

# CONSTRUCTING REPARAMETERIZATION INVARIANT METRICS ON SPACES OF PLANE CURVES

MARTIN BAUER, MARTINS BRUVERIS, STEPHEN MARSLAND, PETER W. MICHOR

**ABSTRACT.** Metrics on shape spaces are used to describe deformations that take one shape to another, and to define a distance between shapes. We study a family of metrics on the space of curves, which includes several recently proposed metrics, for which the metrics are characterised by mappings into vector spaces where geodesics can be easily computed. This family consists of Sobolev-type Riemannian metrics of order one on the space  $\text{Imm}(S^1, \mathbb{R}^2)$  of parameterized plane curves and the quotient space  $\text{Imm}(S^1, \mathbb{R}^2)/\text{Diff}(S^1)$  of unparameterized curves. For the space of open parameterized curves we find an explicit formula for the geodesic distance and show that the sectional curvatures vanish on the space of parameterized open curves and are non-negative on the space of unparameterized open curves. For one particular metric we provide a numerical algorithm that computes geodesics between unparameterized, closed curves, making use of a constrained formulation that is implemented numerically using the RATTLE algorithm. We illustrate the algorithm with some numerical tests between shapes.

## 1. INTRODUCTION

The mathematical analysis of shape has been the focus of intense research interest in recent years, not least because of applications in image analysis and computer vision, where methods based on geodesic active contours or ‘snakes’ are used for segmentation, tracking and object recognition [28, 29]. Another source of applications is biomedical image analysis, where the study and comparison of shapes form a large part of the field of computational anatomy [11, 12].

A key problem in shape analysis is to define a distance function between shapes that can measure similarity in a computationally feasible way and act as the basis for object classification. One way to arrive at such a distance function is to equip the space of shapes with a Riemannian metric, which allows the lengths of paths between shapes to be measured. The distance between two shapes can then be defined as the length of the shortest path connecting them.

For the purposes of this paper we consider shapes to be smooth plane curves (open or closed) modulo smooth reparameterizations. A slightly narrower definition would be to define a shape as the outline of a smooth, simply connected domain in the plane; this definition excludes objects like the figure of eight, which we allow in our definition. Mathematically, a shape is represented by a smooth curve  $c : S^1 \rightarrow \mathbb{R}^2$ . Two curves  $c, d$  represent the same shape if there exists a reparameterization map  $\varphi \in \text{Diff}(S^1)$  (that is, the group of smooth, orientation preserving, invertible maps

---

2010 *Mathematics Subject Classification.* 58B20, 58D15, 65D18.

*Key words and phrases.* curve matching, elastic metric, geodesic shooting, reparameterization group, Riemannian shape analysis, shape space .

$\varphi : S^1 \rightarrow S^1$  from the circle onto itself) such that one curve is a reparameterization of the other, i.e.,  $c = d \circ \varphi$ .

We will work with the class of regular or immersed curves  $c$ . A curve  $c$  is regular if it has a non-vanishing tangent, i.e.,  $c'(\theta) \neq 0$ . The diffeomorphism group  $\text{Diff}(S^1)$  acts on the space of immersed curves

$$\text{Imm}(S^1, \mathbb{R}^2) := \{c \in C^\infty(S^1, \mathbb{R}^2) \mid c'(\theta) \neq 0\},$$

from the right via  $(\varphi, c) \mapsto c \circ \varphi$ . Using this setting we can identify a shape with an equivalence class  $[c]$ , that is an element of the quotient space  $\text{Imm}(S^1, \mathbb{R}^2) / \text{Diff}(S^1)$ . This quotient is the shape space of immersed closed curves modulo reparameterizations and will be denoted by  $\mathcal{S}$ . Similarly, one defines the space  $\text{Imm}([0, 2\pi], \mathbb{R}^2)$  of parameterized open curves and the space  $\mathcal{S}_{\text{open}}$  of open shapes.

To arrive at a distance function on shape space requires two steps. First, we define a Riemannian metric (i.e., an inner product measuring the length of infinitesimal deformations of a curve) on the space of immersed curves and compute geodesics on this space. The deformations  $h, k$  are represented by vector fields along the curve  $c$  and the inner product, which depends on the curve, is denoted by  $G_c(h, k)$ . If the metric is invariant under the action of the reparameterization group  $\text{Diff}(S^1)$ , then it induces a Riemannian metric on shape space  $\mathcal{S}$ , which in turn gives rise to the geodesic distance function. The second step is to find the right representatives  $c, d$  of the equivalence classes  $[c]$  and  $[d]$ , such that the geodesic distance  $\text{dist}^{\text{Imm}}(c, d)$  coincides with  $\text{dist}^{\mathcal{S}}([c], [d])$ .

**1.1. Shape metrics and related work.** The simplest reparameterization invariant metric on  $\text{Imm}(S^1, \mathbb{R}^2)$  is the  $L^2$ -metric

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

where we integrate over arc-length,  $ds = |c'(\theta)|d\theta$ . However, the geodesic distance induced by this metric vanishes, i.e., the distance between any two shapes is 0, which renders it unsuitable for shape analysis [20].

One way to overcome this is to add terms involving higher derivatives of  $h$  and  $k$  to the metric, such as:

$$G_c(h, k) = \int_{S^1} \langle h, k \rangle + A \langle D_s h, D_s k \rangle ds,$$

where  $D_s h = \frac{1}{|c'|} h'$  denotes the arc-length derivative of  $h$ , and  $A > 0$  is a constant. This leads to the class of Sobolev-type metrics, which were independently introduced in [8, 22, 29] and studied further in [4, 25].

Another family of metrics, the almost local metrics [5, 6, 21], prevent the geodesic distance from vanishing by introducing a weight function in the integral. Examples of weight functions involving the curvature or length are  $w(\theta) = 1 + A\kappa(\theta)^2$  and  $w(\theta) = \ell(c)^{-1}$ .

Sobolev-type metrics of arbitrary order were studied in [7, 18, 22]. Although they are a natural generalization of the  $L^2$ -metric from a theoretical point of view, their numerical treatment is rather involved, mainly because the geodesic equation of a Sobolev-type metric of order  $k$  is a nonlinear PDE of order  $2k$ . Interestingly, there are special cases of first order metrics for which the geodesic equation admits explicit solutions [28, 31]. Apart from these special cases, there have been some attempts to solve the geodesic equation directly for Sobolev-type metrics of order

1 for curves [23] and surfaces [2]. Metrics of higher order, on the other hand, are still practically untouched.

To avoid dealing with PDEs of high order, one can restrict one's attention to Sobolev-type metrics of order one, in particular to the family of metrics known as 'elastic metrics' and studied in [23], which are of the form

$$(1) \quad G_c^{a,b}(h, k) = \int_{S^1} a^2 \langle D_s h, n \rangle \langle D_s k, n \rangle + b^2 \langle D_s h, v \rangle \langle D_s k, v \rangle ds,$$

with constants  $a, b \in \mathbb{R}^+$  and with  $v$  and  $n$  denoting the unit tangent and normal vectors to  $c$ . The term involving the normal vector can be seen as measuring the bending of the curve  $c$  under the deformation  $h$ , while the derivative of  $h$  in the tangential direction measures the stretching of  $c$ .

Following ideas of [30], it was shown in [31] that it is possible to find explicit formulas for geodesics of the scale-invariant version of the elastic metric with  $b^2 = a^2$  on the space of parameterized curves modulo translations, rotations and scaling. To achieve this a curve  $c$  was represented by the square-root of its velocity vector,  $\sqrt{c'}$ , with  $c'$  being interpreted as a complex number. In this representation the  $G^{1,1}$ -metric has a particularly simple form that allows for explicit formulas for geodesics between two curves and the length of the geodesic.

A similar motivation underlies the introduction of the square root velocity transform (SRVT) in [27]. The SRVT is a transformation that maps a curve  $c$  to  $R(c) = c' / \sqrt{|c'|}$  with the effect that the  $G^{a,b}$ -metric simplifies for  $4b^2 = a^2$ . While not allowing explicit formulas for geodesics to be written down, this map greatly simplifies their numerical computation.

Related to the SRVT is the  $Q$ -transform, introduced in [14], which maps a curve  $c$  to  $Q(c) = \sqrt{|c'|}c$ . This transform generalizes easily to surfaces, but does present theoretical difficulties, as explained in Section 3.7.

There are other ways to define a Riemannian metric on shape space such as large deformation diffeomorphic metric mapping [11, 9], where a Riemannian metric is induced from the diffeomorphism group  $\text{Diff}(\mathbb{R}^2)$  of the ambient space, or the use of conformal welding [26] to represent shapes as diffeomorphisms of the circle.

**1.2. The reparameterization group  $\text{Diff}(S^1)$ .** In this article we will consider the group of smooth, orientation preserving, invertible maps  $\varphi : S^1 \rightarrow S^1$  from the circle onto itself. This group is sometimes denoted by  $\text{Diff}^+(S^1)$ . To shorten notation we will write  $\text{Diff}(S^1)$  instead.

After equipping the manifold  $\text{Imm}(S^1, \mathbb{R}^2)$  with a Riemannian metric, there are various ways to perform the minimization over the reparameterization group. The problem is challenging because of the nonlinear nature of  $\text{Diff}(S^1)$  and the fact that all the spaces involved are infinite-dimensional. One approach is to replace the infinite-dimensional space  $\text{Diff}(S^1)$  by a finite-dimensional one and to perform the minimization over this smaller space. A possible choice for the smaller space, used in [27], are diffeomorphisms  $\varphi$  for whom the Fourier series of  $\sqrt{\varphi'}$  is truncated at a fixed length.

In [9] elements of  $\text{Diff}(S^1)$  were generated as flows of vector fields. The advantage of using vector fields is that they form a linear space, which enables the use of gradient-based optimization algorithms. However, due to the regularization terms needed to ensure convergence of the algorithm, the computed distance failed to be symmetric. An alternative is given in [28], where the authors describe an iterative

procedure directly on  $\text{Diff}(S^1)$  based on geometric considerations. We will expand on this minimization scheme in Section 7.

**1.3. Overview of the paper.** This paper arose from the observation that the SRVT [27] and the  $Q$ -transform [13, 14, 17] are both special cases of a general method for constructing reparameterization invariant metrics on spaces of curves. We show in Section 3 how to use this method to obtain the elastic metric with general parameters  $4b^2 \geq a^2$  as well as Sobolev metrics of higher order. The idea is to construct a transformation that maps  $\text{Imm}(S^1, \mathbb{R}^2)$  isometrically to a submanifold of a flat Riemannian manifold.

In the following we exploit this representation of the space  $\text{Imm}(S^1, \mathbb{R}^2)$  to study in particular the family of elastic metrics and their mathematical properties. We chose to focus on these metrics because they – and especially the one induced by the SRVT – are currently used in shape analysis.

The embedding of  $\text{Imm}(S^1, \mathbb{R}^2)$  into a flat space also allows us to discretize the geodesic equation in a very geometric way. After choosing a discretization of the curve, the infinite-dimensional Riemannian manifold becomes a finite dimensional constrained submanifold of the Euclidean space. In Section 7 we describe this procedure – which can be applied to all metrics constructed in Section 3 – in the particular case of the metric arising from the SRVT. The resulting shooting method to solve the geodesic boundary value problem complements the path-straightening approach used in [24].

When computing the distance between unparameterized curves (Section 8) it is necessary to minimize over the reparameterization group. We give a rigorous interpretation of the iterative procedure proposed in [28] as a gradient descent on  $\text{Diff}(S^1)$  with respect to a specific Riemannian metric. Reparametrization denotes the sliding of the points representing the curve along it. Numerical problems arise when points are compressed too closely together or stretched too far apart. We will argue that this behaviour is not introduced by the discretization, but is a manifestation of geodesic incompleteness of the underlying Riemannian metric: there exist shapes  $[c], [d] \in \mathcal{S}$  with no geodesic connecting them. We conclude the paper by discussing the repercussions of the incompleteness on applications to shape analysis.

## 2. NOTATION AND BACKGROUND MATERIAL

**2.1. Notation.** In this section we introduce some notation that we will use throughout the article. We study the spaces of open and closed, regular, plane curves

$$\begin{aligned}\text{Imm}([0, 2\pi], \mathbb{R}^2) &= \{c \in C^\infty([0, 2\pi], \mathbb{R}^2) \mid c'(\theta) \neq 0\} , \\ \text{Imm}(S^1, \mathbb{R}^2) &= \{c \in C^\infty(S^1, \mathbb{R}^2) \mid c'(\theta) \neq 0\} .\end{aligned}$$

The elastic metric, that we will consider in this article, has a kernel on the manifold of immersions, which consists exactly of the constant vector fields. Therefore we introduce the quotient space of immersions modulo translations  $\text{Imm}(M, \mathbb{R}^2)/\text{Tra}$ , for  $M$  either  $S^1$  or  $[0, 2\pi]$ .

We will also need the diffeomorphism group of both  $[0, 2\pi]$  and  $S^1$ , i.e., the groups of smooth, orientation preserving, invertible maps onto itself:

$$\begin{aligned}\text{Diff}([0, 2\pi]) &= \{\varphi \in C^\infty([0, 2\pi], [0, 2\pi]) \mid \varphi'(\theta) > 0\} , \\ \text{Diff}(S^1) &= \{\varphi \in C^\infty(S^1, S^1) \mid \varphi'(\theta) > 0\} .\end{aligned}$$

These groups are sometimes denoted by  $\text{Diff}^+(S^1)$  and  $\text{Diff}^+([0, 2\pi])$  respectively. To shorten notation we will write  $\text{Diff}(S^1)$  and  $\text{Diff}([0, 2\pi])$  instead.

For a curve  $c$  we denote by  $v = c'/|c'|$  the unit length tangent vector and by  $n$  the unit length normal vector, obtained from  $v$  through a rotation by  $\frac{\pi}{2}$ . The arc-length derivative of a function  $f$  along the curve is denoted by  $D_s f = \frac{1}{|c'|} f'$  and  $ds = |c'|d\theta$  is integration with respect to arc-length. The turning angle  $\alpha$  is defined through the relation  $v = (\cos \alpha, \sin \alpha)$  and the curvature by  $\kappa = \langle D_s v, n \rangle$ .

**2.2. Variational formulas.** In this section we will compute some variational formulae for the quantities that have been introduced in the previous section, i.e., we compute how these terms change if we vary the underlying curve  $c$ . For a smooth map  $F$  from  $\text{Imm}(S^1, \mathbb{R}^2)$  to any convenient vector space we denote by

$$dF(c).h = D_{c,h}F = \left. \frac{d}{dt} \right|_{t=0} F(c + th) = \left. \frac{d}{dt} \right|_{t=0} F(\tilde{c}(t, \cdot))$$

the variation in the direction  $h$ , where  $\tilde{c} : \mathbb{R} \times M \rightarrow \mathbb{R}^2$  is any smooth variation with  $\tilde{c}(0, \theta) = c(\theta)$  and  $\partial_t|_0 \tilde{c}(t, \theta) = h(\theta)$  for all  $\theta$ . Examples of maps  $F$  include  $v$ ,  $n$ ,  $\alpha$ ,  $|c'|$ ,  $\kappa$ .

The following formulae will be used repeatedly throughout the article.

**Lemma 2.3.** *The first variations of the unit tangent vector  $v$ , the normal vector  $n$ , the length element  $|c'|$  and the curvature  $\kappa$  are given by:*

$$\begin{aligned}(2) \quad & D_{c,h}v = \langle D_s h, n \rangle n \\ (3) \quad & D_{c,h}n = -\langle D_s h, n \rangle v \\ (4) \quad & D_{c,h}|c'| = \langle D_s h, v \rangle |c'| \\ (5) \quad & D_{c,h}\kappa = \langle D_s^2 h, n \rangle - 2\kappa \langle D_s h, v \rangle .\end{aligned}$$

*Proof.* The proof of these formulae can be found for example in [22]. □

### 3. CONSTRUCTING REPARAMETERIZATION INVARIANT METRICS VIA SHAPE TRANSFORMATIONS

In this section we describe a general method of constructing reparameterization invariant metrics on the space of plane curves. This method will include the metrics studied by Srivastava et al. [27], Mio et al. [23], the analogue for curves of the surface metric from Kurtek et al. [13, 14] and Younes et al. [31] as special cases.

Let us consider a general transform  $F : \text{Imm}(S^1, \mathbb{R}^2) \rightarrow C^\infty(S^1, \mathbb{R}^n)$  mapping plane curves to  $\mathbb{R}^n$ -valued functions for some  $n \in \mathbb{N}$ . On  $C^\infty(S^1, \mathbb{R}^n)$  we consider the following  $L^2$ -metric,

$$(6) \quad G_q^{L^2}(h, h) = \int_{S^1} |h(\theta)|^2 d\theta ,$$

for  $q \in C^\infty(S^1, \mathbb{R}^n)$ ,  $h \in T_q C^\infty(S^1, \mathbb{R}^n)$ . This is a particularly simple Riemannian metric: it does not depend on the basepoint  $q$ ; it is a flat metric in the sense of Riemannian geometry with geodesic distance given by the  $L^2$ -norm

$$\text{dist}^{L^2}(q_0, q_1)^2 = \int_{S^1} |q_0(\theta) - q_1(\theta)|^2 d\theta.$$

We will define the metric  $G^F$  on  $\text{Imm}(S^1, \mathbb{R}^2)$  to be the pullback of the  $L^2$ -metric under the transform  $F$ :

$$G_c^F(h, h) = G_{F(c)}^{L^2}(D_{c,h}F, D_{c,h}F) = \int_{S^1} |D_{c,h}F|^2 d\theta.$$

To obtain a reparameterisation invariant metric  $G^F$  we need the transform  $F$  to satisfy an equivariance property.

**Theorem 3.1.** *If  $F : \text{Imm}(S^1, \mathbb{R}^2) \rightarrow C^\infty(S^1, \mathbb{R}^n)$  satisfies*

$$(7) \quad F(c \circ \varphi) = \sqrt{\varphi'} F(c) \circ \varphi,$$

*with  $c \in \text{Imm}(S^1, \mathbb{R}^2)$  and  $\varphi \in \text{Diff}(S^1)$  and if  $F$  is infinitesimally injective, i.e.,  $D_{c,\cdot}F$  is injective for all  $c$ , then  $G^F$  is a Riemannian metric on  $\text{Imm}(S^1, \mathbb{R}^2)$  that is invariant under the reparameterization group  $\text{Diff}(S^1)$ .*

*Proof.* The infinitesimal version of the equivariance property is

$$D_{c \circ \varphi, h \circ \varphi} F = \sqrt{\varphi'} (D_{c,h} F) \circ \varphi$$

and thus we see that

$$\begin{aligned} G_{c \circ \varphi}^F(h \circ \varphi, h \circ \varphi) &= \int_{S^1} |D_{c \circ \varphi, h \circ \varphi} F|^2 d\theta \\ &= \int_{S^1} |(D_{c,h} F) \circ \varphi|^2 |\varphi'| d\theta = G_c^F(h, h), \end{aligned}$$

i.e., the metric  $G^F$  is invariant under  $\text{Diff}(S^1)$ .  $\square$

**Remark 3.2.** Examples of such transforms can be constructed in the following way: Take a smooth function  $f \in C^\infty(\mathbb{R}^{2m}, \mathbb{R}^n)$  and define the transform  $F$  as

$$(8) \quad F(c) = \sqrt{|c'|} f \circ (c, D_s c, \dots, D_s^{m-1} c).$$

To see that  $F$  indeed satisfies the required invariance property, note that

$$\sqrt{|(c \circ \varphi)'|} = \left( \sqrt{|c'|} \circ \varphi \right) \sqrt{\varphi'}$$

and that the arc-length derivative is equivariant with respect to reparameterizations. If we write  $D_c$  for the arc-length derivative to emphasize the dependence on the curve  $c$ , the equivariance property reads as  $D_{c \circ \varphi}(h \circ \varphi) = (D_c h) \circ \varphi$ . Thus

$$\begin{aligned} F(c \circ \varphi) &= \left( \sqrt{|c'|} \circ \varphi \right) \sqrt{\varphi'} f \circ (c \circ \varphi, (D_s c) \circ \varphi, \dots) \\ &= \sqrt{\varphi'} \left( \sqrt{|c'|} f \circ (c, D_s c, \dots) \right) \circ \varphi. \end{aligned}$$

The order of the induced metric will correspond to the order of the highest derivative appearing in the definition of the transform  $F$ .

**Remark 3.3.** Often the transform  $F$  will not be infinitesimally injective, but will have a kernel that consists of the constant vector fields. In this case we can regard  $F$  as an infinitesimally injective map  $F : \text{Imm}(S^1, \mathbb{R}^2) / \text{Tra} \rightarrow C^\infty(S^1, \mathbb{R}^n)$  and it induces a Riemannian metric on the space  $\text{Imm}(S^1, \mathbb{R}^2) / \text{Tra}$  of parameterized curves modulo translations.

We note that the same construction can also be applied for metrics on the space  $\text{Imm}([0, 2\pi], \mathbb{R}^2)$  of open curves.

**3.4. The square root velocity transform.** Our first example is the so-called square root velocity transform (SRVT), as introduced in [27]:

$$\begin{aligned} R : \text{Imm}(S^1, \mathbb{R}^2) / \text{Tra} &\rightarrow C^\infty(S^1, \mathbb{R}^2) \\ R(c) &= |c'|^{1/2} v . \end{aligned}$$

The metric induced by this transform is

$$G_c^{1,1/2}(h, h) = \int_{S^1} \langle D_s h, n \rangle^2 + \frac{1}{4} \langle D_s h, v \rangle^2 ds ,$$

which is a special case of the family in (1) of elastic metrics. The transform  $R$  can be written in the form (8) with the function  $f : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ,  $f(x_1, x_2) = x_2$ . To see this, note that  $v = D_s c$ . We will consider a generalization of this transform that allows us to represent metrics of the family in (1) for arbitrary parameters  $a, b$  in Section 4.

**3.5. The transform of Younes et al.** [31]. The method applied in Younes et al. [31] to study a Sobolev-type metric on curves also fits within the setup described in this paper. The basic mapping considered in [31] is given by

$$\begin{aligned} \Phi : C^\infty([0, 2\pi], \mathbb{R}^2 \setminus 0) &\rightarrow \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra} \\ \Phi(q)(\theta) &= \frac{1}{2} \int_0^\theta \begin{pmatrix} q_1^2(u) - q_2^2(u) \\ 2q_1(u)q_2(u) \end{pmatrix} du . \end{aligned}$$

It has the property that it pulls back the Sobolev metric

$$G_c(h, h) = \int_0^{2\pi} |D_s h|^2 ds$$

to the flat  $L^2$ -metric on the space  $C^\infty([0, 2\pi], \mathbb{R}^2 \setminus 0)$ . Here  $\text{Imm}([0, 2\pi], \mathbb{R}^2)$  denotes the manifold of open curves.

To fit this mapping into our framework we consider the inverse of  $\Phi$ :

$$\begin{aligned} \Phi^{-1} : \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra} &\rightarrow C^\infty([0, 2\pi], \mathbb{R}^2 \setminus 0) \\ \Phi^{-1}(c) &= \sqrt{|c'|} \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix} , \end{aligned}$$

with  $\alpha$  denoting the turning angle. If one wants to study closed curves instead of open curves, one needs to restrict the above mappings to the corresponding subspaces.

**Theorem 3.6** (From [31]). *The transform  $\Phi^{-1}$  induces the Sobolev metric of order one*

$$G_c(h, h) = \int_0^{2\pi} |D_s h|^2 ds$$

on the space of parameterized curves modulo translations, i.e., the  $a$ - $b$  metric with  $a = b = 1$ .

**3.7. The  $Q$ -transform.** Another member in this family of shape transformations is the  $Q$ -transform, which was introduced in Mani et al. [17] for curves and in Kurtek et al. [13, 14] for surfaces.

Following [17], let us define for a curve  $c \in \text{Imm}(S^1, \mathbb{R}^2)$  the transform

$$\begin{aligned} Q : \text{Imm}(S^1, \mathbb{R}^2) &\rightarrow C^\infty(S^1, \mathbb{R}^2) \\ Q(c) &= \sqrt{|c'|} c. \end{aligned}$$

This transform is a special case of the family in (8), with the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x) = x$ .

**Theorem 3.8.** *The  $Q$ -transform induces a reparameterization invariant metric on  $\text{Imm}(S^1, \mathbb{R}^2)$ , given by*

$$(9) \quad G_c(h, h) = \int_{S^1} \langle h + \frac{1}{2} \langle D_s h, v \rangle c, h + \frac{1}{2} \langle D_s h, v \rangle c \rangle ds.$$

**Remark 3.9.** Note that the  $Q$ -transform induces a metric on  $\text{Imm}(S^1, \mathbb{R}^2)$ , the space of all parameterized curves. This is in contrast to the other transforms discussed in this section, which induce metrics on the space  $\text{Imm}(S^1, \mathbb{R}^2)/\text{Tra}$  of curves modulo translations.

*Proof.* Using Lemma 2.3 we can compute the variation of  $Q$ . It is given by

$$(10) \quad D_{c,h}Q = \left(h + \frac{1}{2} \langle D_s h, v \rangle c\right) \sqrt{|c'|}.$$

This shows that the metric has the form (9).

It remains to show that the map  $h \mapsto D_{c,h}Q$  is injective for all  $c$ . Since we are only considering immersions that satisfy  $|c'| \neq 0$ , the condition  $D_{c,h}Q = 0$  is equivalent to

$$(11) \quad \langle h', c' \rangle c + 2|c'|^2 h = 0.$$

For all  $\theta$ , where  $c(\theta) = 0$ , this already implies  $h(\theta) = 0$ . On the open set  $\{\theta : c(\theta) \neq 0\}$  we expand the tangent vector  $h$  into a part along  $c$  and a part orthogonal to it,  $h = h_1 c + h_2 c^\perp$  with  $c \perp c^\perp$ . Then (11) can be rewritten as

$$\langle h', c' \rangle c + 2|c'|^2 h_1 c + 2|c'|^2 h_2 c^\perp = 0.$$

This implies that  $h_2 = 0$  everywhere and we are left with

$$(12) \quad \langle c, c' \rangle h_1' + 3|c'|^2 h_1 = 0.$$

As before, we solve for  $\theta$  satisfying  $\langle c(\theta), c'(\theta) \rangle = 0$  and find that  $h_1(\theta) = 0$ . It remains to study the open set, where neither  $c(\theta)$  nor  $\langle c(\theta), c'(\theta) \rangle$  vanish. This open set is the union of disjoint open intervals. Denote one such interval by  $(\theta_0, \theta_1)$ . On this interval every solution of (12) is given by

$$h_1(\theta) = C e^{\int_{\tilde{\theta}}^{\theta} \frac{-3|c'|^2}{\langle c, c' \rangle} d\theta},$$

with  $\tilde{\theta} \in (\theta_0, \theta_1)$ . We are looking for a smooth solution  $h_1$  on all of  $S^1$  and we already know that for  $\theta_0$  and  $\theta_1$  the solution has to satisfy  $h_1(\theta_0) = h_1(\theta_1) = 0$ , because these points lie outside the open set. Thus we see that only the solution with  $C = 0$  can satisfy this. Therefore  $h_1 \equiv 0$  and with it  $h \equiv 0$  on all of  $S^1$ .  $\square$



The injectivity of the map  $D_c.Q$  is essential in order to use the metric for shape comparisons. However, we could not find a proof for this anywhere in the literature.

Questions that have a comparably easy answer for the family of  $R$ -transforms are much more difficult for the  $Q$ -transform. To our knowledge it is unknown whether  $Q$  itself is injective; similarly little is known about its image on either open or closed curves, e.g., whether the image is an open subset or a smooth submanifold of the space  $C^\infty(S^1, \mathbb{R}^2)$ . See Sections 4 and 5 for an answer to these questions in the case of the  $R$ -transform. Even finding a numerically efficient way to invert the  $Q$ -transform presents difficulties, mainly because the  $Q$ -transform scales the curve  $c$  with the object  $\sqrt{|c'|}$ , which is geometrically the square-root of the volume element on the curve and multiplies tangent vectors along  $c$ , not the curve itself.

**3.10. An  $H^2$ -type metric.** We can also use this method to construct higher order metrics, such as the following transform, which induces a second order Sobolev-type metric:

$$K : \text{Imm}(S^1, \mathbb{R}^2) / \text{Tra} \rightarrow C^\infty(S^1, \mathbb{R}^3)$$

$$K(c) = \sqrt{|c'|}(v, \kappa).$$

**Theorem 3.11.** *The transform  $K$  induces a reparameterization invariant metric on  $\text{Imm}(S^1, \mathbb{R}^2) / \text{Tra}$ . It is given by*

$$G_c(h, h) = \int_{S^1} \langle D_s^2 h, n \rangle^2 - 3\kappa \langle D_s^2 h, n \rangle \langle D_s h, v \rangle$$

$$+ \langle D_s h, n \rangle^2 + \frac{1}{4}(1 + 9\kappa^2) \langle D_s h, v \rangle^2 ds.$$

*Proof.* To calculate the formula of the pullback metric we compute the variation of  $K$ . Using Lemma 2.3 we obtain:

$$D_{c,h}K = \sqrt{|c'|} \left( \frac{1}{2} \langle D_s h, v \rangle v + \langle D_s h, n \rangle n \right),$$

from which we can deduce the formula for the pullback metric. The injectivity of the map  $K$  and its derivative is clear, since one can reconstruct the immersion up to translations from the first two components of  $K$ , c.f., Section 4.4.  $\square$

Note that the transform  $K$  is not surjective. In fact, since it maps curves to  $\mathbb{R}^3$ -valued functions, the image will have infinite codimension. This is in contrast to the transform considered in Section 4, whose image is open on the space of open curves and has finite codimension for closed curves. This is the price that has to be paid for increasing the order of the metric. Second order metrics, that are induced in a similar way, are studied in [3].

**3.12. The elastic metric in Mio et al. [23].** The representation of the elastic metric used in [23] can also be rephrased in the spirit of representing metrics as pull-backs of simpler metrics. Define on  $\mathbb{R}^2$  the Riemannian metric

$$g_{(x,y)} = a^2 e^x dx^2 + b^2 e^x dy^2$$

and consider parallel to (6) the  $L^2$ -metric on the space  $C^\infty(S^1, (\mathbb{R}^2, g))$ ,

$$(13) \quad G_q^{L^2(g)}(h, h) = \int_{S^1} g_{q(\theta)}(h(\theta), h(\theta)) d\theta.$$

Then the transform

$$\begin{aligned} R : \text{Imm}(S^1, \mathbb{R}^2) / \text{Tra} &\rightarrow C^\infty(S^1, (\mathbb{R}^2, g)) \\ R(c) &= (\log |c'|, \alpha) \end{aligned}$$

pulls back the  $L^2$ -metric to the elastic metric (1). While in this case the metric (13) is not as simple as the  $L^2$ -metric in (6), it is still simpler than the elastic metric, since it does not contain spatial derivatives.

#### 4. THE $R$ -TRANSFORM FOR OPEN CURVES

We want to generalize the SRVT, as defined in Section 3.4, to study the family of  $a$ - $b$  metrics (1) for arbitrary parameters  $a$  and  $b$ . We will start by studying the metric on the set of open curves  $\text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra}$ , and then investigate the closedness conditions.

Define the  $R$ -transform of a plane curve  $c \in \text{Imm}([0, 2\pi], \mathbb{R}^2)$  by

$$\begin{aligned} R^{a,b} : \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra} &\rightarrow C^\infty([0, 2\pi], \mathbb{R}^3) \\ R^{a,b}(c) &= |c'|^{1/2} \left( a \begin{pmatrix} v \\ 0 \end{pmatrix} + \sqrt{4b^2 - a^2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

Here  $a, b \in \mathbb{R}^+$  are positive numbers with  $4b^2 \geq a^2$ . We will omit the parameters  $a, b$  when the meaning is unambiguous. The  $R$ -transform maps an open plane curve to a space curve. We will see that equipping the space  $C^\infty([0, 2\pi], \mathbb{R}^3)$  with a flat  $L^2$ -metric, i.e., considering the vector space  $C^\infty([0, 2\pi], \mathbb{R}^3)$  with the  $L^2$ -inner product as a Riemannian manifold, will generate Sobolev metrics of order one via the pullback by the  $R$ -transform. For the choice of parameters  $4b^2 = a^2$  the  $R$ -transform reduces to the SRVT of Section 3.4, as studied in [27]. In the following theorem we will show that the metrics induced by the  $R^{a,b}$ -transform coincide with the family of  $a$ - $b$  metrics in (1) as introduced in [23].

First let us note the following properties of the  $R$ -transform.

**Lemma 4.1.** *For a curve  $c \in \text{Imm}([0, 2\pi], \mathbb{R}^2)$  and  $\varphi \in \text{Diff}([0, 2\pi])$  a diffeomorphism we have:*

- *Equivariance under reparametrizations*

$$R(c \circ \varphi) = |\varphi'|^{1/2} \cdot (R(c) \circ \varphi)$$

- *Translation invariance*

$$R(c + p) = R(c) \text{ for } p \in \mathbb{R}^2$$

- *Scaling property*

$$R(\rho \cdot c) = \rho^{1/2} \cdot R(c) \text{ for } \rho \in \mathbb{R}_{>0}$$

- *Preservation of the  $L^2$ -norm under reparametrizations*

$$\int_0^{2\pi} |R(c \circ \varphi)|^2 d\theta = \int_{S^1} |R(c)|^2 d\theta$$

*Proof.* The first property follows from Remark 3.2. The other properties can be verified by simple calculations.  $\square$

**Theorem 4.2** (The pullback metric on  $\text{Imm}/\text{Tra}$ ). *The pullback of the  $L^2$ -inner product on  $C^\infty([0, 2\pi], \mathbb{R}^3)$  to the manifold of immersions,  $\text{Imm}([0, 2\pi], \mathbb{R}^2)$ , by the  $R^{a,b}$ -transform yields the family of reparameterization invariant  $a$ - $b$  metrics on the space  $\text{Imm}([0, 2\pi], \mathbb{R}^2)/\text{Tra}$ :*

$$G_c^{a,b}(h, h) = \int_0^{2\pi} a^2 \langle D_s h, n \rangle^2 + b^2 \langle D_s h, v \rangle^2 ds.$$

*Proof.* The pullback metric is defined via

$$G_c^{a,b}(h, h) = \langle D_{c,h} R, D_{c,h} R \rangle_{L^2},$$

and hence we need to compute the variation of the  $R$ -transform. Using the formulas for the derivatives of the functions  $c \mapsto |c'|^{1/2}$  and  $c \mapsto v$  from Lemma 2.3 we obtain

$$D_{c,h} R = |c'|^{1/2} \left( a \langle D_s h, n \rangle n + \frac{a}{2} \langle D_s h, v \rangle v \right).$$

Therefore the pullback metric is

$$\begin{aligned} G_c^{a,b}(h, h) &= \int_0^{2\pi} \left( \left| a \langle D_s h, n \rangle n + \frac{a}{2} \langle D_s h, v \rangle v \right|^2 + \left( b^2 - \frac{a^2}{4} \right) \langle D_s h, v \rangle^2 \right) |c'| d\theta \\ &= \int_0^{2\pi} a^2 \langle D_s h, n \rangle^2 + b^2 \langle D_s h, v \rangle^2 ds. \end{aligned}$$

Lemma 4.1 tells us that the metric  $G^{a,b}$  is invariant under reparameterizations. This can also be seen directly, since the metric is written only in terms of operations  $D_s, v, n$ , which are equivariant with respect to reparameterizations. The kernel of the transform consists only of constant vector fields, since one can reconstruct  $D_s h$  from  $D_{c,h} R^{a,b}$ . Thus  $R^{a,b}$  induces a metric on  $\text{Imm}(S^1, \mathbb{R}^2)/\text{Tra}$ .  $\square$

**4.3. Geodesic Equation.** This family of metrics can be also written in the form

$$G_c^{a,b}(h, h) = \int_0^{2\pi} \langle P_c^{a,b}(h), h \rangle ds$$

where for each  $c$  the associated pseudo-differential operator  $P_c^{a,b}$  is given by

$$\begin{aligned} (14) \quad P_c^{a,b}(h) &= -D_s (a^2 \langle D_s^2 h, n \rangle n + b^2 \langle D_s^2 h, v \rangle v) \\ &\quad + (\delta_{2\pi} - \delta_0) (a^2 \langle n, D_s h \rangle n + b^2 \langle v, D_s h \rangle v). \end{aligned}$$

Here  $\delta$  denotes the delta distribution.

If we identify the space of immersions modulo translations with the section

$$\text{Imm}([0, 2\pi], \mathbb{R}^2)/\text{Tra} \cong \{c : c(0) = 0\},$$

then the geodesic equation for the  $G^{a,b}$ -metric is

$$\begin{aligned} (15) \quad D_s \left( (A_c c_t)_t + \frac{1}{2} B_c(c_t, c_t) \right) &= 0 \\ (A_c c_t)_t(2\pi) + \frac{1}{2} B_c(c_t, c_t)(2\pi) &= 0, \end{aligned}$$

with

$$\begin{aligned} A_c h &= a^2 \langle D_s h, n \rangle n + b^2 \langle D_s h, v \rangle v \\ B_c(h, h) &= (a^2 \langle D_s h, n \rangle^2 + b^2 \langle D_s h, v \rangle^2) v - 2(b^2 - a^2) \langle D_s h, n \rangle \langle D_s h, v \rangle n. \end{aligned}$$

We can then integrate the equation once and use the boundary conditions to determine the constant of integration. The equation we obtain is

$$(A_c c_t)_t + \frac{1}{2} B_c(c_t, c_t) = 0.$$

On the space of closed curves, the associated operator is the same as in (14), except that the terms involving the delta-distribution disappear. The geodesic equation is the same as (15) minus the boundary conditions.

From the results of Sect. 4.4 we can deduce an explicit solution formula for the geodesic equation on the space of open curves. Given a curve  $c_0$  and an initial velocity  $u_0$ , the solution of the geodesic equation for small time is given by

$$c(t, \theta) = (R^{a,b})^{-1} (R^{a,b}(c_0) + t D_{c_0, u_0} R^{a,b}).$$

Even though the geodesic equation on the space of closed curves is simpler, there exists no explicit solution formula for it; see Section 7 for numerical computations of geodesics on closed curves.

**4.4. Image of the  $R$ -transform.** To characterize the image of the  $R$ -transform we note that we can reconstruct  $c'$  from the first two components of  $R(c)$  via

$$c' = \frac{1}{a^2} \sqrt{R_1^2(c) + R_2^2(c)} \begin{pmatrix} R_1(c) \\ R_2(c) \end{pmatrix}$$

and hence  $a|c'|^{1/2} = \sqrt{R_1^2(c) + R_2^2(c)}$ . This implies that

$$R_3(c) = \frac{\sqrt{4b^2 - a^2}}{a} \sqrt{R_1^2(c) + R_2^2(c)},$$

which can be written in the form

$$(16) \quad (4b^2 - a^2) (R_1(c)^2 + R_2(c)^2) = a^2 R_3(c)^2.$$

Let us define the cone

$$C^{a,b} = \{q \in \mathbb{R}^3 : (4b^2 - a^2)(q_1^2 + q_2^2) = a^2 q_3^2, q_3 > 0\}$$

in  $\mathbb{R}^3$ . Then the image of the  $R$ -transform consists of curves, that lie in  $C^{a,b}$ . That is, for a curve  $q \in C^\infty([0, 2\pi], \mathbb{R}^3)$  we have

$$q \in \text{im } R \iff q(\theta) \in C^{a,b}, \forall \theta \in [0, 2\pi],$$

$$\text{im } R = C^\infty([0, 2\pi], C^{a,b}).$$

In the special case  $4b^2 = a^2$  the  $R^{2a,a}$ -transform has no third component and the image of the  $R^{2a,a}$ -transform is the open set consisting of all curves which avoid  $(0, 0) \in \mathbb{R}^2$ . The latter condition arises from the requirement that the curves be regular.

The inverse of the  $R$ -transform can be computed using the identity

$$c' = \frac{1}{2ab} |R| \begin{pmatrix} R_1(c) \\ R_2(c) \end{pmatrix}.$$

Therefore

$$R^{-1} : \text{im } R \rightarrow \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra}$$

$$R^{-1}(q)(\theta) = p_0 + \frac{1}{2ab} \int_0^\theta |q(\theta)| \begin{pmatrix} q_1(\theta) \\ q_2(\theta) \end{pmatrix} d\theta.$$

Of course, the inverse of the  $R$ -transform is defined only up to translation, which manifests itself as the freedom to choose the starting point  $p_0 \in \mathbb{R}^2$  of the integration.

The cone  $C^{a,b}$  is a flat hypersurface in  $\mathbb{R}^3$ , since there is an isometric covering map from the polar coordinate domain  $\{(r, \varphi) : r > 0\}$  to  $C^{a,b}$  given by

$$q(r, \varphi) = \left( \frac{r}{m} \cos(m\varphi), \frac{r}{m} \sin(m\varphi), \frac{\sqrt{4b^2 - a^2}}{2b} r \right),$$

with  $m = \frac{2b}{a}$ . The computation

$$(4b^2 - a^2)(q_1^2 + q_2^2) = \frac{4b^2 - a^2}{m^2} r^2 = a^2 q_3^2,$$

checks that the image of this map is the cone  $C^{a,b}$ . From

$$dq_1^2 + dq_2^2 + dq_3^2 = \frac{1}{m^2} dr^2 + r^2 d\varphi^2 + \frac{4b^2 - a^2}{4b^2} dr^2 = dr^2 + r^2 d\varphi^2$$

we see that the map  $q(r, \varphi)$  is an isometry from the Euclidean metric in  $\mathbb{R}^2$ , which has the expression  $dr^2 + r^2 d\varphi^2$  in polar coordinates, to the natural metric on the cone  $C^{a,b}$ .

The inverse map is determined only up to a multiple of  $2\pi$  and is given by

$$\Phi(q) = \begin{pmatrix} r(q) = \frac{2b}{\sqrt{4b^2 - a^2}} q_3 \\ \varphi(q) = \frac{a}{2b} \left( \arctan\left(\frac{q_2}{q_1}\right) + 2k\pi \right) \end{pmatrix}$$

with  $k \in \mathbb{Z}$ . Using the inverse map we can write the distance function on the cone

$$\begin{aligned} (17) \quad \text{dist}^2(q, \bar{q}) &= (x_1(q) - x_1(\bar{q}))^2 + (x_2(q) - x_2(\bar{q}))^2 \\ &= r(q)^2 + r(\bar{q})^2 - 2r(q)r(\bar{q}) \cos(\varphi(q) - \varphi(\bar{q})) \\ &= \min_{k \in \mathbb{Z}} \frac{4b^2}{4b^2 - a^2} \left( q_3^2 + \bar{q}_3^2 - \right. \\ &\quad \left. - 2q_3\bar{q}_3 \cos\left(\frac{a}{2b} \left( \arctan\frac{q_2}{q_1} - \arctan\frac{\bar{q}_2}{\bar{q}_1} \right) + \frac{a}{b} k\pi \right) \right). \end{aligned}$$

The minimum appears because the angle is only determined up to a multiple of  $2\pi$ .

**Theorem 4.5.** *The metric  $G^{a,b}$  on open curves is flat. Geodesics are the preimages under the  $R$ -transform of geodesics on the flat space  $\text{im } R$ .*

*A path of curves  $q : \mathbb{R} \times [0, 2\pi] \rightarrow C^{a,b}$  in  $\text{im } R$  is a geodesic if for each  $\theta \in [0, 2\pi]$  the curve  $t \mapsto q(t, \theta)$  is a geodesic in  $C^{a,b}$ .*

*The geodesic distance between  $c, \bar{c} \in \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra}$  is given by the integral over the pointwise distances,*

$$\int_0^{2\pi} \text{dist}(R(c)(\theta), R(\bar{c})(\theta)) d\theta.$$

*However, the minimum over  $k \in \mathbb{Z}$  is not to be taken pointwise, but only once for all values of  $\theta$ . This corresponds to choosing a continuous lift of the curve  $R(c)$  via  $\Phi$ .*

*Proof.* Since the cone  $C^{a,b}$  is flat in the sense of Riemannian geometry, so is the space  $C^\infty([0, 2\pi], C^{a,b})$  of curves in  $C^{a,b}$  with respect to the  $L^2$ -metric given by

$$G_q(h, k) = \int_0^{2\pi} \langle h, k \rangle d\theta,$$

for  $q \in C^\infty([0, 2\pi], C^{a,b})$  and  $h, k$  tangent vectors at  $q$ . Note that this metric does not depend on the basepoint  $q$ . The metric is the same at all points in the space. Note also that the image of the  $R$ -transform

$$\text{im } R = C^\infty([0, 2\pi], C^{a,b}),$$

equals the set of curves that lie in the cone  $C^{a,b}$ . It is a property of the  $L^2$ -metric that geodesics in  $\text{im } R$  are given by paths of curves  $q(t, \theta)$ , such that for each fixed  $\theta$ , the curve  $q(\cdot, \theta)$  is a geodesic in the cone  $C^{a,b}$ . The length of this geodesic will be given by an expression of the form (17) with some  $k \in \mathbb{Z}$ , not necessarily the smallest one. The length of the path  $q(t, \theta)$  in  $\text{im } R$  is given by the integral over the lengths of each point-wise path  $q(\cdot, \theta)$ . Since we have a continuous family of geodesics on  $C^{a,b}$ , the value for  $k$  will be the same for all  $\theta \in [0, 2\pi]$ . Hence, the geodesic distance between two elements  $q, \bar{q} \in \text{im } R$  is given by the integral

$$\int_0^{2\pi} \text{dist}(q(\theta), \bar{q}(\theta)) d\theta,$$

where the minimum over  $k \in \mathbb{Z}$  is taken only once for all values of  $\theta$ .  $\square$

See [10, Theorem 9.1] for more details on the flat  $L^2$ -metric.

## 5. THE $R$ -TRANSFORM FOR CLOSED CURVES

In this section we want to consider the  $R$ -transform acting on closed curves. First note that Theorem 4.2 remains valid, if we replace open curves by closed ones.

Restricting our attention to closed curves  $C^\infty(S^1, \mathbb{R}^2)$  means that we impose additional constraints on the image of the  $R$ -transform. From the inversion formula

$$R^{-1}(q)(\theta) = \frac{1}{a^2} \int_0^\theta \sqrt{q_1(\theta)^2 + q_2(\theta)^2} \begin{pmatrix} q_1(\theta) \\ q_2(\theta) \end{pmatrix} d\theta,$$

we see that a curve  $q$  is the image of a closed curve only if the condition

$$F(q) = \int_0^\theta \sqrt{q_1(\theta)^2 + q_2(\theta)^2} \begin{pmatrix} q_1(\theta) \\ q_2(\theta) \end{pmatrix} d\theta = 0$$

is satisfied. Let us denote the image of the  $R$ -transform, restricted to closed curves, by

$$\mathcal{C}^{a,b} = \{q \in C^\infty(S^1, C^{a,b}) : F(q) = 0\}.$$

Note the difference between the cone  $C^{a,b}$ , which is a submanifold of  $\mathbb{R}^3$  and the space  $\mathcal{C}^{a,b}$ , which is a submanifold in the space of curves.

**Theorem 5.1.** *The image  $\mathcal{C}^{a,b}$  of the manifold of closed curves under the  $R$ -transform is a codimension 2 submanifold of the flat space  $C^\infty(S^1, C^{a,b})$ .*

A basis of the orthogonal complement  $(T_q \mathcal{C}^{a,b})^\perp$  is given by the two vectors

$$\begin{aligned} U_1(q) &= \frac{1}{\sqrt{q_1^2 + q_2^2}} \begin{pmatrix} 2q_1^2 + q_2^2 \\ q_1 q_2 \\ 0 \end{pmatrix} + \frac{2}{a} \sqrt{4b^2 - a^2} \begin{pmatrix} 0 \\ 0 \\ q_1 \end{pmatrix}, \\ U_2(q) &= \frac{1}{\sqrt{q_1^2 + q_2^2}} \begin{pmatrix} q_1 q_2 \\ q_1^2 + 2q_2^2 \\ 0 \end{pmatrix} + \frac{2}{a} \sqrt{4b^2 - a^2} \begin{pmatrix} 0 \\ 0 \\ q_2 \end{pmatrix}. \end{aligned}$$

*Proof.* A basis of  $(T_q \mathcal{C}^{a,b})^\perp$  can be computed by projecting the gradients of the two components of the function  $F = (F_1, F_2)$  to the tangent space of  $C^\infty(S^1, C^{a,b})$ . Let  $q \in \mathcal{C}^{a,b}$  be a curve and  $h \in C^\infty(S^1, \mathbb{R}^3)$  a tangent vector. Then

$$\begin{aligned} D_{q,h} F_1 &= \int_0^{2\pi} \frac{q_1 h_1 + q_2 h_2}{\sqrt{q_1^2 + q_2^2}} q_1 + \sqrt{q_1^2 + q_2^2} h_1 d\theta \\ D_{q,h} F_2 &= \int_0^{2\pi} \frac{q_1 h_1 + q_2 h_2}{\sqrt{q_1^2 + q_2^2}} q_2 + \sqrt{q_1^2 + q_2^2} h_2 d\theta. \end{aligned}$$

Thus the two gradients are

$$\begin{aligned} \text{grad}^{L^2} F_1(q) &= \frac{q_1}{\sqrt{q_1^2 + q_2^2}} \begin{pmatrix} q_1 \\ q_2 \\ 0 \end{pmatrix} + \sqrt{q_1^2 + q_2^2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\ \text{grad}^{L^2} F_2(q) &= \frac{q_2}{\sqrt{q_1^2 + q_2^2}} \begin{pmatrix} q_1 \\ q_2 \\ 0 \end{pmatrix} + \sqrt{q_1^2 + q_2^2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Differentiating the governing equation for the cone (16) we obtain that the tangent space  $T_q C^\infty(S^1, C^{a,b})$  is given by all curves  $h$ , which satisfy the pointwise condition

$$(18) \quad (4b^2 - a^2)(q_1 h_1 + q_2 h_2) = a^2 q_3 h_3.$$

A projection (not necessarily orthogonal) of the vectors  $\text{grad}^{L^2} F_i(q)$  can be found by choosing  $\lambda_i$  such that the vector  $\text{grad}^{L^2} F_i(q) + \lambda_i(0, 0, 1)^T$  satisfies (18). A simple computation shows that for the component  $F_i$  the right value is  $\lambda_i = 2q_i \frac{\sqrt{4b^2 - a^2}}{a}$ . This completes the proof.  $\square$

We can now use the Gram-Schmidt procedure to compute an orthonormal basis of  $(T_q \mathcal{C}^{a,b})^\perp$ , which is better suited for computational purposes. The formulas however do not reveal more structure.

**5.2. The curvature of the space  $\mathcal{C}^{2a,a}$ .** The curvature of a Riemannian manifold is a symmetric bilinear form on the space of skew bivectors. The normalized quadratic version is called sectional curvature, which in turn is the Gauß curvature of the geodesic 2-submanifold spanned by the bivector.

As mentioned previously the  $R$ -transform has no third component in the case  $b = 2a$  and the cone  $C^{2a,a}$  reduces to the space  $\mathbb{R}^2 \setminus 0$ . Let  $\widetilde{U}_1(q), \widetilde{U}_2(q) \in C^\infty(S^1, \mathbb{R}^2)$  be the orthonormal basis of  $(T_q \widetilde{\mathcal{C}}^{2a,a})^\perp$  derived via the Gram-Schmidt procedure from the basis  $U_1(q), U_2(q)$ . Using this ONB we can express the curvature of the cone  $\widetilde{\mathcal{C}}^{2a,a}$ .

To do so we take a constant vector field  $q \mapsto (q, h)$  on  $C^\infty(S^1, \mathbb{R}^2)$  and its orthonormal projection

$$X_h(q) = h - \langle \widetilde{U}_1(q), h \rangle \widetilde{U}_1(q) - \langle \widetilde{U}_2(q), h \rangle \widetilde{U}_2(q) \in T_q \mathcal{C}^{2a,a}.$$

Then we take the flat covariant derivative in  $\text{Imm}(S^1, \mathbb{R}^2)$

$$\overline{\nabla}_{X_h(q)} X_k(q) = D_{q, X_h(q)} X_k(q) = d(X_k(q))(X_h(q)).$$

The orthonormal projection of this vector field onto  $(T_q \mathcal{C}^{2a,a})^\perp$  is then equal to the value of the second fundamental form  $S \in (T_q \mathcal{C}^{2a,a})^\perp$ , i.e.:

$$\begin{aligned} S(X_h(q), X_k(q)) &= \langle d(X_k(q))(X_h(q)), \widetilde{U}_1(q) \rangle_{L^2} \widetilde{U}_1(q) \\ &\quad + \langle d(X_k(q))(X_h(q)), \widetilde{U}_2(q) \rangle_{L^2} \widetilde{U}_2(q). \end{aligned}$$

The curvature of  $\mathcal{C}^{2a,a}$  at  $q$  is then given by the Gauß-equation [19, 26.4]:

$$\begin{aligned} \langle R(X_h(q), X_k(q))X_k(q), X_h(q) \rangle_{L^2} \\ = -\|S(X_h(q), X_k(q))\|_{L^2}^2 + \langle S(X_k(q), X_k(q)), S(X_h(q), X_h(q)) \rangle_{L^2}. \end{aligned}$$

## 6. THE INDUCED METRIC ON SHAPE SPACE $\mathcal{S}$

In the rest of the paper we will mainly consider closed curves. However the results can be easily reformulated for the case of open curves.

The shape space  $\mathcal{S}$  denotes the space of unparameterized plane curves, which can be represented as the quotient  $\mathcal{S} := \text{Imm}(S^1, \mathbb{R}^2)/\text{Diff}(S^1)$  of parameterized curves modulo parameterizations. Associated to this quotient is the natural projection  $\pi : \text{Imm}(S^1, \mathbb{R}^2) \rightarrow \text{Imm}(S^1, \mathbb{R}^2)/\text{Diff}(S^1)$ , which maps a curve  $c$  to its image  $C = \pi(c)$ .

Given a reparameterization invariant metric  $G$  on  $\text{Imm}(S^1, \mathbb{R}^2)$  there exists a unique Riemannian metric  $\overline{G}$  on shape space, such that the projection  $\pi$  is a Riemannian submersion. Associated to the projection  $\pi$  is the decomposition of the tangent bundle  $T\text{Imm}(S^1, \mathbb{R}^2)$  into horizontal and vertical parts. The vertical bundle  $\text{Ver}$  is the kernel of the projection  $\pi$ , i.e.,  $\text{Ver} = \ker T\pi$ , and the horizontal bundle  $\text{Hor}(c) = \text{Ver}(c)^\perp \subset T_c \text{Imm}(S^1, \mathbb{R}^2)$  is defined as the orthogonal complement of  $\text{Ver}$  with respect to the Riemannian metric  $G$ . The action of  $\text{Diff}(S^1)$  on  $\text{Imm}(S^1, \mathbb{R}^2)$  induces an infinitesimal action of its Lie algebra  $\mathfrak{X}(S^1)$ , given by

$$\zeta_\mu(c) = c' \mu \in T_c \text{Imm}(S^1, \mathbb{R}^2)$$

with  $\mu \in \mathfrak{X}(S^1)$ . This defines a vector field  $\zeta_\mu$  on the space  $\text{Imm}(S^1, \mathbb{R}^2)$ . The vertical bundle consists of the image of all infinitesimal vector fields

$$\text{Ver}(c) = \{ \zeta_\mu(c) : \mu \in \mathfrak{X}(S^1) \}.$$

When we fix a curve  $c$ , there is a one-to-one correspondence between the vector fields on the circle  $\mathfrak{X}(S^1)$  and the space  $\text{Ver}(c)$  given by the map  $\mu \mapsto \zeta_\mu(c)$ .

From the theory of Riemannian submersions [19] it follows that:

- Geodesics on shape space  $\mathcal{S}$  with respect to  $\overline{G}$  correspond to horizontal geodesics on the manifold  $\text{Imm}(S^1, \mathbb{R}^2)$  of parameterized curves with respect to  $G^{a,b}$ . Horizontal geodesics on  $\text{Imm}(S^1, \mathbb{R}^2)$  are those, whose tangent vector lies in the horizontal bundle, i.e.,  $\partial_t c(t) \in \text{Hor}(c(t))$ .



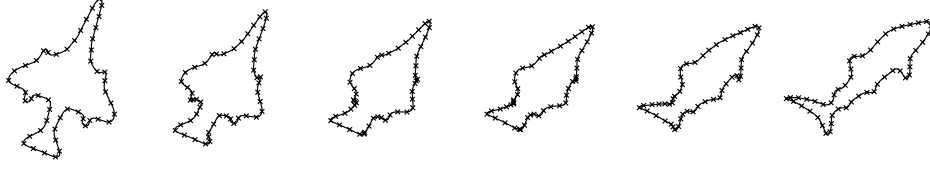


FIGURE 1. Example of a minimal geodesic between two parameterized shapes.

- The geodesic distance on shape space can be computed using the formula

$$\text{dist}(C_0, C_1) = \inf_{\varphi \in \text{Diff}(S^1)} \text{dist}(c_0, c_1 \circ \varphi),$$

where  $\text{dist}$  on the right hand side denotes the geodesic distance on the space  $\text{Imm}(S^1, \mathbb{R}^2)$  with respect to the  $G^{a,b}$ -metric.

- The curvature of the shape space can be calculated using O’Neil’s curvature formula, see for example [19]. Given two orthonormal vector fields  $X, Y$  on the space  $\mathcal{S}$  of unparameterized curves, the sectional curvature  $K$  is given by

$$K_S(X, Y) = K_{\text{Imm}}(\tilde{X}, \tilde{Y}) + \frac{3}{4} |[\tilde{X}, \tilde{Y}]^{\text{vert}}|^2.$$

Here  $\tilde{X}, \tilde{Y}$  are horizontal lifts of the vector fields  $X, Y$  to  $\text{Imm}(S^1, \mathbb{R}^2)$  and  $[\tilde{X}, \tilde{Y}]^{\text{vert}}$  denotes the vertical projection of the vector field  $[\tilde{X}, \tilde{Y}]$ .

We can now apply this construction to the metric  $G^{a,b}$  and the projection  $\pi : \text{Imm}(S^1, \mathbb{R}^2) / \text{Tra} \rightarrow \mathcal{S} / \text{Tra}$

**Theorem 6.1.** *Consider the shape space of open curves modulo translations with the metric that is induced from the  $G^{a,b}$ -metric on parameterized curves. The curvature of this space is non-negative.*

*Proof.* Theorem 4.5 shows that the space  $\text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra}$  of parameterized curves is flat, which implies  $K_{\text{Imm}}(\tilde{X}, \tilde{Y}) = 0$  in O’Neil’s curvature formula. The only remaining term is clearly non-negative:  $K(X, Y) = \frac{3}{4} |[\tilde{X}, \tilde{Y}]^{\text{vert}}|^2$ .  $\square$

## 7. NUMERICAL COMPUTATION OF GEODESICS FOR PARAMETERIZED CURVES

In this section we will describe a way to numerically compute the shortest path between two parameterized closed curves  $c_0$  and  $c_1$ . We will consider the special case of the metric where  $4b^2 = a^2 = 1$ . In this case the  $R$ -transform maps plane curves into plane curves,

$$R : \text{Imm}([0, 2\pi], \mathbb{R}^2) / \text{Tra} \rightarrow C^\infty([0, 2\pi], \mathbb{R}^2),$$

$$R(c) = |c'|^{1/2} v.$$

The cone  $C^{a,a/2}$  regarded as a subset of  $\mathbb{R}^2$  simplifies to  $C^{a,a/2} = \mathbb{R}^2 \setminus 0$ . Thus the image of the  $R$ -transform of open curves is the set  $\text{im } R = C^\infty([0, 2\pi], \mathbb{R}^2 \setminus 0)$  of all curves that avoid the origin, which is an open subset of the space of all curves. Restricting ourselves to closed curves, we obtain from Theorem 5.1 that the image  $\mathcal{C}$  of the  $R$ -transform is a codimension 2 submanifold of  $C^\infty(S^1, \mathbb{R}^2)$ .

Geodesics in the space  $\text{Imm}(S^1, \mathbb{R}^2)/\text{Tra}$  correspond under the  $R$ -transform to geodesics on  $\mathcal{C}$  with the Riemannian metric induced by the flat  $L^2$ -metric on  $C^\infty(S^1, \mathbb{R}^2)$ . Using Theorem 5.1 we can implement a projection operator which, given a curve  $q \in \mathcal{C}$  and a tangent vector  $p$  to  $q$ , computes its orthogonal projection onto  $T_q\mathcal{C}$ .

```

function PROJ( $q, p$ )
   $U_1, U_2 \leftarrow$  ONB of  $(T_q\mathcal{C})^\perp$ 
  return  $p - \langle p, U_1 \rangle U_1 - \langle p, U_2 \rangle U_2$ 
end function

```

The forward computation of the geodesic starting from a curve  $q \in \mathcal{C}$  with initial velocity  $p \in T_q\mathcal{C}$  can be seen as a constrained optimization problem. A geodesic on  $\mathcal{C}$  is a minimum of

$$\int_0^1 \int_{S^1} \frac{1}{2} |u(t)|^2 + \langle p(t), \dot{q}(t) - u(t) \rangle d\theta + \lambda(t) \cdot F(q(t)) dt .$$

Since we use the Euclidean metric on the ambient space, we will not distinguish between velocities  $u(t)$  and momenta  $p(t)$ . The variable  $\lambda(t) \in \mathbb{R}^2$  is a Lagrange multiplier, which enforces the constraint  $F(q) = 0$  with

$$(19) \quad F(q) = \int_{S^1} |q(\theta)| q(\theta) d\theta ,$$

and the cone  $\mathcal{C} = F^{-1}(0)$ . These are holonomic constraints and there exist several methods, that can solve this problem; see, e.g., [15] for an overview. We have chosen to use the RATTLE algorithm [1], which is a symplectic integrator that preserves the constraints through an explicit projection of the velocity vector, forcing it to lie on the constraint surface. The variational equations in continuous time have the form

$$\begin{aligned} \dot{q}(t) &= p(t) \\ \dot{p}(t) &= \lambda(t) \cdot \nabla F(q(t)) . \end{aligned}$$

We denote by  $N$  the number of time-steps used by the integrator and set  $\Delta t = \frac{1}{N}$ .

```

function EXP( $q, p, N$ )
   $q_0, p_0 \leftarrow q, p$ 
  for  $i \leftarrow 0, \dots, N-1$  do
     $\bar{p} \leftarrow p_i + \frac{\Delta t}{2} \lambda \cdot \nabla F(q_i)$ 
     $q_{i+1} \leftarrow q_i + \Delta t \bar{p}$ 
    estimate initial value for  $\lambda$  to enforce constraint
    iteratively adapt  $\lambda$  so that  $q_{i+1} \in \mathcal{C}$ 
     $p_{i+1} \leftarrow \bar{p} + \frac{\Delta t}{2} \mu \cdot \nabla F(q_{i+1})$ 
     $\mu$  chosen such that  $p_{i+1} \in T_{q_{i+1}}\mathcal{C}$ 
  end for
  return  $q, p$ 
end function

```

The next step is to solve the boundary value problem for parameterized shapes. Given two curves  $q_0, q_1 \in \mathcal{C}$  we need to find the initial velocity  $p$ , such that the endpoint  $\text{Exp}_{q_0} p$  is close to  $q_1$ . We do so via a fixed-point iteration, using the fact that our space  $\mathcal{C}$  is a submanifold of a flat space.

```

function LOG( $q_0, q_1, N$ )

```

```

 $p \leftarrow \Delta t \text{PROJ}(q_0, q_1 - q_0)$ 
 $\tilde{q} \leftarrow \text{EXP}(q_0, p, 1)$ 
 $p \leftarrow \text{PROJ}(q_0, \tilde{q} - q_0)$ 
while  $|\text{EXP}(q_0, p, N) - q_1| > \varepsilon$  do
     $\tilde{q} \leftarrow \text{EXP}(q_0, p, N)$ 
     $p \leftarrow p + \alpha \text{PROJ}(q_0, q_1 - \tilde{q})$ 
end while
return  $p$ 
end function
    
```

To construct the initial guess for  $p$  we project the straight line between  $q_0$  and  $q_1$  to  $T_{q_0}\mathcal{C}$ . This would be a valid initial guess. Based on experiments, however we found that computing one step of the forward-shooting algorithm to obtain  $\tilde{q}$  and projecting the straight line  $\tilde{q} - q_0$  back to  $T_{q_0}\mathcal{C}$  leads to a better initial guess. The iteration consists of computing the endpoint  $\tilde{q} = \text{Exp}_{q_0} p$  of the geodesic with initial velocity  $p$  and updating  $p$  in the direction  $q_1 - \tilde{q}$ .

The parameter  $\alpha$  controls the step-size of the iteration. It can be either fixed for the whole minimization or chosen adaptively. In our experiments we chose  $\alpha$  as large as possible, while still ensuring that the distance decreased, compared to the last iteration.

It is also possible to view this iteration as an approximation to a gradient descent for the geodesic distance

$$E(p) = \frac{1}{2} \text{dist}(\text{Exp}_{q_0} p, q_1)^2.$$

The derivative of this function is

$$T_p E(\delta p) = G_{\text{Exp}_{q_0} p} \left( -\text{Log}_{\text{Exp}_{q_0} p} q_1, T_p \text{Exp}_{q_0}(\delta p) \right).$$

If we approximate the logarithm by the straight line,  $\text{Log}_{\text{Exp}_{q_0} p} q_1 \approx q_1 - \text{Exp}_{q_0} p$ , the differential of the exponential map by the identity,  $T_p \text{Exp}_{q_0}(\delta p) \approx \delta p$  and the parallel transport along the geodesic  $t \mapsto \text{Exp}_{q_0}(tp)$  from  $\text{Exp}_{q_0} p$  to  $q_0$  by the projection to  $T_{q_0}\mathcal{C}$ , then we obtain

$$T_p E(\delta p) \approx G_{q_0} \left( -\text{Proj}_{q_0} (q_1 - \text{Exp}_{q_0} p), \delta p \right)$$

and hence the approximation of the gradient is given by

$$\nabla_p E \approx -\text{Proj}_{q_0} (q_1 - \text{Exp}_{q_0} p).$$

For the spatial discretization we replaced the curve  $q \in \mathcal{C}$  by a finite number of points  $(q^j)_{1 \leq j \leq n}$ . Spatially discrete geodesics correspond to minima of

$$\int_0^1 \sum_{j=1}^n \frac{1}{2} |w^j(t)|^2 + \langle p^j(t), \dot{q}^j(t) - w^j(t) \rangle + \lambda(t) \cdot F(q^j) dt.$$

This variational principle now describes a finite-dimensional Hamiltonian system and we can apply the above algorithms without changes to compute geodesics.

This approach to discretization works for all metrics that can be represented via a transform as in Section 3, provided we can characterize the image of this transform, similarly to (19).

I <sub>1</sub>	I <sub>2</sub>	I <sub>1</sub> → I <sub>2</sub>		I <sub>2</sub> → I <sub>1</sub>	
		distance	time (s)	distance	time (s)
cat	cow	17.106	1.93	17.106	2.25
cat	dog	21.349	2.59	21.349	2.24
cat	donkey	25.273	2.59	25.273	2.59
cow	dog	18.389	2.22	18.389	2.21
cow	donkey	20.206	2.56	20.206	2.26
dog	donkey	14.983	2.25	14.983	1.91
shark	airplane	20.488	2.65	20.488	2.34

TABLE 1. Geodesic distances between parameterized shapes, together with the time in seconds to compute the distance. It can be seen the the distances are symmetric between pairs of parameterized shapes.

**7.1. Numerical Results.** In this section we present experiments demonstrating our numerical computations with parameterized curves. The algorithms described in the previous section were implemented in Python using the NumPy and SciPy libraries. The shapes used for the experiments come from the database of closed binary shapes collected by the LEMS Vision Group at Brown university (<http://www.lems.brown.edu/~dmc>). Each curve was initially parameterized by a set of 300 points positioned equidistantly along the curve, and each shape is centered at an arbitrary point (the position of the shape does not matter, as the metric is invariant under translations). The geodesics were computed using  $N = 25$  time steps.

Figure 1 shows a set of shapes along the minimal geodesic path between two pairs of shapes. The geodesics are sampled at timesteps 0, 5, 10, 15, 20 and 25. The first and last images in each row thus show the template and target curves respectively. The template curve is always parameterized proportional to arc length. Table 1 shows the geodesic distances forwards and backwards between parameterized shapes, along with the time for the computation (based on a Python implementation running on a virtual server with 3.6Gb of memory and access to a dual core 3GHz Xeon processor). It can be seen that these distances are symmetric, as expected.

## 8. NUMERICAL COMPUTATION OF GEODESICS ON UNPARAMETERIZED CURVES

**8.1. Finding the optimal parameterization.** In this section we build on the previous section and describe our method for finding geodesics between unparameterized shapes. As was explained in Section 6, this corresponds to finding horizontal geodesics or equivalently finding the minimum of

$$\text{dist}(C, D) = \inf_{\psi \in \text{Diff}(S^1)} \text{dist}(c, d \circ \psi),$$

with  $C = \pi(c)$ ,  $D = \pi(d)$  and where  $\text{dist}$  on the right denotes the geodesic distance on parameterized curves, computed as described in Section 7. We will compute this minimum using a gradient descent algorithm on

$$(20) \quad E(\varphi) = \frac{1}{2} \text{dist}(c, d \circ \varphi^{-1})^2.$$

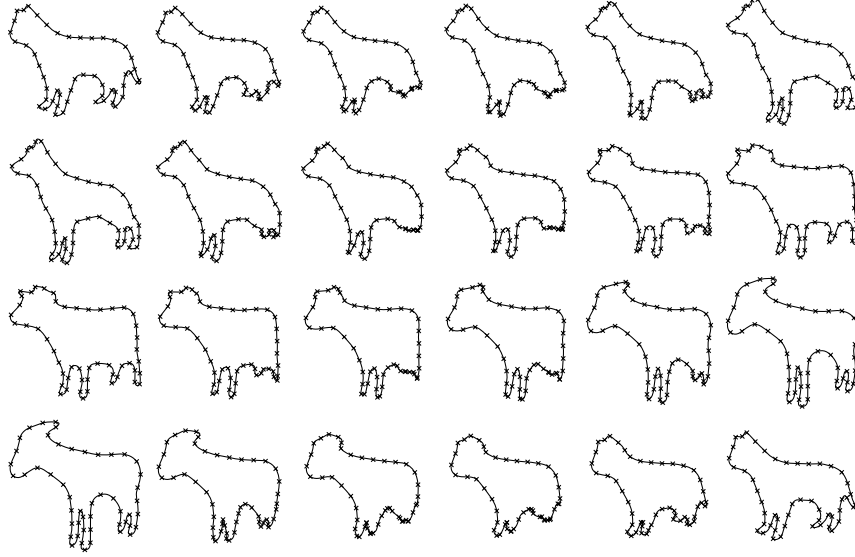


FIGURE 2. Examples of geodesics between several parametrized shapes. The geodesics shown are from the cat to the dog in the first row, from the dog to the cow in the second, from the cow to the donkey in the third and from the donkey back to the cat in the last row thus forming a geodesic quadrangle.

We use  $d \circ \varphi^{-1}$  instead of  $d \circ \varphi$ , in order to have a left action of  $\text{Diff}(S^1)$  on the space of curves. We want to compute the right-trivialized gradient  $\nabla_\varphi E$  of  $E$ , defined as  $\langle \nabla_\varphi E, \mu \rangle_{\mathfrak{X}(S^1)} = T_\varphi E(\mu \circ \varphi)$  for  $\mu \in \mathfrak{X}(S^1)$  and a choice of an inner product  $\langle \cdot, \cdot \rangle_{\mathfrak{X}(S^1)}$  on  $\mathfrak{X}(S^1)$ .

The gradient of

$$G(p) = \frac{1}{2} \text{dist}(p, q)^2$$

is given by  $\nabla_p G = -\text{Log}_p q$ . Hence

$$T_\varphi E(\delta\varphi) = G_{d \circ \varphi^{-1}}(-\text{Log}_{d \circ \varphi^{-1}} c, -(d \circ \varphi^{-1})' \delta\varphi \circ \varphi^{-1})$$

and by writing  $\delta\varphi = \mu \circ \varphi$  for some  $\mu \in \mathfrak{X}(S^1)$  we get

$$T_\varphi E(\mu \circ \varphi) = G_{d \circ \varphi^{-1}}(\text{Log}_{d \circ \varphi^{-1}} c, (d \circ \varphi^{-1})' u).$$

After introducing an inner product on  $\mathfrak{X}(S^1)$ , we could compute the gradient of  $E(\varphi)$  by solving

$$\langle \nabla_\varphi E, u \rangle_{\mathfrak{X}(S^1)} = G_{d \circ \varphi^{-1}}(\text{Log}_{d \circ \varphi^{-1}} c, (d \circ \varphi^{-1})' u).$$

There is however a better expression for the gradient, which fits better with the action of  $\text{Diff}(S^1)$  on  $\text{Imm}(S^1, \mathbb{R}^2)$ . Consider the related function

$$\overline{E}(\varphi) = \frac{1}{2} \text{dist}(c \circ \varphi^{-1}, d)^2.$$

Its gradient can be computed using the invariance of the metric under  $\varphi$ , i.e.,  $\overline{E}(\varphi) = E(\varphi^{-1})$ , which implies

$$\begin{aligned} T_\varphi \overline{E}(\delta\varphi) &= T_{\varphi^{-1}} E \left( -(\varphi^{-1})' \delta\varphi \circ \varphi^{-1} \right) \\ &= -T_{\varphi^{-1}} E \left( \left( \frac{1}{\varphi'} \delta\varphi \right) \circ \varphi^{-1} \right) \\ &= G_{d \circ \varphi} \left( -\text{Log}_{d \circ \varphi} c, (d \circ \varphi)' \frac{1}{\varphi'} \delta\varphi \right) \\ &= G_{d \circ \varphi} \left( -\text{Log}_{d \circ \varphi} c, (d' \circ \varphi) \delta\varphi \right). \end{aligned}$$

Using the invariance of  $\text{Log}$  and the metric  $G_c$  under reparameterizations,

$$\begin{aligned} \text{Log}_{d \circ \varphi} c \circ \varphi &= (\text{Log}_d c) \circ \varphi \\ G_{d \circ \varphi}(h \circ \varphi, k \circ \varphi) &= G_d(h, k), \end{aligned}$$

we get

$$\begin{aligned} T_\varphi \overline{E}(\delta\varphi) &= G_{d \circ \varphi} \left( -(\text{Log}_d(c \circ \varphi^{-1})) \circ \varphi, (d' \circ \varphi) \delta\varphi \right) \\ &= G_d \left( -\text{Log}_d(c \circ \varphi^{-1}), d'(\delta\varphi \circ \varphi^{-1}) \right), \end{aligned}$$

and hence the gradient can be obtained by solving

$$\langle \nabla_\varphi \overline{E}, \mu \rangle_{\mathfrak{X}(S^1)} = -G_d(\text{Log}_d(c \circ \varphi^{-1}), d' \mu).$$

Since the functions  $E$  and  $\overline{E}$  differ only by exchanging  $c$  and  $d$ , we can also express the gradient of  $E$  by:

$$(21) \quad \langle \nabla_\varphi E, \mu \rangle_{\mathfrak{X}(S^1)} = -G_c(\text{Log}_c(d \circ \varphi^{-1}), c' \mu).$$

As the inner product on the space of vector fields we can use the one induced by the identification of  $\mathfrak{X}(S^1)$  with the vertical space at the curve  $c$  using the infinitesimal action  $\zeta_\mu(c) = c' \mu$ . The inner product is thus given by

$$(22) \quad \langle \mu, \nu \rangle_{\mathfrak{X}(S^1)} = G_c(c' \mu, c' \nu).$$

**Theorem 8.2.** *The (right-trivialized) gradient of the energy*

$$E(\varphi) = \frac{1}{2} \text{dist}(c, d \circ \varphi^{-1})^2$$

*with respect to the inner product (22) is given by the vector field on  $S^1$  corresponding to the vertical projection of  $-\text{Log}_c(d \circ \varphi^{-1})$ , i.e.*

$$c' \nabla_\varphi E = -\text{Ver}_c(\text{Log}_c(d \circ \varphi^{-1})).$$

*It can be computed by solving the equation*

$$G_c(c' \nabla_\varphi E, c' \mu) = -G_c(\text{Log}_c(d \circ \varphi^{-1}), c' \mu) \quad \forall \mu \in \mathfrak{X}(S^1)$$

*for  $\nabla_\varphi E \in \mathfrak{X}(S^1)$ .*

*Proof.* Note that the equation

$$G_c(h, c' \mu) = -G_c(\text{Log}_c d \circ \varphi^{-1}, c' \mu)$$

is satisfied for all  $\mu \in \mathfrak{X}(S^1)$  if and only if

$$h = -\text{Ver}_c(\text{Log}_c(d \circ \varphi^{-1})).$$

Each element in the vertical space corresponds via the infinitesimal action to one element of  $\mathfrak{X}(S^1)$ . The theorem now follows from combining equations (21) and (22).  $\square$

The next algorithm computes the element  $\mu \in \mathfrak{X}(S^1)$  corresponding to the orthogonal projection of a vector  $h \in T_c \text{Imm}(S^1, \mathbb{R}^2)$  to the vertical subspace  $\text{Ver}(c)$ .

```

function VER( $c, h$ )
   $\nu \leftarrow$  test function on  $\mathfrak{X}(S^1)$ 
   $\mu \leftarrow$  solution of  $G_c(c'\mu, c'\nu) = G_c(h, c'\nu)$ 
  return  $\mu$ 
end function

```

We use a finite element method with Lagrange elements of first order to numerically compute the vertical projection. A more explicit formula for the inner product  $G_c(h, k)$  is given by

$$G_c(h, k) = \int_{S^1} \frac{\langle h', k' \rangle}{|c'|} - \frac{3}{4} \frac{\langle h', c' \rangle \langle k', c' \rangle}{|c'|^3} d\theta.$$

The algorithm to find geodesics between unparameterized shapes  $C, D$  takes as input two parameterizations  $c, d$  of these shapes such that  $\pi(c) = C$  and  $\pi(d) = D$  and finds the diffeomorphism  $\psi \in \text{Diff}(S^1)$ , such that the geodesic distance  $\text{dist}(c, d \circ \psi)$  is minimal.

```

function SOLVEBVP( $c, d, N$ )
   $\psi \leftarrow \text{Id}_{S^1}$  ▷ Notation  $\psi := \varphi^{-1}$  in (20)
  while  $\text{dist}(c, d \circ \psi)$  is not minimal do
     $h \leftarrow \text{LOG}(c, d \circ \psi, N)$ 
     $\mu \leftarrow \text{VER}(c, h)$  ▷ Notation  $\mu := -\nabla_\varphi E$ 
     $\eta \leftarrow \text{FLOW}(-\mu, \text{Id}_{S^1}, \alpha)$ 
     $\psi \leftarrow \psi \circ \eta$ 
  end while
  return  $\psi$ 
end function

```

To understand the algorithm note that  $\psi$  corresponds to  $\varphi^{-1}$  in Theorem 8.2. There is no need to compute  $\varphi$  itself, as only  $\varphi^{-1}$  is necessary to compute the reparameterization of the curve  $d$ . In each iteration of the algorithm we first compute the gradient  $\mu = -\nabla_\varphi E$  with the help of Theorem 8.2. A continuous gradient descent would take the form  $\partial_t \varphi = -\nabla_\varphi E \circ \varphi$ , while a first order time-discretization would be

$$\varphi_{i+1} = \text{Fl}^\mu(\alpha, \text{Id}_{S^1}) \circ \varphi_i$$

with  $\text{Fl}^\mu(\alpha, \text{Id}_{S^1})$  denoting the flow of the vector field  $\mu = -\nabla_\varphi E$  up to time  $\alpha$  starting from  $\text{Id}_{S^1}$  at time 0. Since we are only interested in  $\psi = \varphi^{-1}$ , we can rewrite this as

$$\begin{aligned} \psi_{i+1} &= \varphi_{i+1}^{-1} = \varphi_i^{-1} \circ \text{Fl}^\mu(\alpha, \text{Id}_{S^1})^{-1} \\ &= \psi_i \circ \text{Fl}^{-\mu}(\alpha, \text{Id}_{S^1}). \end{aligned}$$

The gradient descent step is repeated until the relative decrease of  $\text{dist}(c, d \circ \psi)$  falls below a prescribed threshold. Similarly to the function  $\text{Log}_{q_0}(q_1)$  in Section 7, the parameter  $\alpha$  is the step-size and is chosen adaptively to ensure the distance decreases in each step.

**8.3. Adaptive grid refinement.** The behaviour of shortest paths between unparameterized curves can be understood as a combination of stretching and bending.

Bending of a parameterized curve is numerically well behaved, since uniformly sampled curves will stay approximately uniformly sampled. However, stretching tends to expand a very short section of the curve into a much larger one and will lead to a curve that is under-sampled in the expanded area, unless the original parameterization was chosen to counter this effect. See Figure 3 for an example of this.

In the gradient descent algorithm from Section 8.1 we start with a curve that is sampled uniformly at the points

$$x_0 = 0, x_1 = \frac{2\pi}{n}, \dots, x_{n-1} = \frac{n-1}{n}2\pi.$$

At each iteration, given a grid  $x_0, \dots, x_{n-1}$  for the curve  $c$ , the curve  $d$  would be sampled at the points  $\psi(x_0), \dots, \psi(x_{n-1})$ . Stretching of the curve  $c$  would correspond to a large derivative  $\psi'(x)$  or a large distance  $|\psi(x_{i+1}) - \psi(x_i)|$  between two consecutive points. We add points to the grid, whenever the distance exceeds that of a uniform grid, i.e.

$$|\psi(x_{i+1}) - \psi(x_i)| > \frac{2\pi}{n_0}.$$

Here  $n_0$  is the size of the original grid. By this we ensure that the target curve  $d \circ \psi$  will not be under-sampled.

Similarly, we remove a point  $x_i$  whenever the neighbouring points would be sufficient to provide enough resolution, i.e., these conditions are satisfied:

$$|x_{i+1} - x_{i-1}| < \frac{2\pi}{n_0} \text{ and } |\psi(x_{i+1}) - \psi(x_{i-1})| < \frac{2\pi}{n_0}.$$

**8.4. Numerical Results.** In order to compute the vertical projection of a tangent vector as described in Section 8.1 we used the finite element library FEniCS [16]. To find the optimal parameterization the gradient descent algorithm from Section 8.1 with a maximum of 100 iterations was used.

Figure 4 shows examples of some of the geodesics. From a mathematical point of view we would expect the geodesics between two shapes to be symmetric, i.e., interchanging the curves  $c_0$  and  $c_1$ , representing the shapes, should have no effect. However, if we look at the second and third rows of the figure, we see that the path between the cow and the dog is not completely symmetrical. This behaviour originates in the incompleteness of the metric. The optimal geodesic from the dog to the cow (shown backwards in the last line) shrinks a leg of the dog and at the same time grows another leg of the cow from the body, and such growth tends to originate from a single point. Thus, one point of the dog expands to create the whole leg of the cow, while the whole leg of the dog collapses to one point on the cow.

In Figure 3 we show a more pronounced example of this behaviour. The template shape  $c_0$  is an ellipse and the target shape  $c_1$  is an ellipse with the added fold both discretized by equidistantly spaced points. There is no problem to compute the geodesic between  $c_0$  and  $c_1$ . However in the process of computing the geodesic between the shapes  $\pi(c_0)$  and  $\pi(c_1)$  we see that the horizontal geodesic wants to create the whole fold out of a single point on the ellipse. The development of this singularity means that there exists no geodesic in  $\mathcal{S}$  between the shapes  $\pi(c_0)$  and  $\pi(c_1)$ . This behaviour was also described in [31, Sect. 4.2] for a related metric and it is a property of the Riemannian metric, rather than an artefact of the discretization.





FIGURE 3. This image shows the minimal geodesic between the two shapes in the first row. The template shape is an ellipse while the target shape is an ellipse with a large fold. The bottom left image shows the minimal geodesic computed with the grid adaptively refined as in Section 8.3. In the bottom right image the geodesic is computed without grid refinement, which leads to a loss of resolution along the fold. Since the fold is growing out of a point it is necessary to refine the grid in the neighbourhood of this point to accurately capture the fold.

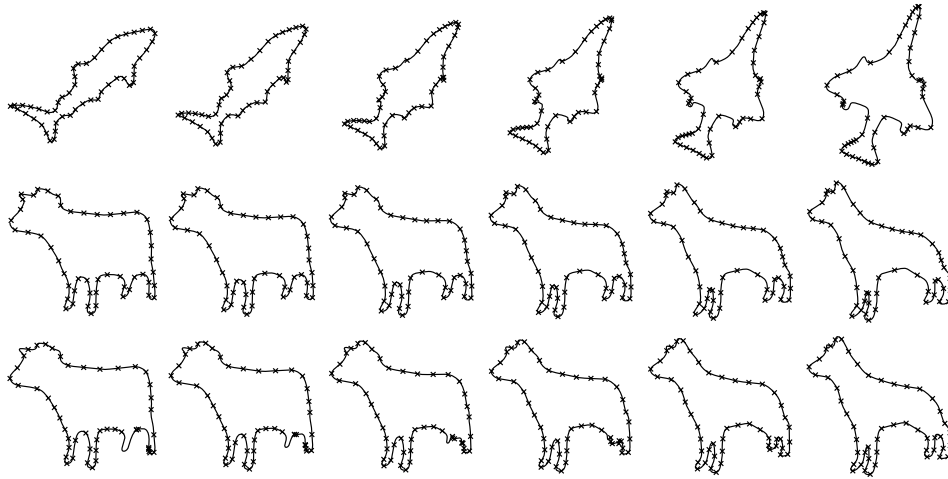


FIGURE 4. Minimal geodesics between various shapes with optimized parameterization. The bottom row shows the backwards geodesic to the middle row (template and target curves reversed) but printed backwards, so that it can be more easily compared with the line above. Mathematically this should be perfectly symmetric, but since there is stretching and compression along the geodesics it is not, and the geodesic distances computed will differ.

The incompleteness of the metric has consequences for its use in shape analysis. If geodesics don't necessarily exist between any two shapes, then it is not possible to linearize shape space using the Riemannian exponential map. Statistical methods that are based on the exponential map will be blind to potentially large parts of shape space.

## 9. CONCLUSIONS

Riemannian metrics on shape space of plane curves are of great interest in a wide variety of applications in image analysis and computational anatomy. Due to the infinite-dimensional nature of shape space, metrics that allow for efficient computations of geodesics and distance between shapes are particularly useful.

In this paper we generalize the  $R$ -transform, first introduced in the work of [27], to the class of all elastic metrics whose coefficients satisfy  $4b^2 \geq a^2$ , where  $a$  and  $b$  are parameters controlling the degree of bending and stretching of the curve respectively (see Section 1.1). This transformation allows us to obtain efficient algorithms for computing the geodesic distance between shapes as well as to gain a better understanding of the geometry of the space. For the case of open curves we show that the space of parameterized open curves is a flat space in the sense of Riemannian geometry and obtain explicit formulas for geodesics. As a consequence, it follows that the shape space of unparameterized open curves has positive sectional curvature. For closed curves the situation is more difficult, since the space of parameterized curves is not flat anymore, but the representation is still very useful from a numerical point of view.

We have presented numerical results for geodesics between both parametrized and unparameterized closed curves. For parameterized curves the computed metric is symmetric (that is, the distance between two curves  $c$  and  $d$  does not depend which is the source and which the target), while for unparametrized curves incompleteness of the space makes the task of computing geodesics much more difficult. Further work is required to understand this behaviour and to either develop numerical methods capable of dealing with this situation or to find Riemannian metrics for which the geodesic boundary value problem on shape space is solvable. We also plan to demonstrate the use of the metrics for the classification of sets of shapes.

## ACKNOWLEDGMENTS

This research was partly supported by the FWF-Projects P2462511 and P21030-N13 as well as by an Advanced Grant from the European Research Council and the Royal Society of New Zealand Marsden Fund. We thank Colin Cotter for helpful discussions and valuable comments.

## REFERENCES

- [1] H. C. Andersen. Rattle: A velocity version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1):24 – 34, 1983.
- [2] M. Bauer and M. Bruveris. A new Riemannian setting for surface registration. In *3rd MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pages 182–194, 2011.
- [3] M. Bauer, M. Bruveris, and P. W. Michor.  $R$ -transforms for Sobolev  $H^2$ -metrics on spaces of plane curves. To appear in: *Geometry, Imaging and Computing*, 2013.
- [4] M. Bauer, P. Harms, and P. W. Michor. Sobolev metrics on shape space of surfaces. *J. Geom. Mech.*, 3(4):389–438, 2011.
- [5] M. Bauer, P. Harms, and P. W. Michor. Almost local metrics on shape space of hypersurfaces in  $n$ -space. *SIAM J. Imaging Sci.*, 5:244–310, 2012.
- [6] M. Bauer, P. Harms, and P. W. Michor. Curvature weighted metrics on shape space of hypersurfaces in  $n$ -space. *Differential Geom. Appl.*, 30(1):33–41, 2012.
- [7] M. Bauer, P. Harms, and P. W. Michor. Sobolev metrics on shape space, ii: Weighted sobolev metrics and almost local metrics. *J. Geom. Mech.*, 4(4):365 – 383, 2012.
- [8] G. Charpiat, R. Keriven, and O. Faugeras. Shape statistics for image segmentation with priors. In *Conference on Computer Vision and Pattern Recognition*, 2007.

- [9] C. Cotter, A. Clark, and J. Peiró. A reparameterisation based approach to geodesic constrained solvers for curve matching. *International Journal of Computer Vision*, 99:103–121, 2012.
- [10] D. G. Ebin and J. Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)*, 92:102–163, 1970.
- [11] J. Glaunès, A. Qiu, M. Miller, and L. Younes. Large deformation diffeomorphic metric curve mapping. *International Journal of Computer Vision*, 80:317–336, 2008.
- [12] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quart. Appl. Math.*, 56:617–694, 1998.
- [13] S. Kurtek, E. Klassen, Z. Ding, S. Jacobson, J. Jacobson, M. Avison, and A. Srivastava. Parameterization-invariant shape comparisons of anatomical surfaces. *IEEE Transactions on Medical Imaging*, 30(3):849–858, 2011.
- [14] S. Kurtek, E. Klassen, J. Gore, Z. Ding, and A. Srivastava. Elastic geodesic paths in shape space of parametrized surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP(99):1, 2011.
- [15] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Number 14 in Cambridge Monographs on Applied and Computational Mathematics. Springer, 2004.
- [16] A. Logg and G. Wells. DOLFIN: Automated finite element computing. *ACM Transactions on Mathematical Software*, 37(2), 2010.
- [17] M. Mani, S. Kurtek, C. Barillot, and A. Srivastava. A comprehensive Riemannian framework for the analysis of white matter fiber tracts. In *IEEE Symposium on Biomedical Imaging*, pages 1101–1104, 2010.
- [18] A. C. G. Mennucci, A. Yezzi, and G. Sundaramoorthi. Properties of sobolev-type metrics in the space of curves. *Interfaces and Free Boundaries*, 8(4):423–445, 2008.
- [19] P. W. Michor. *Topics in differential geometry*, volume 93 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.
- [20] P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math.*, 10:217–245 (electronic), 2005.
- [21] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)* 8 (2006), 1–48, 2006.
- [22] 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(1):74–113, 2007.
- [23] W. Mio, A. Srivastava, and S. Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73:307–324, 2007.
- [24] C. Samir, P.-A. Absil, A. Srivastava, and E. Klassen. A gradient-descent method for curve fitting on Riemannian manifolds. *Found. Comput. Math.*, 12(1):49–73, 2012.
- [25] J. Shah. An  $H^2$  type Riemannian metric on the space of planar curves. In *1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pages 40–46, 2006.
- [26] E. Sharon and D. Mumford. 2D-Shape Analysis Using Conformal Mapping. *International Journal of Computer Vision*, 70(1):55–75, 2006.
- [27] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, 2011.
- [28] G. Sundaramoorthi, A. Mennucci, S. Soatto, and A. Yezzi. A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. *SIAM Journal on Imaging Sciences*, 4(1):109–145, 2011.
- [29] G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Sobolev active contours. *International Journal of Computer Vision*, 73:345–366, 2007.
- [30] L. Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.*, 58(2):565–586 (electronic), 1998.
- [31] 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. Rend. Lincei (9) Mat. Appl.*, 19(1):25–57, 2008.