangent vector in the Riemannian metric, yielding the class of Sobolev metrics [34, 36]

$$G_c(h, k) = \int_{M^1} \left\langle h, k \right\rangle \, ds + \left\langle D_s^n h, D_s^n k \right\rangle \, ds, \quad n \geq 1,$$

where $D_s = \frac{1}{|c'|} \partial_\theta$ denotes derivative with respect to arc length. Closely related to first order Sobolev metrics is the family of elastic $G^{a,b}$ -metrics, proposed as a model for shape analysis in [38, 39]. These are metrics of the form

$$G_c(h, k) = \int_{M^1} a \left\langle D_s h^\top, D_s k^\top \right\rangle + b \left\langle D_s h^\perp, D_s k^\perp \right\rangle \, ds.$$

Here $a, b > 0$ are constants and $D_s h^\top = \langle D_s h, D_s c \rangle D_s c$ and $D_s h^\perp = D_s h - D_s h^\top$ denote the decomposition of $D_s h$ into vectors tangent and orthogonal to the curve c . The first term can be interpreted as penalizing stretching of the curve, whereas the second term measures bending of the curve. Thus, one expects to be able to model a variety of behaviors by varying the constants a and b . However, so far only two special cases have been implemented numerically³: Younes *et al.* [52] studied the metric for planar curves when $a = b = 1$ and Srivastava *et al.* [30, 46] studied it for \mathbb{R}^d -valued curves when $a = 1$ and $b = 1/4$. In both cases, there exist transformations that allow one to compute explicit formulas for geodesics and geodesic distance. These transformations allowed for the development of efficient and fast numerical algorithms both on parametrized and unparametrized curves and have been proven successful for applications in shape analysis. In [6], these transformations have been extended to a wider class of parameters and metrics. In an upcoming preprint Kurtek and Needham [32] propose a different numerical framework for the class of $G^{a,b}$ -metrics based on a generalization of the transformation of Younes *et al.* [52].

It has been recently shown in [13, 15, 40], that adding second derivatives to the metric allows one to obtain completeness results for the Riemannian manifolds in question: the geodesic equation is globally well-posed,

¹Note, that the invariance of the metric is only a necessary condition for a metric on the space of parametrized curves to induce a metric on unparametrized curves. However, all the metrics considered in this article do induce Riemannian metrics on the quotient space. For details, see [36].

²This result has been later extended to the space of parametrized curves in [3].

³The article [37] studied the family of elastic metrics with different choices of parameters on the space of *parametrized* curves. This analysis was, however, not extended to the space of *unparametrized* curves, which is the more relevant object for applications.

the metric completion consists of all H^2 -immersions, and the metric extends to a strong Riemannian metric on the space of H^2 -immersions. Furthermore, any two curves in the same connected component can be joined by a minimizing geodesic. For a more detailed overview of various Riemannian metrics on the space of curves we refer to the overview articles [7, 8].

1.1. Contributions of the article

In this article, we present the first numerical implementation of the geodesic initial and boundary value problem for a family of first and second order metrics on the space of open and closed unparametrized curves. Our code is available under an open source license.⁴ This family includes in particular the elastic $G^{a,b}$ -metrics with arbitrary parameters a and b as well as scale-invariant Sobolev metrics. To be precise, we study metrics of the form

$$G_c(h, k) = \int_{M^1} a_0(\ell_c) \langle h, k \rangle + a_1(\ell_c) \langle D_s h^\top, D_s k^\top \rangle + b_1(\ell_c) \langle D_s h^\perp, D_s k^\perp \rangle + a_2(\ell_c) \langle D_s^2 h, D_s^2 k \rangle \, ds.$$

where $a_0, a_1, b_1, a_2 \in C^\infty(\mathbb{R}_{>0}, \mathbb{R}_{\geq 0})$ are smooth positive functions of the curve length ℓ_c . All metrics in this class are invariant with respect to Euclidean motions thus they induce Riemannian metrics on the shape space of unparametrized curves modulo Euclidean motions. If the coefficient functions a_j and b_1 are chosen appropriately one obtains scale-invariant metrics, which then induce Riemannian metrics on the space of unparametrized curves modulo similarity transformations.

In future applications, this numerical framework will allow us to choose the constants of the metric in a data-driven way as opposed to the *ad hoc* methods employed currently. As a first step towards this goal we will show in selected examples how the choice of constants influences minimal geodesics between two shapes and how it affects the point-to-point registration. In the experiments section, we also compare this intrinsic approach to curve matching with the LDDMM framework, where differences between curves are measured extrinsically using a Riemannian metric on the diffeomorphism group of the ambient space.

On the space of closed curves we additionally extend the completeness results, that were obtained first in [13, 15] for Sobolev metrics with constant coefficients and then in [16] for length-weighted Sobolev metrics to this class of elastic metrics. For open curves we find a counter example showing that second order metrics with constant coefficients are not metrically complete.

From a numerical point of view, we introduce a different method for handling the boundary conditions when solving the geodesic boundary value problem on the space of unparametrized curves. Mathematically, an unparametrized curve corresponds to the orbit $c \circ \text{Diff}(M^1)$ but there is no numerically convenient way to discretize this group action. This problem has been approached in various ways, *e.g.* in previous work [5] the authors optimized simultaneously over both the geodesic path and the reparametrization of the target curve. In this paper, we continue to develop the idea inspired from previous works on curve and surface registration based on diffeomorphic models like [21, 25, 27]: to enforce the constraint that the endpoint of the geodesic path $c(1)$ and the given target curve c_1 belong to the same equivalence class of unparametrized curves, we use an auxiliary reparametrization-invariant distance function. The construction of those distance functions, unrelated to Sobolev metrics, follow the principles of geometric measure theory, which have so far been used as fidelity terms in combination with models like the large deformation diffeomorphic metric mapping (LDDMM) framework. In the present article, building on our previous conference publication [2], we examine a fairly general class of kernel metrics on immersed open and closed curves that are induced from the representation of curves as oriented varifolds. We examine the rigorous conditions to obtain distance functions on unparametrized immersed curves. Since this framework provides us with smooth proximity measures between unparametrized curves, it can be used to formulate the geodesic boundary value problem without having to explicitly estimate the reparametrizations.

⁴<https://www.github.com/h2metrics/h2metrics>

Our numerical implementation takes full advantage of the flexibility provided by the varifold-based distance. We formulate two algorithms: an exact matching algorithm using an augmented Lagrangian approach and an inexact matching algorithm based on optimizing a relaxed functional that incorporates the geodesic energy and the varifold-based distance. The latter algorithm is more flexible when the given data is noisy and similar approaches are used in most deformation-based shape analysis frameworks.

2. METRICS ON SPACES OF OPEN AND CLOSED CURVES

2.1. Shape spaces of curves

Let $d \geq 2$ be the dimension of the ambient space and M^1 the parameter space for a curve. For open curves we have $M^1 = I = [0, 2\pi]$ and for closed curves $M^1 = S^1 = \mathbb{R}/2\pi\mathbb{Z}$. In both cases M^1 is a compact, one-dimensional manifold. For a curve $c : M^1 \rightarrow \mathbb{R}^d$ we write $c'(\theta) = \frac{d}{d\theta}c(\theta)$ to denote its derivative.

Definition 2.1. Let M^1 be S^1 or I . The space of smooth, regular curves with values in \mathbb{R}^d is

$$\text{Imm}(M^1, \mathbb{R}^d) = \left\{ c \in C^\infty(M^1, \mathbb{R}^d) : c'(\theta) \neq 0 \quad \forall \theta \in M^1 \right\}.$$

The notation Imm stands for *immersions*. The space $\text{Imm}(M^1, \mathbb{R}^d)$ is an open subset of the Fréchet space $C^\infty(M^1, \mathbb{R}^d)$ and therefore itself a Fréchet manifold. Its tangent space $T_c \text{Imm}(M^1, \mathbb{R}^d)$ at any curve c is the vector space $C^\infty(M^1, \mathbb{R}^d)$ itself.

We will call curves in $\text{Imm}(M^1, \mathbb{R}^d)$ parametrized curves, because as maps from the parameter space M^1 to \mathbb{R}^d they carry with them a parametrization. We will later define the space of unparametrized curves in Definition 2.3.

Two curves that differ only by their parametrization represent the same geometric object. In the context of shape analysis it is therefore natural to consider them as equal, *i.e.*, we identify the curves c and $c \circ \varphi$, where φ is a reparametrization. As the reparametrization group we use the group of smooth diffeomorphisms of M^1 ,

$$\text{Diff}(M^1) = \left\{ \varphi \in C^\infty(M^1, M^1) : \varphi \text{ bij. and } \varphi^{-1} \in C^\infty(M^1, M^1) \right\},$$

which is an infinite-dimensional regular Fréchet Lie group [31]. For the two cases studied in this article these groups are

$$\begin{aligned} \text{Diff}(S^1) &= \left\{ \varphi \in C^\infty(S^1, S^1) : \varphi'(\theta) \neq 0 \quad \forall \theta \in S^1 \right\}, \\ \text{Diff}(I) &= \left\{ \varphi \in C^\infty(I, I) : \varphi'(\theta) \neq 0, \varphi(\{0, 2\pi\}) = \{0, 2\pi\} \right\}. \end{aligned}$$

To define the quotient space of unparametrized curves we need to restrict ourselves to *free immersions*, *i.e.*, those upon which the diffeomorphism group acts freely:

Definition 2.2. Let M^1 be S^1 or I . The space of free immersions with values in \mathbb{R}^d is

$$\text{Imm}_f(M^1, \mathbb{R}^d) = \left\{ c \in \text{Imm}(M^1, \mathbb{R}^d) : (c \circ \varphi = c \Rightarrow \varphi = \text{Id}_{M^1}) \right\}.$$

This restriction is only necessary for technical reasons to be able to define a manifold structure on the quotient space; in applications almost all curves are freely immersed, in particular the subset of free immersions is dense [18].

Definition 2.3. The space of unparametrized curves

$$B_{i,f}(M^1, \mathbb{R}^d) = \text{Imm}_f(M^1, \mathbb{R}^d) / \text{Diff}(M^1),$$

is the orbit space of the group action of $\text{Diff}(M^1)$ restricted to all free immersions.

This space is a Fréchet manifold although constructing charts is nontrivial in this case [18]. The space $B_{i,f}(M^1, \mathbb{R}^d)$ is sometimes referred to as the *pre-shape space*, while the *shape space* is obtained from the pre-shape space by additionally taking the quotient with respect to the group $S(d)$ of similarity transformations of \mathbb{R}^d or one of its subgroups. Here

$$S(d) = (\mathbb{R}_{>0} \times SO(d)) \ltimes \mathbb{R}^d,$$

where $\mathbb{R}_{>0}$ is the scaling group, $SO(d)$ is the rotation group and \mathbb{R}^d is the translation group. The composition of two transformations is given by

$$(r_1, A_1, w_1) \cdot (r_2, A_2, w_2) = (r_1 r_2, A_1 A_2, r_2^{-1} A_2^{-1} w_1 + w_2).$$

The group $S(d)$ acts on curves from the left *via*

$$\text{Imm}(M^1, \mathbb{R}^d) \times S(d) \rightarrow \text{Imm}(M^1, \mathbb{R}^d) \quad (c, (r, A, w)) \mapsto rA(c + w).$$

Note that elements of $S(d)$ are orientation-preserving. We do not include the reflection $c \mapsto -c$ in this group. This can be done, but is not relevant for the applications considered below.

Let H be a subgroup of $S(d)$. Common choices for H are the translation group \mathbb{R}^d , the group of Euclidean motions $SE(d) = SO(d) \ltimes \mathbb{R}^d$ and $S(d)$ itself. The shape space of unparametrized curves modulo similarities of type H is the quotient

$$\mathcal{S}_H(M^1, \mathbb{R}^d) = B_{i,f}(M^1, \mathbb{R}^d)/H = \text{Imm}_f(M^1, \mathbb{R}^d)/H \times \text{Diff}(M^1).$$

We will write simply $\mathcal{S}(M^1, \mathbb{R}^d)$ instead of $\mathcal{S}_H(M^1, \mathbb{R}^d)$, when the meaning of H is clear from the context.

2.2. Notation

We denote the Euclidean inner product on \mathbb{R}^d by $\langle \cdot, \cdot \rangle$. For any fixed immersed curve c , we denote differentiation and integration with respect to arc length by $D_s = \frac{1}{|c_\theta|} \partial_\theta$ and $ds = |c_\theta| d\theta$, respectively. The length of the curve is $\ell_c = \int_{M^1} ds = \int_{M^1} |c'| d\theta$. We will omit the subscript and write $\ell = \ell_c$ if the curve c is clear from the context. The unit length tangent vector to c is $v = D_s c = \frac{c'}{|c'|}$. We can decompose any vector field along the curve, $h \in T_c \text{Imm}(M^1, \mathbb{R}^d)$, into components tangential and normal to the curve, $h = \langle h, v \rangle v + (h - \langle h, v \rangle v)$ and we denote them by $h^\top = \langle h, v \rangle v$ and $h^\perp = h - h^\top$. In fact we will apply this decomposition to the derivative $D_s h$ and we write $D_s h^\top = (D_s h)^\top$ and $D_s h^\perp = (D_s h)^\perp$. Note that D_s does not commute with these projections and $D_s(h^\top) \neq (D_s h)^\top$ in general.

2.3. Higher order elastic metrics

Here, we introduce the class of Riemannian metrics that will be used in the remainder of the article. We also show that these metrics possess certain desirable completeness properties and these will serve as theoretical justification for our numerical framework.

Definition 2.4. A *second order elastic Sobolev metric with length-weighted coefficients* is a Riemannian metric on the space $\text{Imm}(M^1, \mathbb{R}^d)$ of parametrized curves of the form

$$G_c(h, k) = \int_{M^1} a_0(\ell) \langle h, k \rangle + a_1(\ell) \langle D_s h^\top, D_s k^\top \rangle + b_1(\ell) \langle D_s h^\perp, D_s k^\perp \rangle + a_2(\ell) \langle D_s^2 h, D_s^2 k \rangle ds,$$

where $a_0, a_1, b_1, a_2 \in C^\infty(\mathbb{R}_{>0}, \mathbb{R}_{\geq 0})$ are smooth positive functions of the curve length ℓ with $a_0(\ell) > 0, a_2(\ell) > 0$ and $h, k \in T_c \text{Imm}(S^1, \mathbb{R}^d)$ are tangent vectors at c .

In the remainder of this article, we will restrict our attention to two special subfamilies. First, the family of *elastic metrics with constant coefficients*,

$$G_c^1(h, k) = \int_{M^1} a_0 \langle h, k \rangle + a_1 \langle D_s h^\top, D_s k^\top \rangle + b_1 \langle D_s h^\perp, D_s k^\perp \rangle + a_2 \langle D_s^2 h, D_s^2 k \rangle \, ds, \quad (2.1)$$

where $a_0, a_1, b_1, a_2 \in \mathbb{R}_{\geq 0}$ and $a_0, a_2 > 0$; second, the family of *scale-invariant elastic metrics*,

$$G_c^2(h, k) = \int_{M^1} \frac{a_0}{\ell^3} \langle h, k \rangle + \frac{a_1}{\ell} \langle D_s h^\top, D_s k^\top \rangle + \frac{b_1}{\ell} \langle D_s h^\perp, D_s k^\perp \rangle + \ell a_2 \langle D_s^2 h, D_s^2 k \rangle \, ds, \quad (2.2)$$

where the coefficient functions are of the form $\ell \mapsto \lambda \ell^\rho$ and $a_0, a_1, b_1, a_2 \in \mathbb{R}_{\geq 0}$ with $a_0, a_2 > 0$. Note that the symbols D_s , ds , \perp and \top hide the nonlinear dependency of the metric on the base point c . In the following remark, we discuss the invariance properties of the metrics (2.1) and (2.2).

Remark 2.5. Because we use arc length operations in the definition of length-weighted elastic Sobolev metrics, Definition 2.4, the resulting metrics are invariant under the action of the diffeomorphism group $\text{Diff}(M^1)$. They are also invariant under the Euclidean motion group $SE(d)$, but only the family G^2 of scale-invariant metrics is also invariant under the scaling group. The invariance properties of these metrics will allow us later to define induced Riemannian metrics on the shape space of unparametrized curves.

First, however, we will study properties of length-weighted elastic metrics on the space of parametrized curves. Let G be such a metric. The Riemannian length of a path $c: [0, 1] \rightarrow \text{Imm}(M^1, \mathbb{R}^d)$ is

$$L(c) = \int_0^1 \sqrt{G_{c(t)}(c_t(t), c_t(t))} \, dt,$$

with $c_t = \partial_t c$ the time derivative of the path c . We denote by \mathcal{P} the space of all smooth paths,

$$\mathcal{P} = C^\infty([0, 1], \text{Imm}(M^1, \mathbb{R}^d)).$$

The geodesic distance induced by G between two given curves c_0, c_1 is defined as the infimum of the lengths of all paths connecting these two curves, *i.e.*,

$$\text{dist}^G(c_0, c_1) = \inf \left\{ L(c) : c \in \mathcal{P}, c(0) = c_0, c(1) = c_1 \right\}.$$

It is a general result in Riemannian geometry that the squared geodesic distance is also the infimum over all paths of the Riemannian energy,

$$E(c) = \int_0^1 G_{c(t)}(c_t(t), c_t(t)) \, dt. \quad (2.3)$$

Geodesics are critical points of the energy functional and the first order condition for critical points, $DE(c) = 0$ is the geodesic equation. For elastic metrics the geodesic equation is a partial differential equation for the function $c = c(t, \theta)$. Since we are working in infinite dimensions the existence of geodesics is a nontrivial question. For elastic Sobolev metrics we have the following existence results for geodesics, which are based on the results in [13, 15, 36]. They will serve as the theoretical foundation of the proposed numerical framework.

Theorem 2.6. *Let G be a second order elastic metric, either scale-invariant or with constant coefficients on the space of closed curves. Then*

1. *The Riemannian manifold $(\text{Imm}(S^1, \mathbb{R}^d), G)$ is geodesically complete, i.e., given any initial conditions $(c_0, u_0) \in T \text{Imm}(S^1, \mathbb{R}^d)$ the solution of the geodesic equation for the metric G with initial values (c_0, u_0) exists for all times.*
2. *The metric completion of the space $\text{Imm}(S^1, \mathbb{R}^d)$ equipped with the geodesic distance dist^G is the space $\mathcal{I}^2(S^1, \mathbb{R}^d)$ of immersions of Sobolev class H^2 ,*

$$\mathcal{I}^2(S^1, \mathbb{R}^d) = \left\{ c \in H^2(S^1, \mathbb{R}^d) : c'(\theta) \neq 0 \quad \forall \theta \in S^1 \right\}.$$

Furthermore, any two curves in the same connected component of the space $\mathcal{I}^2(S^1, \mathbb{R}^d)$ can be joined by a minimizing geodesic.

Remark 2.7 (Incompleteness for open curves). In the following we will present a counterexample for the above completeness result for the metric G^1 in the case of open curves. Therefore, we consider the path $c(t, \theta) = ((1-t)\theta, 0)$ for $t \in [0, 1]$. Note that this path leaves the space of immersions for $t = 1$, since $c(1, \theta) = (0, 0)$. We have

$$c_t = (-\theta, 0), \quad c_\theta = (1-t, 0), \quad \langle D_s c_t, v \rangle = \frac{-1}{1-t}, \quad (D_s c_t)^\perp = (0, 0) \quad D_s^2 c_t = (0, 0).$$

Using this we calculate the G^1 -length of c :

$$\begin{aligned} L(c) &= \int_0^1 \sqrt{G_{c(t)}^1(c_t(t), c_t(t))} dt \\ &= \int_0^1 \sqrt{\int_0^{2\pi} a_0 \theta^2 + \frac{a_1}{(1-t)^2} (1-t) d\theta} dt \\ &= \int_0^1 \sqrt{\frac{a_0(2\pi)^3(1-t)}{3} + \frac{a_1 2\pi}{(1-t)}} dt \\ &\leq C \int_0^1 \sqrt{1 + \frac{1}{1-t}} dt \\ &= C(\sqrt{2} + \cosh^{-1}(\sqrt{2})) < \infty, \end{aligned}$$

where $C > 0$ is a constant. Thus, we have found a path that leaves the space of immersions with finite G^1 -length. This shows that the space $(\mathcal{I}^2([0, 2\pi], \mathbb{R}^d), \text{dist}^{G^1})$ is metrically incomplete.

We conjecture that for the scale invariant metric G^2 the completeness results would also hold on the space of open curves if a_0, a_1 and a_2 are non-zero. Note that the above path c has indeed infinite length with respect to the metric G^2 .

Proof of Theorem 2.6. Denote by G^1 an elastic metric with constant coefficients and by G^2 a scale-invariant elastic metric of the form (2.1) and (2.2). We first observe that both G^1 and G^2 extend to smooth Riemannian metrics on the Hilbert manifold $\mathcal{I}^2(S^1, \mathbb{R}^d)$. This follows directly from Sobolev embedding and multiplication theorems.

We will next show that $\mathcal{I}^2(S^1, \mathbb{R}^d)$ equipped with these metrics is metrically complete. For closed curves, metric completeness of second order Sobolev metrics on the space $\mathcal{I}^2(S^1, \mathbb{R}^d)$ has been shown for the metric

$$\bar{G}_c^1(h, k) = \int_{S^1} \langle h, k \rangle + \langle D_s h, D_s k \rangle + \langle D_s^2 h, D_s^2 k \rangle \, ds,$$

in [10, 13, 15] and for the metric

$$\bar{G}_c^2(h, k) = \int_{S^1} \frac{1}{\ell^3} \langle h, k \rangle + \frac{1}{\ell} \langle D_s h, D_s k \rangle \, ds + \ell \langle D_s^2 h, D_s^2 k \rangle \, ds,$$

in [16]. Because we can find uniform constants C_1, C_2 such that

$$C_1 \bar{G}_c^1(h, h) \leq G_c^1(h, h) \leq C_2 \bar{G}_c^1(h, h) \quad \text{and} \quad C_1 \bar{G}_c^2(h, h) \leq G_c^2(h, h) \leq C_2 \bar{G}_c^2(h, h),$$

holds for all $(c, h) \in T\mathcal{I}^2(S^1, \mathbb{R}^d)$, where \bar{G}_c^1, \bar{G}_c^2 are defined as above, it follows that the metrics G^1 and G^2 are also metrically complete on $\mathcal{I}^2(S^1, \mathbb{R}^d)$.

Metric completeness of G^1 and G^2 on $\mathcal{I}^2(S^1, \mathbb{R}^d)$ implies geodesic completeness. To see that the space $\text{Imm}(S^1, \mathbb{R}^d)$ of smooth immersions is also geodesically complete, we use that the geodesic equation preserves smoothness of the initial conditions: if the initial curve and velocity field are C^∞ -smooth, then so is every curve along the geodesic, see [14, 15, 23].

To show the existence of minimizing geodesics we use ([13], Rem. 5.4). There the existence of minimizing geodesics for metrics on $\mathcal{I}^2(S^1, \mathbb{R}^d)$ is proven provided that:

- they are uniformly bounded and uniformly coercive with respect to the background $\|\cdot\|_{H^n(d\theta)}$ -norm on metric balls in the geodesic distance;
- they are of the form

$$G_c(h, h) = \sum_{k=1}^N \|A_k(c)h\|_{F_k}^2,$$

with some Hilbert spaces F_k and smooth maps $A_k : \mathcal{I}^2 \rightarrow L(H^2, F_k)$, where the maps A_k are required to have the property:

$$c^j \rightarrow c \text{ weakly in } H_t^1 \mathcal{I}_\theta^2 \quad \Rightarrow \quad A_k(c^j) \dot{c}^j \rightarrow A_k(c) \dot{c} \text{ weakly in } L^2(I, F_k).$$

In our case $H_t^1 \mathcal{I}_\theta^2 = H^1([0, 1], \mathcal{I}^2(S^1, \mathbb{R}^d))$. In our case $N = 4$, $F_k = L^2(S^1, \mathbb{R}^d)$ and $A_1(c)h = \sqrt{a_0}h$, $A_2(c)h = \sqrt{a_1}D_s h^\top$, $A_3(c)h = \sqrt{b_1}D_s h^\perp$ and $A_4(c)h = \sqrt{a_2}D_s^2 h$. The necessary convergence properties follow from [13, Lemma 5.9]. \square

For applications in matching the central task is to obtain stable and fast algorithms to calculate the induced geodesic distance. Our framework is based on discretizing the Riemannian energy and minimizing it over all (discrete) paths. For some first order metrics there exist transformations that can significantly speed up these calculations, because they yield explicit formulas for the geodesic distance on open parametrized curves. In related work on open planar curves Kurtek and Needham [32] will follow this approach to obtain fast numerical algorithms for first order metrics. The aim of the present article is to develop a numerical framework for a wider class of metrics on open and closed curves, that should allow one to model a variety of different matching behaviors. Furthermore, we plan to further enhance our framework in future work in order to be able to deal with surfaces in addition to curves. For these reasons we do not take advantage of these explicit formulas.

2.4. Elastic metrics on shape spaces of curves

We will now use the metrics defined in Section 2.3 to induce Riemannian metrics on shape spaces of unparametrized curves, as defined in Section 2.1. In this section, we fix a subgroup H of the group $SE(d)$ and let $\mathcal{S}_f(M^1, \mathbb{R}^d) = (\text{Imm}_f(M^1, \mathbb{R}^d) / \text{Diff}(M^1)) / H$. Using invariance properties of the metrics, we obtain the following result concerning the induced metrics on the quotient space:

Theorem 2.8. *Let G be an elastic metric of type (2.1) or (2.2). Then G induces a Riemannian metric on the quotient space $\mathcal{S}_f(M^1, \mathbb{R}^d)$ such that the projection*

$$\pi : \text{Imm}_f(M^1, \mathbb{R}^d) \rightarrow \mathcal{S}_f(M^1, \mathbb{R}^d)$$

is a Riemannian submersion. For the metric G^2 the result holds also for quotient spaces with respect to scale.

Remark 2.9. Note, that the invariance of the metric is only a necessary condition for a metric on the space of parametrized curves to induce a metric on unparametrized curves. To guarantee the existence of the induced quotient metric one has to verify the existence of the horizontal bundle for each specific metric. This has been achieved for all the metrics studied in this article [36].

We have to restrict ourselves to free immersions to obtain a smooth structure on the quotient $\mathcal{S}_f(M^1, \mathbb{R}^d)$. However, in numerical calculations we will work with the full space $\text{Imm}(M^1, \mathbb{R}^d)$ and the quotient $\mathcal{S}(M^1, \mathbb{R}^d) = \text{Imm}(M^1, \mathbb{R}^d) / \text{Diff}(M^1) / H$. This quotient space has singularities at non-free immersions, but the singularities are very mild [18].

The geodesic distance on $\text{Imm}(M^1, \mathbb{R}^d)$ gives rise to a distance on the quotient space $\mathcal{S}(M^1, \mathbb{R}^d)$, which coincides – for shapes sufficiently close to each other – with the geodesic distance of the induced Riemannian metric on $\mathcal{S}_f(M^1, \mathbb{R}^d)$. Here, we use the fact that $\mathcal{S}_f(M^1, \mathbb{R}^d)$ is an open dense subset of $\mathcal{S}(M^1, \mathbb{R}^d)$ and thus the geodesic distances coincide at least as long as a minimizing deformation does not encounter points in $\mathcal{S}(M^1, \mathbb{R}^d)$ that do not belong to $\mathcal{S}_f(M^1, \mathbb{R}^d)$.

This distance can be then calculated using paths in $\text{Imm}(M^1, \mathbb{R}^d)$ connecting c_0 to the orbit $c_1 \circ \text{Diff}(M^1) \circ H$, i.e., for $\pi(c_0), \pi(c_1) \in \text{Imm}(M^1, \mathbb{R}^d) / \text{Diff}(M^1) / H$ we have,

$$\text{dist}(\pi(c_0), \pi(c_1)) = \inf \left\{ L(c) : c \in \mathcal{P}, c(0) = c_0, c(1) \in c_1 \circ \text{Diff}(M^1) \circ H \right\}. \quad (2.4)$$

We have the following completeness result for the quotient space of closed curves modulo reparametrizations:

Theorem 2.10. *Let G be an elastic metric of type (2.1) or (2.2) with $a_0, a_2 > 0$ on the shape space $\mathcal{S}(S^1, \mathbb{R}^d)$. Then the metric completion of $\mathcal{S}(S^1, \mathbb{R}^d)$ equipped with the quotient distance is the space of all Sobolev shapes of class H^2 ,*

$$\mathcal{I}^2(S^1, \mathbb{R}^d) / \mathcal{D}^2(S^1) / H.$$

Here, \mathcal{D}^2 denotes the H^2 -Sobolev completion of the diffeomorphism group $\text{Diff}(S^1)$. Furthermore, the metric completion is a length space and any two shapes in the same connected component can be joined by a minimizing geodesic.

The proof of this theorem is verbatim the same as in Section 6 from [13].

3. ORIENTED VARIFOLD METRICS

We will derive an efficient relaxation term for the matching constraint using distances on the space of curves that originate from geometric measure theory and which have been applied extensively in shape analysis and computational anatomy. Heuristically, the philosophy is to induce a metric on the space of unparametrized curves using their representations as generalized distributions.

Several models for such distributions and their associated metrics have been proposed: measures [26], currents [12, 22, 27] or varifolds [21]. The most recent work [29] introduces the general representation of a curve as an oriented varifold, which combines the different approaches into a single framework. Varifolds can be used, in principle, not just to define distances between curves but between embedded submanifolds of any dimension and codimension although numerical implementations exist only for curves and surfaces in \mathbb{R}^2 and \mathbb{R}^3 . In the present work, we will focus only on smooth, open or closed curves. However, we have to be careful when applying the varifold framework because, technically, we are dealing with immersed and not embedded curves. In what follows, we give a brief reminder of the framework of oriented varifolds as it applies to curves while addressing the distinction between immersed and embedded curves in more detail.

3.1. Representation of curves as oriented varifolds

Let $C_0(\mathbb{R}^d \times S^{d-1})$ be the space of continuous functions vanishing at infinity. An oriented varifold is intuitively a joint distribution of point positions and directions. Mathematically, we define them as follows:

Definition 3.1. An oriented varifold is an element of the distribution space $C_0(\mathbb{R}^d \times S^{d-1})^*$, i.e., a signed measure on the product $\mathbb{R}^d \times S^{d-1}$.

The analogue of Dirac distributions in the context of oriented varifolds are the distributions $\delta_{(x,u)}$ with $(x,u) \in \mathbb{R}^d \times S^{d-1}$, defined by $\delta_{(x,u)}(\omega) = \omega(x,u)$ for all test functions $\omega \in C_0(\mathbb{R}^d \times S^{d-1})$.

Next we define the natural representation of curves as oriented varifolds.

Definition 3.2. The varifold application $\mu : c \mapsto \mu_c$ associates to any immersion $c \in \text{Imm}(M^1, \mathbb{R}^d)$ the oriented varifold μ_c defined, for any $\omega \in C_0(\mathbb{R}^d \times S^{d-1})$, by

$$\mu_c(\omega) = \int_{M^1} \omega \left(c(\theta), \frac{c'(\theta)}{|c'(\theta)|} \right) ds. \quad (3.1)$$

Note that (3.1) writes informally as $\mu_c = \int_{M^1} \delta_{(c(\theta), u(\theta))} ds$ where $u(\theta) = \frac{c'(\theta)}{|c'(\theta)|}$ is the unit tangent vector at θ ; in other words μ_c can be interpreted as the weighed combination of Diracs at the point positions of the curve $c(\theta)$ with attached vectors $u(\theta)$ and infinitesimal weights given by the arclength $ds = |c'(\theta)| d\theta$.

A key property is that μ_c is actually *independent* of the parametrization. Indeed, a straightforward change of variables in (3.1) shows that for any positive reparametrization φ in $\text{Diff}^+(M^1) = \{\varphi \in \text{Diff}(M^1) \mid \varphi'(\theta) > 0\}$, one has $\mu_{c \circ \varphi} = \mu_c$. It follows that the map $c \mapsto \mu_c$ projects to a well-defined map from the quotient space $B_i^+(M^1, \mathbb{R}^d) = \text{Imm}(M^1, \mathbb{R}^d)/\text{Diff}^+(M^1)$ of *oriented* unparametrized immersed curves into the space of varifolds:

$$\begin{array}{ccc} \text{Imm}(M^1, \mathbb{R}^d) & \xrightarrow{\pi^+} & B_i^+(M^1, \mathbb{R}^d) \\ & \searrow \mu & \downarrow \\ & & C_0(\mathbb{R}^d \times S^{d-1})^* \end{array}$$

Note that neither the original map μ_c nor the resulting quotient map are surjective as there are many varifolds that are not curves (e.g. a single Dirac). Whether the quotient map is injective is a question with a more nuanced answer and we will discuss it more thoroughly in the following.

Remark 3.3. The varifold representation remains sensitive to the orientation of a curve because, if \check{c} is the curve c with the opposite orientation then we have, in general, $\mu_{\check{c}} \neq \mu_c$. In all of the applications considered in this paper curves can be naturally and consistently oriented and hence orientation represents a relevant piece of information that can be exploited for curve matching (cf. the discussion in [29]). However, as we will explain in the next subsection, we can also consider the quotient spaces of unoriented curves and varifolds by constraining

test functions to be symmetric with respect to the second variable u ; this corresponds to the framework of unoriented varifolds of [21].

3.2. Kernel metrics

The varifold application embeds unparametrized curves in a common space of distributions. This suggests that we can construct distances on the space of curves by restricting distances or pseudo-distances defined on the space of varifolds. The natural choice would be the distance induced by the norm on $C_0(\mathbb{R}^d, S^{d-1})^*$ that is dual to the supremum norm on $C_0(\mathbb{R}^d \times S^{d-1})$. However, this yields a fundamentally nonsmooth distance for curves, as it essentially measures the exact overlap between two curves. Thus, to obtain more reasonable distances, one needs to restrict oneself to more regular spaces of test functions equipped with stronger norms.

A practical approach is to consider a Hilbert space \mathcal{H} of test functions, continuously embedded in $C_0(\mathbb{R}^d, S^{d-1})$. In this case, \mathcal{H} is a *reproducing kernel Hilbert space (RKHS)* and it is generated by a positive definite kernel k on the product space $\mathbb{R}^d \times S^{d-1}$. Following [29], we require the kernel k to have some additional structure, namely, k has to be a product of a radial kernel ρ on \mathbb{R}^d and a zonal kernel γ on S^{d-1} , *i.e.*,

$$k(x, u, y, v) \doteq \rho(|x - y|^2) \gamma(u \cdot v) \quad (3.2)$$

for all (x, u) and (y, v) in $\mathbb{R}^d \times S^{d-1}$. Here, ρ defines a continuous, positive radial basis function with $\rho(t) \rightarrow 0$ as $|t| \rightarrow \infty$ and $\gamma : [-1, 1] \rightarrow \mathbb{R}$ defines a continuous zonal function on the sphere. The general theory of reproducing kernels [1] states that under these assumptions on the kernel k , the Hilbert space \mathcal{H} is uniquely determined by k . We denote by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ the inner product on \mathcal{H} . The Riesz duality map then induces an inner product on \mathcal{H}^* . Moreover, using the dual map $C_0(\mathbb{R}^d \times S^{d-1})^* \rightarrow \mathcal{H}^*$, it then also induces a pseudo-distance on the space of varifolds (which is actually a distance if \mathcal{H} is additionally assumed to be dense in $C_0(\mathbb{R}^d \times S^{d-1})$). Having chosen a kernel of the form (3.2), we will use the generic notation $\langle \cdot, \cdot \rangle_{\text{Var}}$ for the associated inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}^*}$ on the space of varifolds.

For the purpose of this paper, we are in fact interested in the metric that is induced on the space of curves by the varifold map $c \mapsto \mu_c$. This metric is given by $d^{\text{Var}}(c_1, c_2) = \|\mu_{c_1} - \mu_{c_2}\|_{\text{Var}} = \langle \mu_{c_1} - \mu_{c_2}, \mu_{c_1} - \mu_{c_2} \rangle_{\text{Var}}^{1/2}$. The reproducing kernel property implies – *cf.* [29] for details – that for any two curves c_1, c_2 we have

$$\langle \mu_{c_1}, \mu_{c_2} \rangle_{\text{Var}} = \iint_{M^1 \times M^1} \rho(|c_1(\theta_1) - c_2(\theta_2)|^2) \gamma\left(\frac{c'_1(\theta_1)}{|c'_1(\theta_1)|} \cdot \frac{c'_2(\theta_2)}{|c'_2(\theta_2)|}\right) ds_1 ds_2, \quad (3.3)$$

leading to a similar closed-form expression for $d^{\text{Var}}(c_1, c_2)^2$. Equation (3.3) shows that d^{Var} can be interpreted as a localized comparison between the relative positions of points and tangent lines of the two curves, quantified by the choice of kernel functions ρ and γ . We will see later how to efficiently evaluate these expressions numerically for discrete curves.

In general d^{Var} only defines a pseudo-distance. In order to be able to separate any two curves, one needs to ensure that the space of test functions \mathcal{H} is large enough. A sufficient condition for this is given by the following theorem which is a particular case of Proposition 4 in [29]:

Theorem 3.4. *Assume that ρ and γ are C^1 -functions, ρ is C_0 -universal and $\gamma(1) > 0$. Then, if $d^{\text{Var}}(c_1, c_2) = 0$, we have $\text{Im}(c_1) = \text{Im}(c_2)$ *i.e.*, the images of c_1 and c_2 in \mathbb{R}^d coincide.*

The kernel ρ is said to be *C_0 -universal* if its associated RKHS is dense in $C_0(\mathbb{R}^d, \mathbb{R})$. This is the case for Gaussian, Cauchy and Wendland kernels for example, we refer to [17] or [44] for details on the construction and characterization of such kernels.

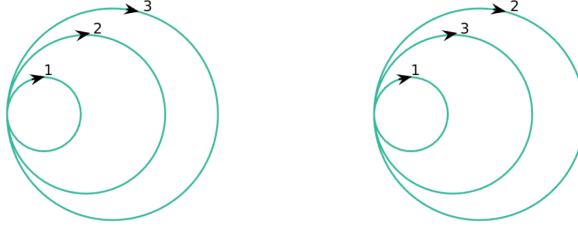


FIGURE 1. Example of two distinct oriented immersed curves that are however equal in the space of oriented varifolds. The numbers reflect the order of crossing of each piece.

Note that while we have $\text{Im}(c_1) = \text{Im}(c_2)$ in the result of Theorem 3.4, it is not necessarily the case that the orientations of c_1 and c_2 coincide. This is in particular not true when γ defines an orientation-invariant kernel (cf. discussion below). This can be enforced, however, under the following conditions:

Corollary 3.5. *In addition to the assumptions of Theorem 3.4, if the function γ is such that $\gamma(-t) \neq \gamma(t)$ for all $t \neq 0$, then d^{Var} defines a distance on the space of oriented, unparametrized, embedded curves.*

The proof follows from very similar arguments than the one of Theorem 3.4 in [29] that we do not repeat for concision. Note that the last statement is equivalent to saying that the varifold application μ into \mathcal{H}^* is injective if restricted to $\text{Emb}(M^1, \mathbb{R}^d)/\text{Diff}^+(M^1)$.

Unfortunately, it is easy to see that this result cannot be extended to the larger space of free immersions. Indeed, the varifold representation μ_c of a curve c takes into account only the image of an immersed curve and its orientation. The example in Figure 1 (also mentioned in [12] for the specific case of currents) shows two distinct elements of $B_i^+(S^1, \mathbb{R}^d) = \text{Imm}(S^1, \mathbb{R}^d)/\text{Diff}^+(S^1)$, both projections of free immersions, that are nevertheless equal as oriented varifolds. Yet, the previous result on embedded curves can be generalized to the space of immersed curves with a finite number of transverse self-intersections. Note that any such immersion is already a free immersion.

Theorem 3.6. *With the assumptions of Corollary 3.5, if c_1 and c_2 are two immersions with a finite number of transverse self-intersections then $d^{\text{Var}}(c_1, c_2) = 0$ if and only if the two curves coincide in $B_i^+(M^1, \mathbb{R}^d)$.*

Proof. Let c_1 and c_2 be two such immersions with $d^{\text{Var}}(c_1, c_2) = 0$. We need to show that there exists $\varphi \in \text{Diff}(M^1)$ such that $c_2 = c_1 \circ \varphi$.

Using Theorem 3.4 we already know that $\text{Im}(c_1) = \text{Im}(c_2)$. Let us denote by $(p_1, \dots, p_m) \in (\mathbb{R}^d)^m$ the self-intersection points of c_1 , which are the same for c_2 as the two curves have the same image. For any $x \in \text{Im}(c_1) \setminus \{p_1, \dots, p_m\}$, there exist unique preimages $c_1^{-1}(x)$ and $c_2^{-1}(x)$ in M^1 . Let us denote by Θ_1^s and Θ_2^s the preimages of the self-intersection points under c_1 and c_2 respectively. We set $\varphi(\theta) = c_1^{-1} \circ c_2(\theta)$ for $\theta \in M^1 \setminus \Theta_2^s$.

First, note that φ is smooth on $M^1 \setminus \Theta_2^s$ and that $c_2(\theta) = c_1 \circ \varphi(\theta)$ for all $\theta \in M^1 \setminus \Theta_2^s$. In addition, using Corollary 3.5, we also have that the orientation coincide on each connected component of $c_1(M^1 \setminus \Theta_2^s)$ and therefore $\varphi'(\theta) > 0$ on $M^1 \setminus \Theta_2^s$.

Now, let $i \in \{1, \dots, m\}$ and $c_2^{-1}(p_i) = \{\theta_1^i, \dots, \theta_{m_i}^i\}$. Since all self-intersections are transverse, we have that the $c_2'(\theta_k^i)$ are all distinct from one another. Similarly, as the two curves have the same image, we can write $c_1^{-1}(p_i) = \{\tilde{\theta}_1^i, \dots, \tilde{\theta}_{m_i}^i\}$ and with the adequate ordering we can also assume that the two vectors $c_1'(\tilde{\theta}_k^i)$ and $c_2'(\theta_k^i)$ are collinear. Then, setting for all i and k , $\varphi(\theta_k^i) = \tilde{\theta}_k^i$, we obtain a bijection $\varphi : M^1 \rightarrow M^1$ such that $c_2 = c_1 \circ \varphi$. Moreover, the above construction makes φ a smooth function that satisfies $c_2'(\theta) = c_1'(\varphi(\theta))\varphi'(\theta)$ for all $\theta \in M^1$. \square

In certain situations, it may be more relevant to work with unoriented immersed curves *i.e.* with the space $B_i(M^1, \mathbb{R}^d) = \text{Imm}(M^1, \mathbb{R}^d)/\text{Diff}(M^1)$. An equivalent result holds by instead considering orientation-invariant kernels which are such that $\gamma(-t) = \gamma(t)$ for all $t \in [-1, 1]$. Indeed one can easily see from (3.3) that the resulting

metric d^{Var} is also invariant to orientation changes in either curve; this corresponds to the subclass of unoriented varifold metrics, *cf.* [21, 29]. Theorem 3.6 can be then replaced by:

Corollary 3.7. *With the assumptions of Theorem 3.4 and if γ is an even function, two immersed curves with finite numbers of transverse self-intersections are equal in $B_i(M^1, \mathbb{R}^d) = \text{Imm}(M^1, \mathbb{R}^d) / \text{Diff}(M^1)$ if and only if we have $d^{\text{Var}}(c_1, c_2) = 0$.*

In summary, although the varifold metrics introduced here may not always distinguish two given immersed curves, we see from Theorem 3.6 and Corollary 3.7 that this will only occur in pathological situations such as shown in Figure 1. We will typically ignore such cases in the practical curve matching applications of this paper.

3.3. Varifold distance as a constraint

The invariance of varifold-induced distances under reparametrizations makes them a natural tool for enforcing the exact matching constraint in the geodesic boundary value problem for elastic metrics. Indeed, the geodesic distance $\text{dist}(\pi(c_0), \pi(c_1))^2$ can be computed in the following way,

$$\text{dist}(\pi(c_0), \pi(c_1))^2 = \inf \left\{ E(c) : c \in \mathcal{P}, c(0) = c_0, d^{\text{Var}}(c(1), c_1)^2 = 0 \right\}, \quad (3.4)$$

where $E(c)$ is the Riemannian energy of the path c and \mathcal{P} is the space of all smooth paths in $\text{Imm}(M^1, \mathbb{R}^d)$. The squared varifold distance, which can be calculated explicitly *via* (3.3), is used as a smooth constraint enforcing the endpoint condition $\pi(c(1)) = \pi(c_1)$. In contrast with the direct approach of calculating $\text{dist}(\pi(c_0), \pi(c_1))^2$ *via* (2.4), the formulation (3.4) does not require optimization over reparametrizations. Never the less the optimal point correspondences can be inferred from our method. The equivalence between the two formulations is rigorous provided the curves c_0, c_1 and the kernel k satisfy the assumptions of Corollary 3.7. In that case, we have $d^{\text{Var}}(c(1), c_1)^2 = 0 \Leftrightarrow c(1) = c_1 \circ \varphi$ for some $\varphi \in \text{Diff}(M^1)$.

Remark 3.8. Note that this corresponds to the problem of matching unoriented, unparametrized, immersed curves, *i.e.*, elements of $B_i(M^1, \mathbb{R}^d)$. In certain other situations, one could assume that curves have been consistently oriented from the start and wish to solve the matching problem for oriented curves in $B_i^+(M^1, \mathbb{R}^d)$ instead. In that case, it is not difficult to see that we can also reformulate the problem as (3.4) by choosing an orientation-sensitive metric for d^{Var} .

3.4. Invariance to similarities

As mentioned at the end of Section 2.1, it is often important to compare curves modulo the positive similarity group $S(d)$ and therefore quotient out these transformations in the estimation of distance and geodesic. We first focus on the particular subgroup $SE(d)$ of Euclidean motions. In that case, we have seen that both families of elastic metrics G_c^1 in (2.1) and G_c^2 in (2.2) are invariant to the action of $SE(d)$. Thanks to the particular form of k in (3.2), it turns out that this is also the case of the kernel-based distances d^{Var} , *i.e.*, we always have $d^{\text{Var}}(A.(c_1 + w), A.(c_2 + w)) = d^{\text{Var}}(c_1, c_2)$ for any two curves c_1, c_2 and rigid motion $(A, w) \in SE(d)$. Consequently, the invariant matching problem:

$$\inf \left\{ E(c) : c \in \mathcal{P}, (A, w) \in SE(d), c(0) = c_0, c(1) = A.(c_1 + w) \right\}$$

becomes once again equivalent to

$$\inf \left\{ E(c) : c \in \mathcal{P}, (A, w) \in SE(d), c(0) = c_0, d^{\text{Var}}(c(1), A.(c_1 + w))^2 = 0 \right\},$$

which we can then tackle like previously in either the relaxed or augmented Lagrangian formulation, jointly over the path c and the finite-dimensional variables (A, w) .

The case of scale-invariance is however more involved in the present setting. While the second family of elastic metrics G_c^2 is invariant to rescaling, this is not true for the oriented varifold metrics of Section 3.2. In fact, it is quite easy to see that no metric originating from a kernel of the form of (3.2) is scale-invariant as this would impose that $\rho(\lambda^2 t) = \rho(t)/\lambda^2$ for all λ and t and thus lead to a singularity at 0 for the function ρ . In most applications [20, 21, 29], it is rather customary to specify kernels with an intrinsic notion of scale by setting for instance the kernel defined by ρ to be a Gaussian $\rho(|x - y|^2) = e^{-\frac{|x-y|^2}{\sigma^2}}$ or a sum of Gaussian for multiscale applications. In the context of this work, we point out the lack of invariance of d^{Var} to rescaling will not constitute an issue as the varifold metric is only used as a surrogate for the matching constraint of the two immersed curves, and thus only invariance with respect to reparametrizations is necessary.

4. IMPLEMENTATION

In this section, we will describe how to discretize and solve the constrained optimization problem (3.4) using both an inexact one-shot method and a iterative augmented Lagrangian scheme which enforces a better constraint satisfaction. Our code is available on GitHub.⁵

4.1. A B-spline discretization

In order to evaluate the energy functional (2.3) and the constraint (3.3) we discretize paths of curves using tensor product B-splines on knot sequences of orders n_t in time and n_θ in space (typically we choose $n_t = 2$ and $n_\theta = 3$). This produces $N_t \times N_\theta$ basis splines, with N_t and N_θ being the number of control points in each variable respectively (typical values we shall take in the experimental section are $N_\theta = 100$ and $N_t = 10$), and we can write

$$c(t, \theta) = \sum_{i=1}^{N_t} \sum_{j=1}^{N_\theta} c_{i,j} B_i(t) C_j(\theta). \quad (4.1)$$

Here, $B_i(t)$ are B-splines defined by an equidistant simple knot sequence on $[0, 1]$ with full multiplicity at the boundary knots, and $C_j(\theta)$ are defined by an equidistant simple knot sequence on $[0, 2\pi]$ with periodic boundary conditions or full multiplicity at the boundary for closed or open curves, respectively; for details see Section 3 of [5]. The full multiplicity of the boundary knots in t implies

$$c(0, \theta) = \sum_{j=1}^{N_\theta} c_{1,j} C_j(\theta), \quad c(1, \theta) = \sum_{j=1}^{N_\theta} c_{N_t,j} C_j(\theta).$$

Thus, the initial curve $c(0)$ is given by the control points $c_{1,j}$ only, which we can utilize later for the constraint satisfaction. In terms of the standard differential operator ∂_θ (as opposed to arc-length differentiation D_s) the Riemannian metrics (2.1) and (2.2) read as

$$\begin{aligned} G_c^1(h, k) = & \int_0^{2\pi} a_0 |c'| \langle h, k \rangle + \frac{a_1}{|c'|} \langle h'^\top, k'^\top \rangle + \frac{b_1}{|c'|} \langle h'^\perp, k'^\perp \rangle + \frac{a_2}{|c'|^7} \langle c', c'' \rangle^2 \langle h', k' \rangle \\ & - \frac{a_2}{|c'|^5} \langle c', c'' \rangle \left(\langle h', k'' \rangle + \langle h'', k' \rangle \right) + \frac{a_2}{|c'|^3} \langle h'', k'' \rangle \, d\theta. \end{aligned}$$

⁵<https://www.github.com/h2metrics/h2metrics>

$$G_c^2(h, k) = \int_0^{2\pi} \frac{a_0}{\ell^3} |c'| \langle h, k \rangle + \frac{a_1}{\ell |c'|} \langle h'^\top, k'^\top \rangle + \frac{b_1}{\ell |c'|} \langle h'^\perp, k'^\perp \rangle + \frac{\ell a_2}{|c'|^7} \langle c', c'' \rangle^2 \langle h', k' \rangle \\ - \frac{\ell a_2}{|c'|^5} \langle c', c'' \rangle \left(\langle h', k'' \rangle + \langle h'', k' \rangle \right) + \frac{\ell a_2}{|c'|^3} \langle h'', k'' \rangle d\theta.$$

Plugging these expressions into (2.3) gives an explicit expression, which we leave out, for the energy of a given path. For a B-spline path, we approximate the integrals in the energy functional (2.3) and varifold distance (3.3) using Gaussian quadrature with quadrature sites placed between knots where the curves are smooth. This yields a fast and robust way to evaluate the energy of paths. The same is true for the evaluation of the derivatives found in Appendix A.

4.2. The optimization procedures

We will now describe two different methods to approximately factor out the action of the diffeomorphism group. At the end we will comment on the action of Euclidean motions and scalings.

In Section 3.3, we showed that in order to factor out the diffeomorphism group we have to solve an optimization problem under the constraint that the end point of the curve satisfies $d^{\text{Var}}(c(1), c_1)^2 = 0$; this corresponds to an *exact* matching of the end point and the target curve. Inspired by the paradigms of other methods like LDDMM, as a simple method we consider an *inexact* matching problem where we only desire that the constraint violation of the end point is small instead of requiring it to be exactly zero. To this end a fixed large value of λ is chosen and the following relaxed quadratic penalty functional is considered

$$\inf \left\{ E(c) + \lambda d^{\text{Var}}(c(1), c_1)^2 : c \in \mathcal{P}, c(0) = c_0 \right\}. \quad (4.2)$$

Here, $\lambda > 0$ is a balance parameter between the elastic energy and the varifold fidelity term. This is particularly well-suited to noisy situations in which exact matching might lead to irrelevant solutions. Note that exact matching is still theoretically recovered in the limit $\lambda \rightarrow +\infty$. To solve the unconstrained optimization problem we use the HANSO library [42], which utilizes an L-BFGS method. This approach does not yield a geodesic with the correct endpoint but with an appropriate choice of λ the varifold distance term is small in practice. In [2] we employed this method, but the problem seemed quite sensitive to the choice of λ : too small and a bad matching is achieved; too big and the optimization algorithm has difficulty finding a solution.

In order to reduce the sensitivity of the solution to the choice of the weight parameter λ and to possibly solve the exact matching problem, we also propose an augmented Lagrangian scheme. In practice it will not be feasible for a B-spline path to satisfy $d^{\text{Var}}(c(1), c_1) = 0$ exactly, hence we would rather relax the constraint to an inequality

$$d^{\text{Var}}(c(1), c_1) \leq \varepsilon,$$

for some small chosen constraint error tolerance $\varepsilon > 0$. In order to solve this inequality constrained minimization, we use a simple adaptation of the augmented Lagrangian scheme, see [41]. We introduce the augmented Lagrangian functional

$$\mathcal{L}(c, \lambda, \mu) = E(c) - \lambda d^{\text{Var}}(c(1), c_1)^2 + \frac{\mu}{2} d^{\text{Var}}(c(1), c_1)^4; \quad (4.3)$$

here λ plays the role of the (real-valued) Lagrange multiplier associated to the constraint $d^{\text{Var}}(c(1), c_1)^2 = 0$. Notice that if $\mu = 0$ then the functional is the same as the quadratic penalty with the sign of λ flipped. In general this method should be better conditioned than the quadratic penalty method, and convergence can be guaranteed for the penalty parameter μ above a finite threshold, and not only for $\mu \rightarrow \infty$. The constrained problem can be then solved by simultaneously minimizing \mathcal{L} over c while updating the Lagrange multiplier λ .

We approximately solve the sequence of unconstrained problems given by

$$c_k = \underset{c \in \mathcal{P}, c(0)=c_0}{\operatorname{argmin}} \mathcal{L}(c, \lambda_k, \mu_k), \quad (4.4)$$

where μ_k is a given sequence of positive scalars which weights the constraint error penalty term, λ_k is the current estimate of the Lagrange multiplier which is updated *via* the rule

$$\lambda_{k+1} = \lambda_k - \mu_k d^{\text{Var}}(c_k(1), c_1)^2. \quad (4.5)$$

At each iteration we check if the soft constraint is satisfied, if so we accept the current value of μ and continue, if not we increase μ in order to enforce the constraints. At each iteration step, we need to solve an unconstrained minimization problem, for this we use the L-BFGS method in the HANSO library [42]. In practice we only need to solve the sequence of problems with a sequence of gradient tolerances $\tau_k \rightarrow 0$. For small k the tolerance can be chosen quite high to quickly terminate the optimization algorithm. The whole method is summarized in Algorithm 1. As opposed to the quadratic penalty method, the augmented Lagrangian method seemed less sensitive to the choice of sequence of μ_k but at the cost of solving several unconstrained optimization problems instead of a single one. If solving each unconstrained optimization problem is difficult, it might be computationally inefficient to use an augmented Lagrangian method.

In order to additionally factor out the action of the Euclidean motion group $SE(d)$ and scalings, we can simply replace the constraint terms involving the varifold distance by

$$d^{\text{Var}}(c(1), rA(c_1 + b)), \quad (r, A, b) \in \mathbb{R}^+ \ltimes SE(d) \quad (4.6)$$

and add (r, A, b) to the list of variables in each unconstrained minimization subproblem. Observe that there are several orderings of the group actions that would have been possible, we choose to translate first in order to be able to center the curves before rotating them. Finally, we want to add some remarks on alternative methods:

Algorithm 1. Augmented Lagrangian.

Input: Curves c_0, c_1 to be matched.
Set $\mu_0 > 0, \lambda_0 \leq 0, \tau_0 > 0, \tau_{final} > 0, c_{init}^0(t, \theta) = c_0(\theta)$
for $k = 0, 1, 2, \dots, k_{\max}$ **do**
 $c_k = \underset{c}{\operatorname{argmin}} \mathcal{L}(c, \lambda_k, \mu_k)$, with stopping criteria $\|\nabla_c \mathcal{L}(c_k, \lambda_k, \mu_k)\| < \tau_k$.
 if $d^{\text{Var}}(c_k(1), c_1)^2 \leq \varepsilon$ and $\tau_k \leq \tau_{final}$ **then**
 return c_k
 end if
 $\lambda_{k+1} = \lambda_k - \mu_k d^{\text{Var}}(c_k(1), c_1)^2$
 if $\|d^{\text{Var}}(c_k(1), c_1)^2\| < \varepsilon$ **then**
 $\mu_{k+1} = \mu_k$
 else
 $\mu_{k+1} = \varrho \mu_k$
 end if
 if $\tau_k < \tau_{final}$ **then**
 $\tau_{k+1} = \frac{1}{2} \tau_k$
 else
 $\tau_{k+1} = \tau_k$
 end if
 $c_{init}^{k+1} = c_k$
end for

Remark 4.1 (Discretization of $\text{Diff}(M)$). To solve the geodesic boundary value problem on shape space, we have proposed in [5] a method that also discretizes the reparametrization group $\text{Diff}(S^1)$ using B-splines. The action of the reparametrization group is by composition, which does not preserve the B-spline space, as degrees are added. To overcome this we added an L^2 projection step, the composition $c \circ \varphi$ is projected back into a fixed lower order spline space. This has the disadvantage that the projection can smooth out details of the original curve, depending on how many control points are used and which parts of the curve are reparametrized. Furthermore, these methods require a good choice of an initial path, which turned out to be a nontrivial obstacle in examples where the shapes under consideration are sufficiently different from each other. The inexact matching algorithm presented in this paper does not have these problems as we can always choose to initialize the optimization procedure with the constant path.

Remark 4.2 (Dynamic programming). In the SRVF framework [9, 46], a dynamic programming method is usually used to find a global solution of the geodesic boundary value problem on shape space. This relies heavily on access to a fast formula for the geodesic distance between parametrized curves (for the SRVF metric and open curves there even exists an explicit analytic formula). For the bigger class of metrics considered in this article calculating the geodesic distance between parametrized curves is comparable fast to calculating the geodesic distance between unparametrized curves. Thus dynamic programming, which relies on iteratively calculating geodesic distances between parametrized curves, is not well-suited for this particular class of metrics.

5. EXPERIMENTS

The choice of constants in the metric does matter – it influences both the path of the minimal geodesic as well as the parametrization of the endpoint. To illustrate this we computed geodesics between the two curves c_0, c_1 shown in the first and last column of Figure 2. Both curves have length 2π and are parametrized by arc length. We use the metric

$$G_c(h, k) = \int_I a_1 \langle D_s h^\top, D_s k^\top \rangle + b_1 \langle D_s h^\perp, D_s k^\perp \rangle + a_2 \langle D_s^2 h, D_s^2 k \rangle \, ds ,$$

with the following choices of constants:

$$\begin{aligned} (a_1, b_1, a_2) &= (10, 0.1, 10^{-3}) \\ (a_1, b_1, a_2) &= (10, 1, 10^{-3}) \\ (a_1, b_1, a_2) &= (10, 10, 10^{-3}) . \end{aligned}$$

In other words, we change the relative weighting of the normal and tangential components in the H^1 -part of the metric. The geodesics are computed modulo translations.

We see that the geodesic in the first row bends the curve to flatten the bump in the initial curve, c_0 , and to create the bump in the target curve, c_1 . Note that the tip of the bump in c_0 is being matched to the fold at the bottom of c_1 . However, in the third row the bump is translated from c_0 to c_1 resulting in stretching and compression along the curve; in particular the tip of c_0 is matched to the tip of c_1 . This is expected, because in this example $b_1 = 10$ and thus bending is more costly now. The middle row shows intermediate behavior between the both extremes.

5.1. Scale-invariant metrics

In applications to shape analysis the scale of the curves often has no natural meaning. To factor out scale differences, it has been proposed to re-scale the curves to fixed length and to perform the analysis on these constant length curves, see [5, 46]. However, for a nonscale-invariant metric the choice of scale, *i.e.*, the common length of the curves, has a large effect on the resulting analysis, as demonstrated in Table 1.

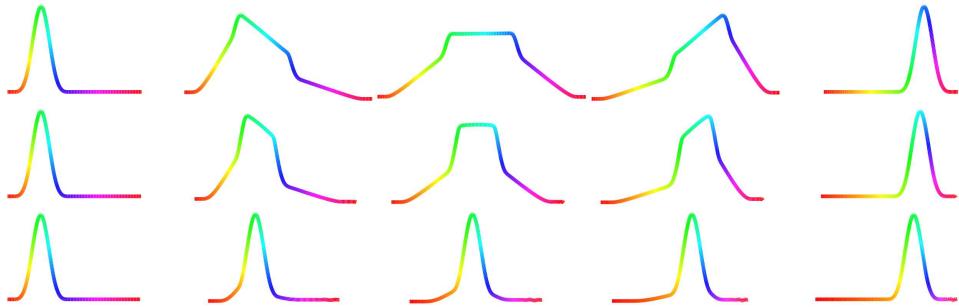


FIGURE 2. Matching of curves with different constants in the metric. The initial and final curves are the same in all rows. The constants are $a_0 = 0$, $a_1 = 10$ and $a_2 = 10^{-3}$ for all rows. The constant b_1 is 0.1 in row one, 1 in row two and 10 in row 3.

TABLE 1. First 4 rows: Geodesic distance for scale and nonscale invariant metrics between different shapes on varying scales. The constants in the metric were chosen to be $a_0 = a_1 = b_1 = 1$, $a_2 = 10^{-4}$. Note that one has to adapt the parameters for the varifold distance to the curve lengths. Fish refers to shapes from the Surrey fish dataset and Corp. Cal. refers to shapes from a collection of outlines of corpi callosi.

	Fish		Corp. Cal.		Corp. Cal.-Fish	
	G^1	G^2	G^1	G^2	G^1	G^2
$d(c_1, c_2)$	0.2423	0.2411	0.1305	0.1305	0.4367	0.4437
$d(2c_1, 2c_2)$	0.2875	0.2411	0.1551	0.1305	0.5231	0.4437
$d(3c_1, 3c_2)$	0.3202	0.2411	0.1670	0.1305	0.5538	0.4436
$d(5c_1, 5c_2)$	0.3521	0.2398	0.1755	0.1305	0.5699	0.4435

Furthermore, even after choosing a “good” scale (or alternatively using a scale-invariant metric) the approach of comparing curves at a fixed length might yield sub-optimal results. In particular, in the presence of noise, rescaling the curves to constant length might artificially change their relative sizes.

Using scale-invariant metrics of the form (2.2) overcomes both of these difficulties: the scale invariance of the metric makes the analysis independent of the choice of scale of the curves and it allows one to consider the induced Riemannian metric on the quotient space of curves modulo scalings by optimizing over all rescalings of the target curve. Consequently, this method automatically fits the optimal size of the target curve with respect to the relative size of the initial curve.

This is shown in Figure 3 where both strategies for dealing with scale variations are compared in a situation where the target curve’s vertices are also corrupted by noise. As fidelity term, we use in this case a varifold metric with a linear function γ which corresponds to the model of currents and was shown (*cf.* [29]) to provide better robustness to such noise. As one can see in the first row of the figure curve length gives in that example a rather bad estimate for the rescaling factor and leads to a quite unnatural mapping where the usual cancellation effects of current fidelity metrics appear when trying to shrink the initial curve. In contrast, using a scale-invariant elastic metric (second row) allows one to jointly estimate a more sensible re-scaling parameter of the target shape (the variable r in (4.6)) together with a more natural path in the space of curves.

5.2. Intrinsic vs extrinsic metric matching

Another benefit of relaxing the constraint with varifold terms is that our new formulation can be more directly compared to another important class of shape space metrics and matching algorithms. We refer to

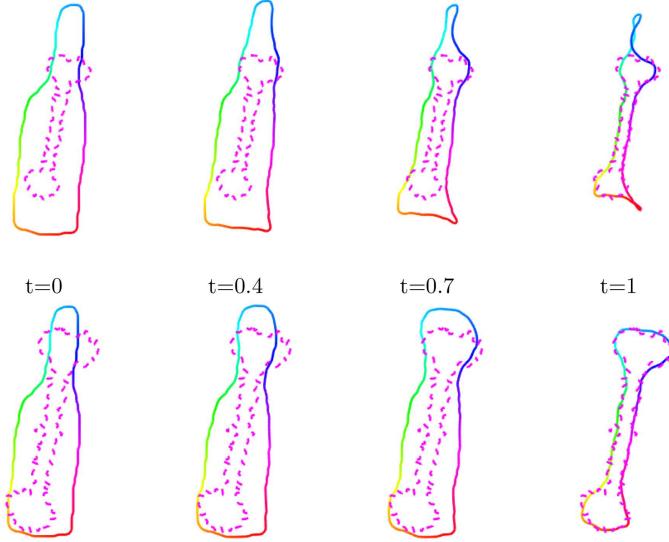


FIGURE 3. Matching of curves with scale differences. The original noisy target curve is twice the size of the source. The first row shows the matching result obtained by length-normalization of the target (magenta curve). The last row is the registration obtained with the scale-invariant metric with simultaneous estimation of the rescaling parameter. The algorithm finds an optimal scaling of the target of 0.505 and the rescaled target is shown in magenta.

those as “extrinsic” as they are usually related to the general model of shape spaces laid out by Grenander in [28]. In this model, distances and geodesics between two shapes are induced by a distance on a certain group of transformations of the entire embedding space that “act” on the shapes: in other words, the distance is quantified by the minimal amount of deformation necessary to map one shape to the other.

In most situations of interest, the groups in question are constructed as subgroups of $\text{Diff}(\mathbb{R}^d)$, the group of diffeomorphisms of the ambient space \mathbb{R}^d , equipped with a right-invariant metric. Multiple models for such groups and metrics have been proposed. In this section, we will focus on comparing our method with one of them: the LDDMM framework originally introduced in [11]. In the case of curves LDDMM inexact matching is typically formulated as the optimal control problem

$$\inf_{v \in L^2([0,1], V)} \int_0^1 \|v(t, \cdot)\|_V^2 dt + \lambda \|\mu_{c(1)} - \mu_{c_1}\|_{W^*}^2 \quad (5.1)$$

on the time-dependent vector field $v \in L^2([0, 1], V)$ where V is a given RKHS of smooth vector fields on \mathbb{R}^d , subject to the constraint $c(0) = c_0$ and $c_t = v(t, c(t))$. Note that the matching constraint is enforced again through a relaxation term based on the varifold metrics of Section 3 (in fact [25] uses metrics with $\gamma(u) = u$ while [21] considers $\gamma(u) = u^2$ in applications). The essential difference between our formulation (4.2) and (5.1) is the fact that in (5.1) the vector field is defined over the whole space \mathbb{R}^d and its energy is measured by the global norm $\|\cdot\|_V$.

This has a few important consequences. One key property of the LDDMM model is that it enforces the global transformation resulting from the flow of v to be diffeomorphic. In particular, it will prevent any self-intersection from occurring along geodesics. In contrast, geodesics for the elastic Sobolev metrics of this paper lie in the space of immersions and, as illustrated in Figure 4, self-intersections can appear in geodesics even if the initial and final curves are embedded curves.

On the other hand, estimating a global diffeomorphic transformations as given by the LDDMM model may prove a particularly difficult or undesirable constraint in certain situations. It is most notably the case when

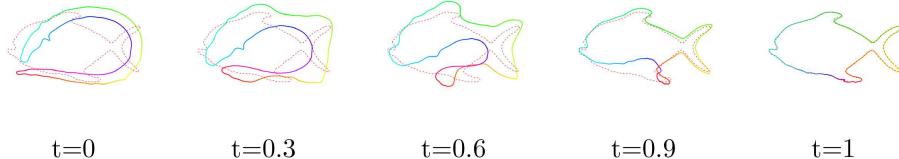


FIGURE 4. An example of elastic Sobolev geodesic between two curves (the target is the red-dotted one). As opposed to extrinsic deformation models like LDDMM, self-intersections can be created in geodesic paths.

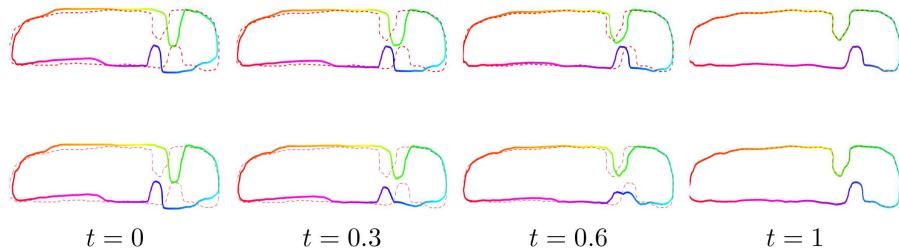


FIGURE 5. Comparison of geodesics between curves with traveling bumps. First line: intrinsic H^2 -metric; second line: LDDMM geodesic. In the LDDMM model one bump is successively flattened and recreated due to the high deformation cost of opposite displacements of close objects. In the intrinsic model the two bumps are merely transported.

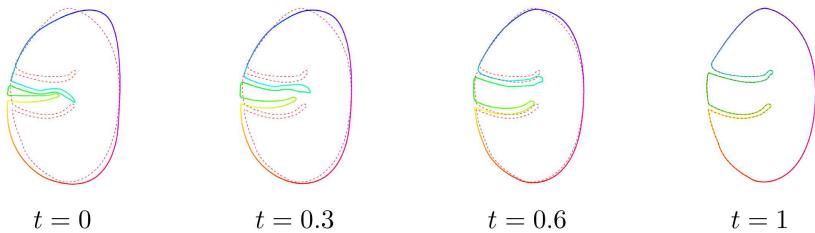


FIGURE 6. Estimated registration between the two curves with an intrinsic H^2 metric.

thin or closely located structures have to be displaced or stretched apart. We illustrate such a phenomenon in the example of Figure 5. The motion of bumps on two opposite sides of a curve is estimated in fundamentally different ways by the intrinsic model for which the two bumps are simply displaced along the curve and by LDDMM where one bump is successively flattened and recreated due to the high deformation cost of opposite displacements of close by objects. Along the same lines, Figures 6 and 7 show another comparison in which two “sulci” have to be moved apart. Since this is again a costly deformation in the LDDMM framework, it is easily prone to reach unnatural solutions if the deformation kernel is too large or to lead to even more unnatural local minima of the functional (5.1) for small deformation kernels. We also point out that similar issues are discussed quite extensively in the recent work [50], which in addition introduces a hybrid model combining a global LDDMM deformation cost with intrinsic H^1 penalties.

5.3. Time comparison

As a last set of experiments, we take a closer look at the computational complexity for both of the previous registration models. As detailed earlier in Section 4, the bulk of the computations for the proposed elastic metric approach at each optimization step is 1) the evaluation of splines and their derivatives to compute the elastic

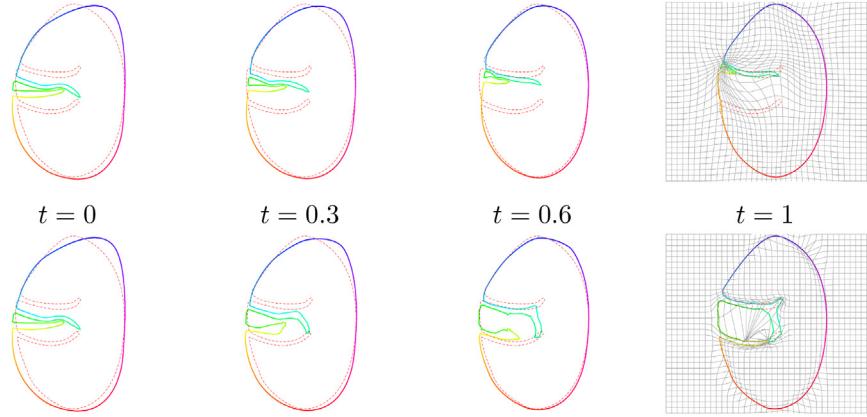


FIGURE 7. Registration using LDDMM. On the first row, using a Gaussian deformation kernel of width 0.2. On the second row, with a Gaussian kernel of width 0.05. Note that if reducing the kernel size allows theoretically to recover finer scale deformations, it is well-known that the LDDMM registration problem then becomes highly sensitive to bad local minima such as the solution shown in the second row.

TABLE 2. Computational time comparison between the proposed algorithm and LDDMM for 380 pairwise registrations of curves with 100 vertices each. We report the total run time, the average number of iterations for the optimization methods and the average time of one iteration.

	Total run time	Avg. n^o of iter./registration	Avg. time/iter.
Elastic	66 min	479	0.02 s
LDDMM	93 min	91	0.16 s

energy and its gradient for which the required number of basic operations is linear in the total number $N_\theta N_t$ of spline control points and 2) the evaluation of the varifold fidelity term which requires $O(N_\theta^2)$ evaluations of the kernels ρ and γ . Standard LDDMM algorithms, in comparison, usually solve (5.1) with a forward-backward shooting procedure involving, at each iteration of the optimization, the integration of Hamiltonian systems of equations with interacting particles, which will typically require of the order of $O(N^2 T)$ evaluations of the kernel of the RKHS V , with N being the number of vertices of the curve and T the number of time steps for the chosen numerical integrator.

We illustrate this comparison empirically by running the two registration algorithms on a subset of 20 closed curves from the Surrey fish dataset and estimating all the 380 pairwise matchings for each method. Note that, although the two optimization problems and part of their underlying parameters differ, the quality of the registration is in both cases measured by the varifold metric. In this experiment, we choose the same scale parameters for the fidelity term and adapt the other parameters relative to the deformation metrics to lead comparable convergence properties and registration accuracy for both methods. All curves are set to $N = 100$ vertices which are also the control points used in the spline representation (thus $N_\theta = N$), and we take $T = 10$ time steps for LDDMM as well as $N_t = 10$ time control points for splines in our proposed algorithm. The implementation of curve matching LDDMM is the one of [19, 29] with the optimization routine given by the same limited memory BFGS algorithm from the HANSO library that is used in this work.

The results are reported in Table 2. There are a few remarks to be made. First, on average, the time for a single iteration of the optimization procedure is significantly lower with the approach of this paper compared to LDDMM, which is consistent with the previous discussion on the theoretical complexity for the computation of deformation energies and gradients in both cases. Second, still on average, the LDDMM algorithm requires

a priori less iterations for BFGS to reach convergence, with the same stopping criterion being used. This is likely due, on the one hand, to the fact that the optimization in LDDMM is performed over the deformation's initial momenta as opposed to the full path of spline parameters in our approach, thus reducing the size of the problem. On the other hand, it is also important to point out that this may be in part due to BFGS occasionally converging to irrelevant local minima (in very few number of iterations) in the case of LDDMM. In this precise experiment, this happens for about 30 registration cases in which the residual varifold cost remains very high at the end of the minimization. In contrast, the convergence seems much more consistent in the case of the elastic method as the total number of iterations and final energies do not vary as significantly from one registration to another.

As additional future comparison, it will be interesting to investigate the influence of the choice of N_θ or N on the computational time and convergence properties for the two models. We postulate that the use of splines represented by their control points instead of directly vertices could allow N_θ to be in practice much smaller than N while still providing consistent registration results for smooth curves.

APPENDIX A. DERIVATIVES OF THE ENERGY FUNCTIONAL

In this appendix, we list the derivatives of the energy functional (2.3) and varifold distance (3.3). The first derivative of the energy

$$\begin{aligned} dE_c(k) = & \int_0^1 \int_0^{2\pi} t_1 \langle c', k' \rangle + t_2 (\langle c'', k' \rangle + \langle c', k'' \rangle) + t_3 \langle \dot{c}'^\perp, k' \rangle + t_4 \langle \dot{c}, \dot{k} \rangle + t_5 \langle \dot{c}', \dot{k}' \rangle \\ & + t_6 \langle \dot{c}'^\top, \dot{k}' \rangle + t_7 \langle \dot{c}'^\perp, \dot{k}' \rangle + t_8 (\langle \dot{c}'', \dot{k}' \rangle + \langle \dot{c}', \dot{k}'' \rangle) + t_9 \langle \dot{c}'', \dot{k}'' \rangle \\ & + d\ell_c(k) \left(a'_0(\ell) |c'| \langle \dot{c}, \dot{c} \rangle + a'_1(\ell) \frac{\langle \dot{c}'^\top, \dot{c}'^\top \rangle}{|c'|} + b'_1(\ell) \frac{\langle \dot{c}'^\perp, \dot{c}'^\perp \rangle}{|c'|} \right. \\ & \left. + a'_2(\ell) \left(\frac{\langle c', c'' \rangle^2}{|c'|} + \frac{\langle \dot{c}'', \dot{c}'' \rangle}{|c'|^3} - \frac{2 \langle c', c' \rangle \langle \dot{c}'', \dot{c}' \rangle}{|c'|^5} \right) \right) d\theta dt, \end{aligned}$$

with

$$\begin{aligned} t_1 = & \frac{a_0}{|c'|} \langle \dot{c}, \dot{c} \rangle - \frac{a_1}{|c'|^3} \langle \dot{c}'^\top, \dot{c}'^\top \rangle - \frac{b_1}{|c'|^3} \langle \dot{c}'^\perp, \dot{c}'^\perp \rangle - 7 \frac{a_2}{|c'|^9} \langle c', c'' \rangle^2 \langle \dot{c}', \dot{c}' \rangle \\ & + 10 \frac{a_2}{|c'|^7} \langle c', c'' \rangle \langle \dot{c}', \dot{c}'' \rangle - 3 \frac{a_2}{|c'|^5} \langle \dot{c}'', \dot{c}'' \rangle, \\ t_2 = & 2 \frac{a_2}{|c'|^7} \langle c', c'' \rangle \langle \dot{c}', \dot{c}' \rangle - 2 \frac{a_2}{|c'|^5} \langle \dot{c}', \dot{c}'' \rangle, \quad t_3 = 2 \frac{a_1 - b_1}{|c'|^3} \langle \dot{c}', c' \rangle, \quad t_4 = 2a_0 |c'|, \\ t_5 = & 2 \frac{a_2}{|c'|^7} \langle c', c'' \rangle, \quad t_6 = 2 \frac{a_1}{|c'|}, \quad t_7 = 2 \frac{b_1}{|c'|} \quad t_8 = -2 \frac{a_2}{|c'|^5} \langle c', c'' \rangle, \quad t_9 = 2 \frac{a_2}{|c'|^3}. \end{aligned}$$

The varifold distance as a function of only its left argument is given by

$$F(c_1) = \langle \mu_{c_1}, \mu_{c_2} \rangle_{\text{Var}} = \iint_{S^1 \times S^1} \rho(|c_1(\theta_1) - c_2(\theta_2)|^2) \gamma(\langle v_1, v_2 \rangle) ds_1 ds_2.$$

with the tangent and normal vectors defined by

$$v_1(\theta_1) = \frac{c'_1(\theta_1)}{|c'_1(\theta_1)|}, \quad v_2(\theta_2) = \frac{c'_2(\theta_2)}{|c'_2(\theta_2)|}$$

The variation of these quantities is simply

$$D_{c_1,h}(v_1) = \langle D_s h, n \rangle n, \quad D_{c_1,h}(v_2) = 0.$$

The derivative is given by the formula

$$\begin{aligned} dF_{c_1}(h) &= \iint_{S^1 \times S^1} 2\rho'(|c_1(\theta_1) - c_2(\theta_2)|^2) \gamma(\langle v_1, v_2 \rangle) \langle c_1(\theta_1) - c_2(\theta_2), h(\theta_1) \rangle ds_1 ds_2 \\ &\quad + \iint_{S^1 \times S^1} \rho(|c_1(\theta_1) - c_2(\theta_2)|^2) \gamma'(\langle v_1, v_2 \rangle) \langle D_{s_1} h, n_1 \rangle \langle n_1, v_2 \rangle ds_1 ds_2 \\ &\quad + \iint_{S^1 \times S^1} \rho(|c_1(\theta_1) - c_2(\theta_2)|^2) \gamma(\langle v_1, v_2 \rangle) \langle D_{s_1} h, v_1 \rangle ds_1 ds_2. \end{aligned}$$

Acknowledgements. We would like to thank Philipp Harms, Eric Klassen, Sebastian Kurtek, Peter Michor, Tom Needham, Anuj Srivastava and the Shape Group at FSU for helpful comments and discussions. Nicolas Charon is supported by the National Science Foundation under Grant No 1819131.

REFERENCES

- [1] N. Aronszajn, Theory of reproducing kernels. *Trans. Am. Math. Soc.* **68** (1950) 337–404.
- [2] M. Bauer, M. Bruveris, N. Charon and J. Møller-Andersen, Varifold-Based Matching of Curves via Sobolev-Type Riemannian Metrics. In *6th MICCAI Workshop on Mathematical Foundations of Computational Anatomy* (2017) 152–163.
- [3] M. Bauer, M. Bruveris, P. Harms and P.W. Michor, Vanishing geodesic distance for the Riemannian metric with geodesic equation the KdV-equation. *Ann. Global Anal. Geom.* **41** (2012) 461–472.
- [4] M. Bauer, M. Bruveris, P. Harms and J. Møller-Andersen, Curve Matching with Applications in Medical Imaging. In *5th MICCAI Workshop on Mathematical Foundations of Computational Anatomy* (2015).
- [5] M. Bauer, M. Bruveris, P. Harms and J. Møller-Andersen, A numerical framework for Sobolev metrics on the space of curves. *SIAM J. Imaging Sci.* **10** (2017) 47–73.
- [6] M. Bauer, M. Bruveris, S. Marsland and P.W. Michor, Constructing reparameterization invariant metrics on spaces of plane curves. *Differ. Geom. Appl.* **34** (2014) 139–165.
- [7] M. Bauer, M. Bruveris and P.W. Michor, Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vis.* **50** (2014) 60–97.
- [8] M. Bauer, M. Bruveris and P.W. Michor, Why use Sobolev metrics on the space of curves, in *Riemannian Computing in Computer Vision*. Springer, Cham (2016) 233–255.
- [9] M. Bauer, M. Eslitzbichler and M. Grasmair, Landmark-guided elastic shape analysis of human character motions. *Inverse Probl. Imaging* **11** (2017) 601–621.
- [10] M. Bauer and P. Harms, Metrics on spaces of immersions where horizontality equals normality. *Differ. Geom. Appl.* **39** (2015) 166–183.
- [11] M.F. Beg, M.I. Miller, A. Trouvé and L. Younes, Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.* **61** (2005) 139–157.
- [12] J. Benn, S. Marsland, R. McLachlan, K. Modin and O. Verdier, Currents and finite elements as tools for shape space (2017). Preprint [arXiv:1702.02780](https://arxiv.org/abs/1702.02780).
- [13] M. Bruveris, Completeness properties of Sobolev metrics on the space of curves. *J. Geom. Mech.* **7** (2015) 125–150.
- [14] M. Bruveris, Regularity of maps between Sobolev spaces. *Ann. Global Anal. Geom.* **52** (2017) 11–24.
- [15] M. Bruveris, P.W. Michor and D. Mumford, Geodesic completeness for Sobolev metrics on the space of immersed plane curves. *Forum Math. Sigma* **2** (2014) e19.
- [16] M. Bruveris and J. Møller-Andersen, Completeness of length-weighted Sobolev metrics on the space of curves (2017). Preprint [arXiv:1705.07976](https://arxiv.org/abs/1705.07976).
- [17] C. Carmeli, E. De Vito, A. Toigo and V. Umanita, Vector valued reproducing kernel Hilbert spaces and universality. *Anal. Appl.* **8** (2010) 19–61.
- [18] V. Cervera, F. Mascaró and P.W. Michor, The action of the diffeomorphism group on the space of immersions. *Differ. Geom. Appl.* **1** (1991) 391–401.
- [19] B. Charlier, N. Charon and A. Trouvé, Fshapes tool kit, 2014. Available at <https://github.com/fshapes/fshapesTk> (2019).
- [20] N. Charon, *Analysis of geometric and functional shapes with extensions of currents: application to registration and atlas estimation*. PhD thesis, ENS Cachan (2013).
- [21] N. Charon and A. Trouvé, The varifold representation of non-oriented shapes for diffeomorphic registration. *SIAM J. Imaging Sci.* **6** (2013) 2547–2580.

- [22] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé and N. Ayache, Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage* **55** (2010) 1073–1090.
- [23] D.G. Ebin and J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. Math.* **92** (1970) 102–163.
- [24] M. Eslitzbichler, Modelling character motions on infinite-dimensional manifolds. *Vis. Comput.* **31** (2015) 1179–1190.
- [25] J. Glaunès, A. Qiu, M. Miller and L. Younes, Large deformation diffeomorphic metric curve mapping. *Int. J. Comput. Vis.* **80** (2008) 317–336.
- [26] J. Glaunès, A. Trouvé and L. Younes, Diffeomorphic matching of distributions: a new approach for unlabelled point-sets and sub-manifolds matching. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2** (2004) 712–718.
- [27] J. Glaunès and M. Vaillant, Surface matching via currents. *Proc. Inform. Process. Med. Imaging (IPMI)*. In Vol. 3565 of *Lect. Notes Comput. Sci.* **3565** (2006) 381–392.
- [28] U. Grenander, General Pattern Theory: A Mathematical Study of Regular Structures. *Oxford Mathematical Monographs*. Clarendon Press, Oxford (1993).
- [29] I. Kaltenmark, B. Charlier and N. Charon, A general framework for curve and surface comparison and registration with oriented varifolds. *IEEE Conf. Comput. Vis. Pattern Recognit.* (2017) 4580–4589.
- [30] E. Klassen, A. Srivastava, W. Mio and S.H. Joshi, Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **26** (2004) 372–383.
- [31] A. Kriegel and P.W. Michor, The Convenient Setting of Global Analysis. Vol. 53 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (1997).
- [32] S. Kurtek and T. Needham, Simplifying transforms for general elastic metrics on the space of plane curves (2018). Preprint [arXiv:1803.10894](https://arxiv.org/abs/1803.10894).
- [33] H. Laga, S. Kurtek, A. Srivastava and S.J. Miklavcic, Landmark-free statistical analysis of the shape of plant leaves. *J. Theor. Biol.* **363** (2014) 41–52.
- [34] A.C. Mennucci, A. Yezzi and G. Sundaramoorthi, Properties of Sobolev-type metrics in the space of curves. *Interface. Free Bound.* **10** (2008) 423–445.
- [35] P.W. Michor and D. Mumford, Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.* **8** (2006) 1–48.
- [36] P.W. Michor and D. Mumford, An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Appl. Comput. Harmon. Anal.* **23** (2007) 74–113.
- [37] W. Mio, J.C. Bowers and X. Liu, Shape of elastic strings in euclidean space. *Int. J. Comput. Vis.* **82** (2009) 96–112.
- [38] W. Mio and A. Srivastava, Elastic-String Models for Representation and Analysis of Planar Shapes. Vol. 2 of *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR* (2004) II-10–II-15.
- [39] W. Mio, A. Srivastava and S. Joshi, On shape of plane elastic curves. *Int. J. Comput. Vis.* **73** (2007) 307–324.
- [40] G. Nardi, G. Peyré and F.-X. Vialard, Geodesics on shape spaces with bounded variation and Sobolev metrics. *SIAM J. Imaging Sci.* **9** (2016) 238–274.
- [41] J. Nocedal and S. Wright, Numerical Optimization. Springer, New York (2006).
- [42] M. Overton, HANSO: hybrid algorithm for non-smooth optimization 2.2, 2016. Available at <https://cs.nyu.edu/overton/software/hanso/> (2019).
- [43] J. Shah, H^0 -type Riemannian metrics on the space of planar curves. *Quart. Appl. Math.* **66** (2008) 123–137.
- [44] B. Sriperumbudur, K. Fukumizu and G. Lanckriet, On the Relation Between Universality, Characteristic Kernels and RKHS Embedding of Measures. Vol. 9 of *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* (2010) 773–780.
- [45] A. Srivastava and E. Klassen, Functional and Shape Data Analysis. *Springer Series in Statistics*. Springer-Verlag, New York (2016).
- [46] A. Srivastava, E. Klassen, S.H. Joshi and I.H. Jermyn, Shape analysis of elastic curves in Euclidean spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **33** (2011) 1415–1428.
- [47] J. Su, S. Kurtek, E. Klassen and A. Srivastava, Statistical analysis of trajectories on Riemannian manifolds: bird migration, hurricane tracking and video surveillance. *Ann. Appl. Stat.* **8** (2014) 530–552.
- [48] J. Su, A. Srivastava, F.D.M. de Souza and S. Sarkar, Rate-invariant analysis of trajectories on Riemannian manifolds with application in visual speech recognition. *IEEE Conf. Comput. Vis. Pattern Recognit.* **6** (2014) 620–627.
- [49] Z. Su, E. Klassen and M. Bauer, The square root velocity framework for curves in a homogeneous space. In *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 07 (2017) 680–689.
- [50] L. Younes, Hybrid Riemannian metrics for diffeomorphic shape registration. *Ann. Math. Sci. Appl.* **3** (2018) 189–210.
- [51] L. Younes, Computable elastic distances between shapes. *SIAM J. Appl. Math.* **58** (1998) 565–586.
- [52] L. Younes, P.W. Michor, J. Shah and D. Mumford, A metric on shape space with explicit geodesics. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur.* **19** (2008) 25–57.