Published in final edited form as:

SIAM J Imaging Sci. 2020 ; 13(1): 445–473. doi:10.1137/19m1265132.

Simplifying Transforms for General Elastic Metrics on the Space of Plane Curves

Tom Needham[†], **Sebastian Kurtek**[‡]

[†]Department of Mathematics, Florida State University, Tallahassee, FL 32306

[‡]Department of Statistics, The Ohio State University, Columbus, OH 43210

Abstract

In the shape analysis approach to computer vision problems, one treats shapes as points in an infinite-dimensional Riemannian manifold, thereby facilitating algorithms for statistical calculations such as geodesic distance between shapes and averaging of a collection of shapes. The performance of these algorithms depends heavily on the choice of the Riemannian metric. In the setting of plane curve shapes, attention has largely been focused on a two-parameter family of first order Sobolev metrics, referred to as elastic metrics. They are particularly useful due to the existence of simplifying coordinate transformations for particular parameter values, such as the well-known square-root velocity transform. In this paper, we extend the transformations appearing in the existing literature to a family of isometries, which take any elastic metric to the flat L^2 metric. We also extend the transforms to treat piecewise linear curves and demonstrate the existence of optimal matchings over the diffeomorphism group in this setting. We conclude the paper with multiple examples of shape geodesics for open and closed curves. We also show the benefits of our approach in a simple classification experiment.

Keywords

elastic shape analysis; statistical shape analysis; infinite-dimensional geometry; Sobolev metrics; curve matching

AMS subject classifications.

Primary, 58B20, 58E50; Secondary, 68U05

1. Introduction.

Shape is a fundamental physical property of objects and plays an important role in various imaging tasks, including identification and tracking. As a result, statistical analysis of shape plays a crucial role in many image-rich application domains such as computer vision, medical imaging, biology, bioinformatics, geology, and biometrics. In statistical shape analysis, shape is viewed as a random object, and one is concerned with developing methods

(tneedham@fsu.edu).

to perform common statistical tasks, including registration, comparison, averaging, summarization of variability, hypothesis testing, regression, and other inferential procedures. Any statistical shape analysis approach requires an appropriate shape representation and an associated metric that enables quantification of shape differences. Evidently, the quality of statistical analyses of shape data is heavily dependent on these choices.

There is a rich literature on statistical analysis of shape, with the most prominent shape representation being landmark-based. Landmarks constitute a finite collection of points that are chosen either by the application expert (anatomical landmarks) or according to some mathematical rule such as high absolute curvature (mathematical landmarks). Once the points are selected, the remaining information regarding the object's outline is discarded. Under this representation, Kendall [25] defined shape as a property of an object that is invariant to its rigid motions and global scaling; this approach is commonly referred to as similarity shape analysis. Since then, there has been continuous development of statistical tools to analyze similarity shapes represented by landmarks; see [17, 43] for a comprehensive set of methods. These approaches combine ideas from differential geometry, algebra, and multivariate statistics to establish rigorous estimation and inferential procedures on the landmark shape space. The main benefit of these approaches is that the resulting shape space is finite-dimensional, making statistical analysis "easier." However, the obvious drawback is that the finite collection of landmarks used to represent shapes of interest results in significant loss of information.

Recently, there has been more emphasis on using a function-based representation of shape, i.e., objects are represented via their boundaries as parameterized curves. Thus, in this case, one must account for possible parameterization variability in addition to rigid motion and global scaling. One set of methods removes this variability by normalizing all parameterizations to arclength [26, 56]. However, such an approach is suboptimal in many real scientific problems due to a lack of appropriate registration. A better approach is to remove such variability in a pairwise manner using an appropriate metric. This is the idea behind elastic shape analysis, where a family of elastic metrics is used for joint registration and comparison. The resulting shape spaces are more complicated than their landmark counterparts, but the benefits of such approaches are clear: (1) there is no need to select landmarks, which can be a tedious and expensive process; (2) the curve representation is able to encode all relevant shape information; and (3) the elastic metric quantifies intuitive shape deformations. Elastic shape analysis is the focus of the current paper, and we provide a formal mathematical setup for this approach in the following section.

1.1. Elastic shape analysis.

A fundamental ingredient in a theory of shape similarity for plane curves is a distance metric on the space of curves $\mathcal{S}$ which is invariant under rigid transformations of the curves. For a pair of plane curves C_1 and C_2, we therefore wish to assign a distance $d(C_1, C_2)$ such that $d(\xi_1 \star C_1, \xi_2 \star C_2) = d(C_1, C_2)$ for any elements ξ_j of the Euclidean isometry group $\mathbb{R}^2 \rtimes \mathrm{SO}(2)$, acting in the natural way.

Under the elastic shape analysis paradigm, the distance function described above arises from a Riemannian metric. This extra structure has obvious benefits over treating $\mathcal{S}$ only as a metric space; for example, it allows the potential to compute geodesic curve deformations and to locally linearize via the logarithm map in order to do statistics in a tangent space. The metric on the space of (unparameterized) curves $\mathcal{S}$ is obtained by treating it as a quotient space, described as follows. Let $I \subset \mathbb{R}$ denote some fixed interval, $\mathrm{Imm}\left(I, \mathbb{R}^2\right) \subset C^\infty\left(I, \mathbb{R}^2\right)$ the open submanifold of smooth immersions, and $\mathrm{Diff}^+(I)$ the Lie group of orientation-preserving diffeomorphisms of I. This group acts on $\mathrm{Imm}\left(I, \mathbb{R}^2\right)$ by reparameterizations. We then represent the space of curves as $\mathcal{S} \approx \mathrm{Imm}\left(I, \mathbb{R}^2\right)/\mathrm{Diff}^+(I)$; in other words, the space of unparameterized curves is realized as the space of parameterized curves identified up to reparameterizations. A choice g of $\mathrm{Diff}^+(I)$-invariant Riemannian metric on the relatively simple space $\mathrm{Imm}\left(I, \mathbb{R}^2\right)$ descends to a well-defined metric on $\mathcal{S}$. If the Riemannian metric is also invariant under Euclidean similarities, then the geodesic distance with respect to this metric induces our desired distance function d via the formula

$$d(C_1, C_2) = \inf_{\sigma_1, \sigma_2 \in \mathrm{Diff}^+(I) \times \left(\mathbb{R}^2 \times SO(2)\right)} \mathrm{dist}_g(\sigma_1 \star c_1, \sigma_2 \star c_2).$$

In this formula, the c_j are arbitrary choices of parameterizations of C_j, dist_g denotes geodesic distance in $\mathrm{Imm}\left(I, \mathbb{R}^2\right)$ with respect to g, and $\star$ denotes the action of the group $\mathrm{Diff}^+(I) \times \left(\mathbb{R}^2 \ltimes \mathrm{SO}(2)\right)$ of shape-preserving transformations on $\mathrm{Imm}\left(I, \mathbb{R}^2\right)$, defined as follows. A triple $\sigma = (\gamma, v, A) \in \mathrm{Diff}^+(I) \times \left(\mathbb{R}^2 \ltimes \mathrm{SO}(2)\right)$ acts on a curve c by reparameterizing by γ, rotating c about $c(0)$ by A, and translating the image of c by v.

The simplest choice of Riemannian metric is the reparameterization-invariant L^2 metric defined for $c \in \mathrm{Imm}\left(I, \mathbb{R}^2\right)$ and $h, k \in T_c\mathrm{Imm}\left(I, \mathbb{R}^2\right) \approx C^\infty\left(I, \mathbb{R}^2\right)$ by the formula

$$\hat{g}_c^{L^2}(h, k) = \int_I \langle h, k \rangle \, \mathrm{d}s.$$

The notation $\hat{g}$ is used to distinguish this metric from the standard (non-reparameterization-invariant) L^2 metric which will appear later in the paper. The nonlinearity of this metric lies in the *measure with respect to arclength* $\mathrm{d}s = |c'(t)| \, \mathrm{d}t$, which provides the desired $\mathrm{Diff}^+(I)$-invariance. It is a surprising fact that geodesic distance on the shape space vanishes with respect to $\hat{g}^{L^2}$ [34], and one must therefore consider more complicated metrics on the space of immersions. Examples in the literature of such metrics include almost-local (weighted L^2) metrics [7, 8, 35] and higher order Sobolev-type metrics [6, 15, 35, 46]. An element of the latter class of metrics is a natural generalization of the reparameterization-invariant L^2 metric, defined by

$$g_c^{\vec{a}}(h,k) = \int_I a_0^2 \left\langle h, k \right\rangle + a_1^2 \langle D_s h, D_s k \rangle + \cdots + a_n^2 \langle D_s^n h, D_s^n k \rangle \, \mathrm{d}s,$$

where $\vec{a} = (a_0, a_1, \ldots, a_n)$ is a vector of weights on the terms, and we use $D_s = \frac{1}{c'(t)} \frac{d}{dt}$ for *derivative with respect to arclength*. The higher order Sobolev metrics no longer suffer the vanishing geodesic phenomenon and are, in fact, geodesically complete when $a_0, a_n > 0$ for $n \geqslant 2$ [12].

A particularly well-studied subfamily of first order (i.e., $n = 1$ in the above notation) Sobolev metrics are the *elastic metrics* introduced in [37]. These form a two-parameter family of metrics $g^{a,b}$ defined by setting $a_0 = 0$ and further decomposing the first order term into tangential and normal components. That is, for $c \in \mathrm{Imm}\left(I, \mathbb{R}^2\right)$, let (T, N) denote the standard moving frame consisting of the unit tangent and unit normal to c, respectively. For $a, b \neq 0$ and $h, k \in T_c \mathrm{Imm}\left(I, \mathbb{R}^2\right)$, we define

$$g_c^{a,b}(h,k) = \int_I a^2 \langle D_s h, N \rangle \langle D_s k, N \rangle + b^2 \langle D_s h, T \rangle \langle D_s k, T \rangle \, \mathrm{d}s \,. \tag{1.1}$$

This metric is invariant under reparameterizations and rigid motions, and so descends to a well-defined metric on the shape space $\mathcal{S}$. While these metrics do not enjoy the geodesic completeness of their higher order counterparts, we will see below that they have a number of useful theoretical properties and, in particular, that geodesic distance is nonvanishing. This paper will focus exclusively on this family of elastic metrics.

1.2. Previous work on elastic metrics.

In order to compute the distance between curves in $\mathcal{S}$, our procedure requires the computation of geodesic distance with respect to the chosen metric. Early approaches to this task accomplished this by explicitly solving the associated variational problems [47, 48, 53]. Focusing on the elastic metric $g^{1,1/2}$, a common technique is to apply the *square-root velocity function* (SRVF) transform, given by

$$\mathrm{Imm}\left(I, \mathbb{R}^2\right) \to C^\infty\left(I, \mathbb{R}^2\right),$$

$$c \mapsto \frac{c'}{|c'|^{1/2}} \,.$$

The theoretical power of the SRVF is the remarkable fact that the pullback of the standard L^2 metric on the target space is the elastic metric $g^{1,1/2}$ [24], whence geodesics with respect to $g^{1,1/2}$ in $\mathrm{Imm}\left(I, \mathbb{R}^2\right)$ can be computed explicitly by pushing forward to the flat target space, computing geodesics there, then pulling the result back. Due to this convenient property, the SRVF transform has been studied extensively from a theoretical perspective [11, 31, 44, 49]

and has seen a wide variety of applications, including classification of plant leaf shapes [30], statistical analysis of protein structures [45], and biomedical imaging of anatomical features in the brain [2]. A similarly simple transform is introduced in [55], where a plane curve c is taken to the curve $\sqrt{c'}$, with the square root taken pointwise by considering c' to be a path in the complex plane; in this case, the map pulls back the L^2 metric to $g^{1/2,1/2}$. A more complicated family of transforms $R_{a,b}$ is defined in [5] for $2b \geq a > 0$, and it is shown that the pullback by $R_{a,b}$ of the L^2 metric on its target is $g^{a,b}$. A different framework for understanding general elastic metrics, with a more explicit focus on the various Lie group actions, is provided in [54]. There has also been substantial effort put toward numerical computations for geodesics in spaces of curves with respect to these metrics (and more general Sobolev-type metrics); see, e.g., [4, 3].

In this paper, we define a two-parameter family of transforms $F_{a,b} : \mathrm{Imm}\left(I, \mathbb{R}^2\right) \to C^\infty\left(I, \mathbb{R}^2\right)$ which is valid for *all* choices of $a, b > 0$. Our main result (Theorem 2.3) is that $F_{a,b}$ pulls back the L^2 metric to the elastic metric $g^{a,b}$. Moreover, we show that $F_{a,b}$ subsumes the SRVF transform, the complex square-root transform, and the $R_{a,b}$-transforms.

1.3. Other formalisms in shape analysis.

Before moving on to our study of elastic metrics, we briefly remark on some other approaches to shape analysis which are similar in spirit in that they define metrics on various shape spaces. The shape analysis literature is quite extensive and varied, and we make no claims that our description is exhaustive.

As was pointed out above, the choice of metric on shape space depends on the shape representation. A natural way to represent a shape (an embedded curve, surface or otherwise) is as a set of points, either abstractly as a continuous object or computationally as an unstructured point cloud of samples (in line with landmark shape analysis already described above). Under this representation, there are several metrics which can be used to compare shapes. Classical metrics which are still in use are the Hausdorff distance (between compact subsets of a metric space) and Fréchet distance (specifically between plane curves). One could consider an embedded shape as a metric space in its own right (using the restriction of the ambient metric) so that Gromov–Hausdorff distance applies as a shape metric [10]. Choosing a probability measure on a shape (e.g., the empirical measure on a finite sampling of the shape) further turns the shape into a metric measure space, so that Gromov–Wasserstein distance provides another shape metric [33]. The Gromov–Hausdorff and Gromov–Wasserstein distances are naturally invariant under rigid motions but are computationally expensive. Unstructured point cloud shape representations can also be compared using techniques from topological data analysis, which enjoy stability with respect to Gromov–Hausdorff distance [16]. These methods are all flexible enough to handle very general classes of shapes, but when shapes come from a fixed class (such as plane curves) they forget that extra structure and only see metric information.

Another shape analysis formalism which is particularly popular in computational anatomy is the Large Deformation Diffeomorphic Metric Mapping framework (see, e.g., [19, 14]), which compares shapes (embedded submanifolds) by looking for an optimal

diffeomorphism of the entire ambient space taking one shape to another. This method comes with its own challenges of rigid motion shape registration and higher computational complexity, but it is intuitively appealing and flexible enough to handle a variety of different shape representations. For example, these ideas can be used to directly compare images without the need to segment shapes [36]. Similarly, ideas from optimal transport can be used to compare images by treating them as probability distributions [21]. Optimal transport methods have also proven useful for comparing anatomical surfaces as embedded submanifolds [9].

The rest of the paper will focus exclusively on an elastic shape analysis framework which is specifically designed for comparing plane curves (although we note that this framework has itself been generalized to treat many other classes of shapes, such as embedded surfaces [23] and neuronal trees [18]). The best choice of framework for shape analysis is largely application-specific, depending on requirements for computational efficiency or robustness to noise and on the particular structure of the available shape data.

1.4. Outline of the paper.

In section 2, we define the transform $F_{a,b}$ and prove our main result. We also consider the important submanifold of closed plane curves, and more precisely compare our transform to those described in the previous subsection. Section 3 describes how various shape-preserving group actions behave in $F_{a,b}$-coordinates. In section 4, we describe the explicit geodesics in the curve spaces. Numerical implementation is formally treated in section 5, where the transform is extended to treat piecewise linear (PL) curves. In particular, we show that optimal registrations over the diffeomorphism group are realized by PL reparameterizations in this setting. Finally, we provide numerical examples[1] in section 6 and suggest future directions in section 7.

2. The $F_{a,b}$-transform.

2.1. Shape spaces of open curves.

The space of plane curve shapes is obtained via a quotient construction, starting with the space of immersions

$$\mathrm{Imm}\left(I,\mathbb{R}^2\right)=\left\{c\in C^\infty\left(I,\mathbb{R}^2\right)\middle|\,|c'(t)|\neq 0\,\forall t\in I\right\},$$

where, without loss of generality, $I=[0,1]$ is fixed. To simplify calculations, we make the identification $\mathrm{Imm}\left(I,\mathbb{R}^2\right)\approx\mathrm{Imm}(I,\mathbb{C})$. We note here that our choice of C^∞ regularity is primarily a matter of convenience and that our results hold essentially without modification for C^1 curves. Decreasing regularity below C^1 does cause some theoretical issues, and these are treated in section 5. Throughout the paper, we use the Whitney-C^∞ topology on $C^\infty\left(I,\mathbb{R}^2\right)$, which turns the space into a tame Fréchet space. With this topology, $\mathrm{Imm}(I,\mathbb{C})$ is

[1]Our code is available for download from the GitHub repository https://github.com/trneedham/Planar-Elastic-Metrics.

an open submanifold of the vector space $C^\infty\left(I, \mathbb{R}^2\right)$. For details about topology and calculus in the tame Fréchet category, see the standard references [22, 27].

There are various shape-preserving Lie group actions which we will quotient by $\mathbb{R}^2$ acting by translations, $\mathbb{R}_{>0}$ acting by scaling, SO(2) acting by rotations, and $\mathrm{Diff}^+(I)$ acting by reparameterizations. The easiest action to deal with is translations, as there is an obvious isomorphism with the space of curves based at zero,

$$\mathrm{Imm}(I,\mathbb{C})/\mathbb{R}^2 = \mathrm{Imm}(I,\mathbb{C})/\mathrm{Tra} \approx \left\{c \in C^\infty(I,\mathbb{C}) \mid c'(t) \neq 0 \forall t \in I,\ c(0) = 0\right\}.$$

We take this identification as a convention in order to simplify notation. Most of our explicit calculations will take place in this space, which we refer to as the *preshape space of curves*.

Our goal is to understand the following quotient space with the full set of shape similarities modded out:

$$\mathrm{Imm}(I,\mathbb{C})/\left(\mathbb{R}^2 \ltimes \mathbb{R}_{>0} \times \mathrm{SO}(2) \times \mathrm{Diff}^+(I)\right) = \mathrm{Imm}(I,\mathbb{C})/\left\{\mathrm{Tra}, \mathrm{Sca}, \mathrm{Rot}, \mathrm{Rep}\right\}.$$

We refer to this quotient space as the *shape space of curves* and denote it by $\mathcal{S}$. Intermediate spaces such as $\mathrm{Imm}(I,\mathbb{C})/\{\mathrm{Tra}, \mathrm{Rot}\}$ will appear frequently, and we will treat them separately as they arise.

2.2. The $F_{a,b}$-transform on preshape space.

For any $a, b > 0$, we define the $F_{a,b}$-transform by the formula

$$F_{a,b}: \mathrm{Imm}(I,\mathbb{C})/\mathrm{Tra} \to C^\infty(I,\mathbb{C}^*),$$
$$c \mapsto 2b|c'|^{1/2}\left(\frac{c'}{|c'|}\right)^{\frac{a}{2b}}.$$

We use the notation $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$. In the formula, all arithmetic operations are taken pointwise on complex numbers. One should immediately notice that, due to the presence of complex exponentiation, $F_{a,b}$ is not well-defined in general. Indeed, writing c' in polar coordinates $r\exp(i\theta)$, the continuous argument function θ is only unique up to a global addition of an integer multiple of 2π. The values of $F_{a,b}$ are then given by

$$2br^{1/2}\exp\left(i(\theta + 2k\pi)\frac{a}{2b}\right) = 2br^{1/2}\exp\left(i\theta\frac{a}{2b}\right)\cdot\exp\left(i\frac{a}{b}k\pi\right)$$

for $k \in \mathbb{Z}$. Therefore $F_{a,b}$ is technically defined as a multivalued function with image set

$$F_{a,b}(c) = \left\{q \cdot \exp\left(i\frac{a}{b}k\pi\right) \mid k \in \mathbb{Z}\right\}, \tag{2.1}$$

where $q = 2br^{1/2}\exp(i\theta a/2b)$ for some arbitrary choice of smooth polar coordinate representation $c' = r\exp(i\theta)$. In this form, it is easy to see that $F_{a,b}$ is a bijection if and only

if $\frac{a}{2b} = 1$. If $\frac{a}{2b}$ is an integer not equal to one, then $F_{a,b}$ is well-defined but many-to-one. If $\frac{a}{2b}$ is not an integer, then $F_{a,b}$ is multivalued, taking finitely many values if and only if $\frac{a}{2b}$ is rational.

For the sake of concreteness, we can locally define $F_{a,b}$ more precisely as follows. Let $c_0 \in \mathrm{Imm}(I, \mathbb{C})/\mathrm{Tra}$ and choose a polar coordinate representation of its derivative $c_0' = r_0 \exp(i\theta_0)$ so that θ_0 is continuous on I. The magnitude function r_0 is unique and such a choice of θ_0 is unique up to addition of an integer multiple of 2π. Moreover, any parameterized curve c which is sufficiently C^∞-close to c_0 has a polar representation $c' = r \exp(i\theta)$ so that θ_0 and θ are C^∞-close (with respect to any metric generating the Whitney topology). The $F_{a,b}$-transform is then defined locally near c_0 by

$$F_{a,b}(c) = 2br^{1/2}\exp\left(i\theta\frac{a}{2b}\right). \tag{2.2}$$

The polar coordinate representation (2.2) shows that $F_{a,b}$ can be represented locally as a well-defined continuous map of Fréchet spaces. In fact, the map is locally smooth. Recall that a map on open subsets of Fréchet spaces $\mathcal{U} \to \mathcal{V}$ is called *smooth* if its composition with any smooth path $\mathbb{R} \to \mathcal{U}$ results in a smooth map $\mathbb{R} \to \mathcal{V}$ (the usual definition of the derivative of a path still makes sense in the Fréchet category) [22]. One can easily check that this property holds for the local representation of $F_{a,b}$. We note that this allows us to take directional derivatives of $F_{a,b}$ in the usual way. Finally, we observe that the transform $F_{a,b}$ is locally bijective with inverse given by the locally smooth map

$$F_{a,b}^{-1}(q)(t) = \frac{1}{4b^2}\int_0^t |q(\tau)|^2\left(\frac{q(\tau)}{|q(\tau)|}\right)^{\frac{2b}{a}} \mathrm{d}\tau, \tag{2.3}$$

which we can make locally well-defined by once again passing to a polar representation. It follows that $F_{a,b}$ is a local diffeomorphism of Fréchet spaces.

Remark 2.1. Formula (2.2) involves a choice of image of $F_{a,b}$. Fortunately, all other choices of image differ from this one by a rotation, and we will see in Theorem 3.2 that this implies that $F_{a,b}$ descends to a well-defined map on quotient spaces of curves modulo rotation.

Remark 2.2. One issue that can arise in the elastic shape analysis approach to shape matching is that curves which are close in Hausdorff distance can be far apart in geodesic distance (see Figure 1). This may be undesirable, depending on the application. One can overcome this issue by restricting analysis to simple curves or by using ad hoc methods to account for such differences.

2.3. Pullback metric.

Let g^{L^2} denote the standard L^2 metric on the space $C^\infty(I, \mathbb{C})$, defined at basepoint $q \in C^\infty(I, \mathbb{C})$ on variations $w, z \in T_q C^\infty(I, \mathbb{C}) \approx C^\infty(I, \mathbb{C})$ by the formula

$$g_q^{L^2}(w,z) = \mathrm{Re}\int_I w\bar{z}\,\mathrm{d}t\,.$$

This is a flat (i.e., sectional curvatures are identically zero) metric on the vector space $C^\infty(I,\mathbb{C})$ which restricts to a flat metric on the open submanifold $C^\infty(I,\mathbb{C}^*)$.

Theorem 2.3. *The L^2 metric g^{L^2} on $C^\infty(I,\mathbb{C}^*)$ pulls back to the elastic metric $g^{a,b}$ on* $\mathrm{Imm}(I,\mathbb{C})$ *under the transform $F_{a,b}$.*

Proof. Let $c \in \mathrm{Imm}(I,\mathbb{C})$ and let h be a tangent vector to c. We first note that, expressing the unit normal N to c as iD_sc and the Euclidean inner product as the real part of $(z,w)\mapsto z\overline{w}$, the elastic metric $g^{a,b}$ can be written as

$$\begin{aligned} g_c^{a,b}(h,k) &= \int_I a^2\langle D_s h, N\rangle\langle D_s k, N\rangle + b^2\langle D_s h, T\rangle\langle D_s k, T\rangle\,\mathrm{d}s \\ &= \int_I \left(a^2\mathrm{Re}\left(\frac{1}{|c'|}h'\cdot - i\frac{\overline{c'}}{|c'|}\right)\mathrm{Re}\left(\frac{1}{|c'|}k'\cdot - i\frac{\overline{c'}}{|c'|}\right) + b^2\mathrm{Re}\left(\frac{1}{|c'|}h'\cdot\frac{\overline{c'}}{|c'|}\right)\mathrm{Re}\left(\frac{1}{|c'|}k'\cdot\frac{\overline{c'}}{|c'|}\right)\right) \\ |c'|\mathrm{d}t &= \int_I \frac{1}{|c'|^3}\left(a^2\mathrm{Im}\left(c'\overline{h'}\right)\mathrm{Im}\left(c'\overline{k'}\right) + b^2\mathrm{Re}\left(c'\overline{h'}\right)\mathrm{Re}\left(c'\overline{k'}\right)\right)\mathrm{d}t\,. \end{aligned} \tag{2.4}$$

Next we compute a directional derivative of $F_{a,b}$ at c. Here we use the idea that $F_{a,b}$ is a well-defined map in a small C^∞ neighborhood of c to do the computation without having to deal with the map being potentially multiple-valued. Using the formula

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0}|c'+\epsilon h'| = \frac{\mathrm{Re}\left(c'\overline{h'}\right)}{|c'|},$$

we have

$$\begin{aligned} DF_{a,b}(c)(h) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} 2b|c'+\epsilon h'|^{1/2}\left(\frac{c'+\epsilon h'}{|c'+\epsilon h'|}\right)^{\frac{a}{2b}} \\ &= \left(\frac{c'}{|c'|}\right)^{\frac{a}{2b}}\left(b|c'|^{-1/2}\frac{\mathrm{Re}\left(c'\overline{h'}\right)}{|c'|} + a|c'|^{1/2}\frac{|c'|}{c'}\frac{|c'|h' - c'\mathrm{Re}\left(c'\overline{h'}\right)/|c'|}{|c'|^2}\right) \\ &= \left(\frac{c'}{|c'|}\right)^{\frac{a}{2b}}|c'|^{-\frac{3}{2}}\left(b\mathrm{Re}\left(c'\overline{h'}\right) - ia\mathrm{Im}\left(c'\overline{h'}\right)\right), \end{aligned}$$

where our formula for the directional derivative at $c \in \mathrm{Imm}(I,\mathbb{C})$ is justified because the space of immersions is an open submanifold of the vector space $C^\infty(I,\mathbb{C})$. Then the pullback metric is given by

$$\begin{aligned} \left(F_{a,b}^* g^{L^2}\right)_c(h,k) &= \int_I \mathrm{Re}\ DF_{a,b}(c)(h)\cdot\overline{DF_{a,b}(c)(k)}\mathrm{d}t \\ &= \int_I |c'|^{-3}\mathrm{Re}\left(b\mathrm{Re}\left(c'\overline{h'}\right) - ia\mathrm{Im}\left(c'\overline{h'}\right)\right)\cdot\left(b\mathrm{Re}\left(c'\overline{k'}\right) + ia\mathrm{Im}\left(c'\overline{k'}\right)\right)\mathrm{d}t, \end{aligned}$$

which easily simplifies to (2.4). ■

We will show in section 2.5 that this theorem is a direct generalization of results appearing in [5, 24, 55].

2.4. The preshape space of closed curves.

We now consider the preshape space of closed loops $\mathrm{Imm}\left(S^1,\mathbb{C}\right)/\mathrm{Tra}$ (i.e., the space of "object outlines"). By identifying S^1 with the quotient $[0,1]/(0\sim 1)$, we can consider the preshape space of closed curves to be a submanifold of the preshape space of open curves of infinite codimension. Under this identification, the $F_{a,b}$-transform can be restricted to $\mathrm{Imm}\left(S^1,\mathbb{C}\right)/\mathrm{Tra}$ and the image of the restricted map will lie in $C^\infty(I,\mathbb{C}^*)$. We wish to characterize the image of the restricted $F_{a,b}$-transform. Using the polar form (2.2) of $F_{a,b}$, we see that for any closed curve c with $c' = r\exp(i\theta)$,

$$\begin{aligned} F_{a,b}(c)(1) &= 2b\sqrt{r(1)}\exp\left(i\frac{a}{2b}\theta(1)\right) \\ &= 2b\sqrt{r(0)}\exp\left(i\frac{a}{2b}(\theta(0)+\theta(1)-\theta(0))\right) \\ &= F_{a,b}(c)(0)\cdot\exp\left(i\frac{a}{b}\pi\mathrm{ind}(c)\right), \end{aligned}$$

where $\mathrm{ind}(c)$ is the Whitney rotation index of the immersed curve c [50]. Observe that this computation does not depend on our choice of polar representation of c'; indeed, if $\tilde{\theta}=\theta+2k\pi$, then $\tilde{\theta}(1)-\tilde{\theta}(0)=\theta(1)-\theta(0)$. It follows that a necessary condition for a complex curve $q\in C^\infty(I,\mathbb{C}^*)$ to be the image of a closed curve under $F_{a,b}$ is that there exists some integer ℓ such that

$$q^{(k)}(1) = q^{(k)}(0)\cdot\exp\left(i\frac{a}{b}\pi\ell\right) \tag{2.5}$$

for all integers $k\in\mathbb{Z}_{\geqslant 0}$. We denote by $V_{a,b}(\ell)$ the codimension-∞ vector subspace of $C^\infty(I,\mathbb{C})$ containing curves q with property (2.5). Let $V^*_{a,b}(\ell)=V_{a,b}(\ell)\cap C^\infty(I,\mathbb{C}^*)$, let $V_{a,b}$ denote the union of all $V_{a,b}(\ell)$, and let $V^*_{a,b}$ denote the union of all $V^*_{a,b}(\ell)$. If we restrict our attention to simple closed curves (i.e., those curves with no self-intersections), then we are only interested in the vector space $V_{a,b}(1)$ and the open submanifold $V^*_{a,b}(1)$.

The discussion above captures the higher order C^k closure conditions for c, but not the two-dimensional C^0 closure condition $c(0)=c(1)$. In fact, the image of $F_{a,b}$ is locally a codimension-2 submanifold of $V_{a,b}$. In order to perform calculations for closed curves, it will be useful to characterize the two-dimensional normal space to this submanifold. Consider the function $f_{a,b}:V^*_{a,b}\to\mathbb{C}$ defined by

$$f_{a,b}(q) := F^{-1}_{a,b}(q)(1) = \frac{1}{4b^2}\int_I |q|^2\left(\frac{q}{|q|}\right)^{\frac{2b}{a}}\,\mathrm{d}t\,. \tag{2.6}$$

The image of $\mathrm{Imm}\left(S^1, \mathbb{C}\right)/\mathrm{Tra}$ in $V^*_{a,b}$ is exactly the set $f^{-1}_{a,b}(0)$. We wish to calculate the gradient to $f_{a,b}(q)$ for q in this submanifold.

The derivative of $f_{a,b}$ at q in the direction of a variation p is given by

$$Df_{a,b}(q)(p) = \frac{1}{2b^2}\int_I \left(\frac{q}{|q|}\right)^{2b/a}\left(\mathrm{Re}(q\bar{p}) - i\frac{b}{a}\mathrm{Im}(q\bar{p})\right)\mathrm{d}t\,.$$

The normal space to the submanifold of closed curves is spanned by the gradients of the real and imaginary parts of $f_{a,b}$. The real component of $Df_{a,b}(q)(p)$ is given by

$$\begin{aligned}\mathrm{Re}\,Df_{a,b}(q)(p) &= \frac{1}{2b^2}\int_I \mathrm{Re}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right)\mathrm{Re}(q\bar{p}) + \frac{b}{a}\mathrm{Im}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right)\mathrm{Im}(q\bar{p})\,\mathrm{d}t \\ &= \int_I \mathrm{Re}\left[\frac{1}{2b^2}\left(\mathrm{Re}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right) - i\frac{b}{a}\mathrm{Im}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right)\right)q\bar{p}\right]\mathrm{d}t,\end{aligned}$$

and it follows that

$$\mathrm{grad}\left(\mathrm{Re}\left(f_{a,b}\right)\right)_q = \frac{1}{2b^2}\left(\mathrm{Re}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right) - i\frac{b}{a}\mathrm{Im}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right)\right)q\,. \tag{2.7}$$

Similarly,

$$\mathrm{grad}\left(\mathrm{Im}\left(f_{a,b}\right)\right)_q = \frac{1}{2b^2}\left(\mathrm{Im}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right) + i\frac{b}{a}\mathrm{Re}\left(\left(\frac{q}{|q|}\right)^{2b/a}\right)\right)q\,. \tag{2.8}$$

An important tool for shape analysis of closed curves is the projection from the preshape space of open curves into the preshape space of closed curves. One cannot compute this projection analytically, but the above characterization of the submanifold of closed curves under the $F_{a,b}$-transform allows us to use a gradient descent algorithm for this purpose—see Figure 2 for a few examples. The algorithm itself is similar in spirit to the one described in [44] for the SRVF transform; we provide an outline of the algorithm without focusing on details for brevity. The algorithm for projecting an open curve q into the preshape space of closed curves follows four steps: (1) compute the Jacobian matrix $J_{i,j} = \delta_{i,j} + 3\int_{S^1} q_i q_j \mathrm{d}t$, $i, j = 1, 2$, where i, j denote the first and second coordinates of q; (2) compute the residual using (2.6) and solve $J\beta = -f_{a,b}(q)$; (3) update $q = q + \epsilon\sum_{i=1}^{2}\beta_i b_i$, where $\epsilon > 0$ is a small (emperically chosen) step size and the basis functions b_1, b_2 are given in (2.7) and (2.8); and (4) rescale q such that its norm is $2b$. We repeat steps (1)–(4) until the residual computed in step (2) becomes small.

2.5. Relation to previous work.

The family of maps $F_{a,b}$ includes transforms which have already appeared in the literature. Indeed,

$$F_{1,\frac{1}{2}}(c) = |c'|^{1/2}\frac{c'}{|c'|} = \frac{c'}{|c'|^{1/2}},$$

so that $F_{1,\frac{1}{2}}$ yields the SRVF transform introduced in [24]. We also have

$$F_{\frac{1}{2},\frac{1}{2}}(c) = |c'|^{1/2}\left(\frac{c'}{|c'|}\right)^{1/2} = \sqrt{c'},$$

and we see that $F_{\frac{1}{2},\frac{1}{2}}$ is the complex square-root transform studied in [55].

The SRVF was already shown to be a special case of a general family of transforms in [5]. There, the authors define a two-parameter family of transforms $R_{a,b}$ for $2b \geqslant a > 0$ by

$$\begin{aligned} R_{a,b}&: \mathrm{Imm}\left(I,\mathbb{R}^2\right)/\mathrm{Tra} \to C^\infty\left(I,\mathbb{R}^3\right) \\ c &\mapsto |c'|^{1/2}\left(\begin{pmatrix}T\\0\end{pmatrix} + \sqrt{4b^2-a^2}\begin{pmatrix}0\\1\end{pmatrix}\right), \end{aligned}$$

where $T = D_s c$. The image of the $R_{a,b}$-transform is an open subset of a cone given by

$$C_{a,b} = \left\{(x,y,z) \in \mathbb{R}^3 \mid \left(4b^2-a^2\right)\left(x^2+y^2\right) = a^2z^2, z > 0\right\}.$$

The limiting cone $C_{1,\frac{1}{2}}$ can clearly be identified with $\mathbb{R}^2$, and then $R_{1,\frac{1}{2}}(c)$ gives the SRVF transform of c. It is shown in [5] that the $R_{a,b}$-transform pulls back the L^2 metric on $C^\infty\left(I,\mathbb{R}^3\right)$ to the elastic metric $g^{a,b}$.

We claim that the $R_{a,b}$-transforms correspond to $F_{a,b}$-transforms when $2b \geqslant a$, so that our main result generalizes [5] to work for all parameter choices. Indeed, writing $\mathbb{R}^3 \approx \mathbb{C} \times \mathbb{R}$, there is a projection map $\mathbb{C}^* \to C_{a,b}$ defined for each parameter choice with $2b \geqslant a > 0$ in polar coordinates by

$$(r,\theta) \mapsto \left(\frac{a}{2b}r\cos\left(\frac{2b}{a}\theta\right), \frac{a}{2b}r\sin\left(\frac{2b}{a}\theta\right), \frac{\sqrt{4b^2-a^2}}{2b}r\right), \tag{2.9}$$

which extends to a local isometry $p_{a,b}: C^\infty(I,\mathbb{C}^*) \to C^\infty\left(I,C_{a,b}\right)$ with respect to the corresponding L^2 metrics. Expressing the map (2.9) in the form

$$w \mapsto \left(\frac{a|w|}{2b}\left(\frac{w}{|w|}\right)^{2b/a}, \frac{\sqrt{4b^2-a^2}}{2b}|w|\right),$$

it is easy to see that $R_{a,b} = p_{a,b} \circ F_{a,b}$.

In fact, this idea can be extended to all parameter values. The cones $C_{a,b}$ can be understood in terms of *Regge cones*; these are building blocks of the Regge calculus used to approximate Riemannian manifolds in theoretical physics [39]. One constructs a two-dimensional Regge cone from polar coordinates $(r, \theta) \in \mathbb{R}^2_{\geqslant 0}$ with standard metric $\mathrm{d}r^2 + r^2\mathrm{d}\theta^2$ by identifying points according to the relation $(r, \theta_1) \sim (r, \theta_2) \Leftrightarrow |\theta_1 - \theta_2| = 2\pi - \theta_0$ for some choice of *deficit angle* θ_0 (allowed to be positive or negative). The map (2.9) is an isometric embedding of the Regge cone with deficit angle $\theta_0 = 1 - \frac{a}{2b}$ onto a flat cone in Euclidean space. For parameters with $4b^2 < a^2$, replacing the trigonometric functions with their hyperbolic counterparts and the coefficient in the third coordinate with $\frac{\sqrt{a^2 - 4b^2}}{2b}$ yields an isometric map of the Regge cone with (negative) deficit angle θ_0 onto a flat cone in Lorentz space (see [20]).

3. Shape preserving group actions.

3.1. Rotation actions and fibers of $F_{a,b}$.

We now treat the fact that $F_{a,b}$ is multivalued for certain parameter choices and noninjective for others. The fibers of $F_{a,b}$ are closely related to the actions of the rotation group SO(2) on $\mathrm{Imm}(I, \mathbb{C})/\mathrm{Tra}$ and $C^\infty(I, \mathbb{C}^*)$. Using the natural identification of SO(2) with S^1, we can represent the rotation actions as complex multiplication. That is, we express rotations in the respective spaces as $\exp(i\psi)c$ and $\exp(i\psi)q$, where $\exp(i\psi) \in S^1$. As usual, multiplication in these formulas is performed pointwise as complex numbers. We have the following correspondence between the actions, which follows by an elementary computation.

Lemma 3.1. *Let* $c \in \mathrm{Imm}(I, \mathbb{C})/\mathrm{Tra}$ *and* $\exp(i\psi) \in \mathrm{SO}(2)$. *Then for all* $a, b > 0$,

$$F_{a,b}(\exp(i\psi)c) = \exp\left(i\psi\frac{a}{2b}\right)F_{a,b}(c).$$

Theorem 3.2. *The* $F_{a,b}$*-transform induces a well-defined isometry*

$$\mathrm{Imm}(I, \mathbb{C})/\{\mathrm{Tra}, \mathrm{Rot}\} \to C^\infty(I, \mathbb{C}^*)/\mathrm{Rot}$$

with respect to the metrics induced by $g^{a,b}$ *and* g^{L^2}, *respectively.*

We abuse notation and continue to denote the induced isometry by $F_{a,b}$. The induced map is defined by

$$F_{a,b}([c]) = [F_{a,b}(c)], \tag{3.1}$$

where we use brackets to denote the SO(2)-orbit of a parameterized plane curve. The right side denotes the equivalence class of *any* branch of $F_{a,b}(c)$ in the case that the map is multivalued.

Proof. We first note that all $F_{a,b}$-images of a curve c in the list (2.1) are related by rotations, so that $[F_{a,b}(c)]$ is a well-defined element of $C^\infty(I,\mathbb{C}^*)/\mathrm{Rot}$. Lemma 3.1 then implies that the induced map is well-defined. Similar arguments hold for the obvious map induced by the local inverse $F_{a,b}^{-1}$, giving a well-defined inverse map

$$F_{a,b}^{-1}: C^\infty(I,\mathbb{C}^*)/\ \mathrm{Rot}\ \to \mathrm{Imm}(I,\mathbb{C})/\left\{\ \mathrm{Tra},\ \mathrm{Rot}\ \right\},$$

so the induced map $F_{a,b}$ is a bijection. Finally, it is easy to see that $g^{a,b}$ and g^{L^2} are invariant under the action of SO(2). Therefore, the local isometry of Theorem 2.3 descends to a global isometry on the quotient spaces. ■

We have the following immediate corollary.

Corollary 3.3. *The $F_{a,b}$-transform induces a well-defined isometric embedding*

$$\mathrm{Imm}\left(S^1,\mathbb{C}\right)/\left\{\ \mathrm{Tra},\ \mathrm{Rot}\ \right\} \hookrightarrow V_{a,b}^*/\ \mathrm{Rot},$$

where $V_{a,b}^*$ is the space of curves defined in section 2.4.

3.2. The scaling action.

The group of positive real numbers $\mathbb{R}_{>0}$ acts on a parameterized curve by uniform scaling. An easy calculation shows that the scaling action interacts with the $F_{a,b}$-transform as follows: for $\lambda \in \mathbb{R}_{>0}$ and $c \in \mathrm{Imm}(I,\mathbb{C})/\mathrm{Tra}$,

$$F_{a,b}(\lambda c) = \lambda^{1/2} F_{a,b}(c)\,.$$

It will be convenient to represent the quotient of the preshape space by this scaling action as

$$\mathrm{Imm}(I,\mathbb{C})/\{\ \mathrm{Tra},\ \mathrm{Sca}\ \} \approx \{c \in \mathrm{Imm}(I,\mathbb{C}) \mid c(0)=0,\ \mathrm{length}\ (c)=1\}\,. \tag{3.2}$$

For $\Sigma = I$ or S^1 and $a,\ b > 0$, define the *Hilbert sphere of radius r* to be the space

$$\mathscr{H}_{a,b}^{\Sigma}(r) = \begin{cases} \left\{ q \in C^\infty(I,\mathbb{C}) \mid \int_\Sigma |q|^2\ \mathrm{d}t = r^2 \right\}, & \Sigma = I, \\ \left\{ q \in V_{a,b} \mid \int_\Sigma |q|^2\ \mathrm{d}t = r^2 \right\}, & \Sigma = S^1\,. \end{cases}$$

Proposition 3.4. *The $F_{a,b}$-transform sends* $\mathrm{Imm}(\Sigma,\mathbb{C})/\{\mathrm{Tra},\ \mathrm{Sca}\}$ *into the Hilbert sphere* $\mathscr{H}_{a,b}^{\Sigma}(2b)$. *It induces an isometry between* $\mathrm{Imm}(\Sigma,\mathbb{C})/\{\mathrm{Tra},\ \mathrm{Sca},\ \mathrm{Rot}\}$ *and its image in* $\mathscr{H}_{a,b}^{\Sigma}(2b)/\mathrm{Rot}$.

Proof. Let $c \in \mathrm{Imm}(\Sigma,\mathbb{C})$ have length one. Using (2.2), we have

$$\int_{\Sigma} |F_{a,b}(c)|^2 \, \mathrm{d}t = \int_{\Sigma} \left(2b|c'|^{1/2}\right)^2 \mathrm{d}t = 4b^2 \int_{\Sigma} |c'| \mathrm{d}t = 4b^2 .$$

The second statement follows by noting that $\mathscr{H}^{\Sigma}_{a,b}(2b)$ is invariant under the rotation action of SO(2) so that we can restrict the isometries of Theorem 3.2 and Corollary 3.3. ■

3.3. The reparameterization action.

The final shape-preserving group action to consider is the action of $\mathrm{Diff}^+(I)$ on $\mathrm{Imm}(I, \mathbb{C})$ by orientation-preserving reparameterizations. An element $\gamma \in \mathrm{Diff}^+(I)$ of the diffeomorphism group also acts on $q \in C^{\infty}(I, \mathbb{C}^*)$ by the formula

$$\gamma * q := \sqrt{|\gamma'|}(q \circ \gamma) . \tag{3.3}$$

Proposition 3.5. *The map $F_{a,b}$ is equivariant with respect to the* $\mathrm{Diff}^+(I)$*-actions on* $\mathrm{Imm}(I, \mathbb{C})$ *and* $C^{\infty}(I, \mathbb{C}^*)$ *defined above.*

Proof. Let $c \in \mathrm{Imm}(\Sigma, \mathbb{C})$ and $\gamma \in \mathrm{Diff}^+(\Sigma)$. Then

$$\begin{aligned} F_{a,b}(c \circ \gamma) &= 2b|\gamma'(c' \circ \gamma)|^{1/2} \left(\frac{\gamma'(c' \circ \gamma)}{|\gamma'(c' \circ \gamma)|}\right)^{\frac{a}{2b}} \\ &= |\gamma'|^{1/2} 2b|c' \circ \gamma|^{1/2} \left(\frac{c' \circ \gamma}{|c' \circ \gamma|}\right)^{\frac{d}{2b}} = |\gamma'|^{1/2} F_{a,b}(c) \circ \gamma . \end{aligned}$$

■

We likewise wish to consider the action of $\mathrm{Diff}^+(S^1)$ on $\mathrm{Imm}\left(S^1, \mathbb{C}\right)$ and the corresponding action in transform space. To understand the reparameterization action in transform space for closed curves, it is convenient to identify $V_{a,b}(\ell)$ with the vector space

$$\widetilde{V}_{a,b}(\ell) = \left\{ q \in C^{\infty}(\mathbb{R}, \mathbb{C}^*) \mid q(t+1) = q(t) \cdot \exp\left(i \frac{a}{b} \pi \ell\right) \right\} .$$

Under this identification, the $\mathrm{Diff}^+(S^1)$-action on transform space is once again given by (3.3) in the sense that a similar equivariance result holds in this setting.

4. Geodesics between curves.

4.1. Geodesics for open curves.

For open curves, Theorem 3.2 provides an isometry of Riemannian manifolds

$$F_{a,b} : \left(\mathrm{Imm}(I, \mathbb{C}) / \left\{\mathrm{Tra}, \mathrm{Rot}\right\}, g^{a,b}\right) \to \left(C^{\infty}(I, \mathbb{C}^*) / \mathrm{Rot}, g^{L^2}\right) .$$

We can therefore compute geodesics between curves in the former space by translating the problem to the simpler target space, $C^\infty(I, \mathbb{C})$, where geodesic paths are simply straight lines. Passing to the open subset $C^\infty(I, \mathbb{C}^*)$, we lose geodesic completeness as the geodesic joining $q_0, q_1 \in C^\infty(I, \mathbb{C}^*)$ given by $q_u = (1 - u)q_0 + uq_1$, $u \in [0, 1]$, may pass through a curve with $q_u(t) = 0$ for some t. Nonetheless, geodesic distance in the larger space still induces a metric on the restricted space, and curves which are sufficiently close will be joined by a well-defined geodesic (see [13, section 3.4] for a discussion of this phenomenon in the SRVF framework). The geodesic distance between a pair of curves in the geodesic completion of the image space is

$$d_{L^2}(q_0, q_1) = \|q_0 - q_1\|_{L^2} = \left(\int_I |q_0 - q_1|^2 \, dt\right)^{1/2}.$$

Geodesics in the quotient $\mathrm{Imm}(I, \mathbb{C})/\{\mathrm{Tra}, \mathrm{Rot}\} \approx C^\infty(I, \mathbb{C}^*)/\mathrm{Rot}$ are realized as geodesics between curves in the total space $C^\infty(I, \mathbb{C})$ after a preprocessing step whereby the curves are aligned over SO(2) using a standard algorithm called Procrustes analysis (essentially a singular value decomposition problem). Furthermore, we calculate explicit geodesics in $\mathrm{Imm}(I, \mathbb{C})/\{\mathrm{Tra}, \mathrm{Sca}, \mathrm{Rot}\}$ by using Proposition 3.4 to transfer the problem to the Hilbert sphere. The geodesic joining a pair of curves $q_0, q_1 \in \mathscr{H}^I_{a,b}(2b)$ is given by spherical interpolation,

$$q_u = \frac{1}{\sin(D)}(\sin((1 - u)D)q_0 + \sin(uD)q_1),$$

where $D = \arccos\left(\langle q_0, q_1 \rangle_{L^2}/4b^2\right)$ is the geodesic distance in the Hilbert sphere. Geodesics in the quotient $\mathscr{H}^I_{a,b}(2b)/\mathrm{Rot}$ are treated by the same optimization procedure as in the flat case.

4.2. Geodesics for closed curves.

Geodesics in $\mathrm{Imm}\left(S^1, \mathbb{C}\right)/\left\{\mathrm{Tra}, \mathrm{Rot}\right\}$ can be treated in a similar manner to the open case; that is, we transfer the problem to the simpler space $V_{a,b}/\{\mathrm{Rot}\}$. Each vector space $V_{a,b}(\ell)$ is flat, so its geodesics are straight lines. However, the fact that the image of the isometry induced by $F_{a,b}$ is codimension-2 in $V_{a,b}$ makes the procedure more complicated.

We first note that the problem has a remarkable simplification in the case that $a = b$, where the closure condition $f_{a,b}(q) = 0$ (see (2.6)) reduces to L^2-orthogonality of the coordinate functions of q. This was exploited by Younes et al. in [55] to give explicit geodesics for closed curves by relating the space of closed curves to an infinite-dimensional Stiefel manifold.

Apparently, such a simplification is unique to the case $a = b$, and the space of closed curves is not isometric (at least, not in any obvious way) to a classical manifold otherwise. Fortunately, the low codimension of the space of closed curves in the flat space $V_{a,b}$ allows us to approximate geodesics in the submanifold numerically. There are several algorithms in

the literature which are easily adapted to our setting, such as parallel transport-based path-straightening [44] and other gradient descent-based [5, 52] methods. In the examples provided in section 6, we use a simplistic projection-based algorithm to approximate geodesics, described as follows. Given two closed curves, $q_0, q_1 \in \mathscr{H}^{S^1}_{a,b}(2b)$, we first construct a geodesic in the preshape space of open curves using spherical interpolation as described in section 4.1. Then we perform a pointwise projection of the open preshape space geodesic into the preshape space of closed curves using the numerical algorithm outlined at the end of section 2.4. We leave the development of more sophisticated algorithms for computing geodesics in the preshape space of closed curves for future work.

4.3. Optimized geodesics in the shape spaces.

To compute geodesics in the shape space of unparameterized curves with respect to the metric induced by $g^{a,b}$, we pass to the quotient of the parameterized curve space by the action of the diffeomorphism group $\mathrm{Diff}^+(I)$ ($\mathrm{Diff}^+(S^1)$ in the case of closed curves). The geodesic between $\mathrm{Diff}^+(I)$-orbits $[c_1]$ and $[c_2]$ of parameterized curves c_j is realized in practice as the geodesic between c_1 and $\tilde{c}_2$ in the total space $\mathrm{Imm}(I, \mathbb{C})$, where $\tilde{c}_2 = c_2 \circ \gamma$ and

$$\gamma = \mathrm{arginf}\left\{\mathrm{dist}_{g^{a,b}}(c_1, c_2(\gamma)) \mid \gamma \in \mathrm{Diff}^+(I)\right\}. \tag{4.1}$$

Here $\mathrm{dist}_{g^{a,b}}$ denotes geodesic distance with respect to $g^{a,b}$. In general, the reparameterization realizing this infimum may fail to be smooth (see [55, section 4.2]), whence the geodesic is actually realized in the larger space of L^2 curves. Precise characterizations of the regularity of solutions to the optimization problem for SRVF parameters $g^{1,1/2}$ have been the subject of several recent articles, and it has been shown that for C^1 input curves, the optimal reparameterization γ is achieved and is differentiable almost everywhere [11]. For applications, the realistic setting considers piecewise linear (PL) curves (defined precisely in the following section), and it is known that in that case, the optimal reparameterization is realized and is also PL [31]. We adapt the methods of [31] to the PL setting for general elastic metrics $g^{a,b}$ in section 5.

A major benefit of our results is that under the $F_{a,b}$-transform, the optimal reparameterization problem (4.1) becomes equivalent to the optimization problem which appears under the SRVF formalism. We are therefore able to utilize existing, highly efficient numerical approaches to approximate solutions of (4.1) (for example, the dynamic programming approach [40]). Once an approximate solution γ is obtained, we are able to easily compute geodesics in the shape space using the techniques of section 4.1 applied to c_1 and the reparameterized curve $\tilde{c}_2$.

The approach described above can be adapted to provide geodesics in the space of curves of fixed length (using the Hilbert sphere geodesics of section 4.1, together with the simple relationship between Euclidean and spherical distance) and geodesics in the space of closed curves (by numerically optimizing over $\mathrm{Diff}^+(I) \times S^1$, where the S^1 factor corresponds to a search for optimal seed points between the two curves).

5. Extending to piecewise linear shapes.

5.1. Problem setup.

A natural question of both theoretical and practical interest is whether the previous results can be extended to spaces of curves of lower regularity. Indeed, the methods above can be used, essentially without modification, to show that the map $F_{a,b}$ can be extended to give an isometry

$$\left(\mathrm{Imm}^1(I,\mathbb{C})/\left\{\mathrm{Tra},\mathrm{Rot}\right\}, g^{a,b}\right) \to \left(C^0(I,\mathbb{C}^*), g^{L^2}\right),$$

where Imm^1 denotes the space of C^1 immersions and $g^{a,b}$ is the appropriately extended elastic metric.

It is common in the literature on the SRVF transform to consider the space $\mathrm{AC}\left(I,\mathbb{R}^N\right)$ of absolutely continuous curves in Euclidean space (*absolutely continuous* curves can be characterized as those curves which are continuous everywhere and differentiable almost everywhere [41]). Recall that the transform $F_{1,\frac{1}{2}}$ recovers the SRVF transform for smooth, immersed plane curves. This map extends to a well-defined map,

$$F_{1,\frac{1}{2}}:\mathrm{AC}(I,\mathbb{C}) \to L^2(I,\mathbb{C}),$$
$$c(t) \mapsto \begin{cases} \frac{c'(t)}{\sqrt{|c'(t)|}}, & c'(t) \text{ exists and is nonzero,} \\ 0 & \text{otherwise,} \end{cases}$$

which is a homeomorphism that pulls back g^{L^2} to $g^{1,\frac{1}{2}}$ at smooth points. Theoretical details of this construction are examined in [31, 11].

One would like to similarly extend the remaining $F_{a,b}$-transforms to spaces of curves of low regularity, but this causes immediate issues. For $a/2b \neq 1$, one of $F_{a,b}$ or its inverse is multivalued pointwise, since the map involves complex exponentiation. In previous sections, we relied on the smoothness of our curves (or at least continuity of derivatives) to choose complex roots coherently in order to ensure $F_{a,b}$ was well-defined up to rotations.

5.2. Extended $F_{a,b}$-transform.

Inspired by our results for smooth curves, we can extend our work to one of the most important spaces of curves from a practical standpoint: piecewise linear curves. Let $\mathrm{PL}(I,\mathbb{C})$ denote the space of PL planar curves; that is, each $c \in \mathrm{PL}(I,\mathbb{C})$ is a continuous curve such that there is a decomposition

$$I = I_1 \cup I_2 \cup \cdots \cup I_k = [t_0 = 0, t_1] \cup [t_1, t_2] \cup \cdots \cup [t_{k-1}, t_k = 1]$$

with $c'(t) = v_j \in \mathbb{C}$ for all $t \in (t_{j-1}, t_j)$, $j = 1, \ldots, k$. We call the points $c(t_j)$ *vertices* of c and the t_j are called *vertex parameters*. A PL curve c is called a *piecewise linear immersion* if $|c'(t)|$

$\neq 0$, where $c'(t)$ is defined. Let $\mathrm{PL}(I, \mathbb{C})/\mathrm{Tra}$ denote the space of PL curves modulo translations, which we identify with the set of curves based at the origin.

We define the *extended $F_{a,b}$-transform* to be the map

$$F_{a,b}(c) = 2br^{1/2}\exp\left(i\frac{a\theta}{2b}\right),$$

where r is the piecewise constant function

$$r(t) = \begin{cases} |v_1| & \text{for } t \in [0, t_1], \\ |v_j| & \text{for } t \in (t_{j-1}, t_j], j = 2, \ldots, k, \end{cases} \tag{5.1}$$

and θ is the piecewise constant function defined recursively by

$$\theta(t) = \begin{cases} \theta_1 = \arctan\dfrac{\mathrm{Im}(v_1)}{\mathrm{Re}(v_1)} & \text{for } t \in [0, t_1], \\ \theta_{j-1} + s_j \cdot \delta\theta_j & \text{for } t \in (t_{j-1}, t_j], j = 2, \ldots, k\,. \end{cases} \tag{5.2}$$

The term $\delta\theta_j \in [0, \pi]$ is the *jth exterior angle* between the $(j-1)$th and jth edges of c and is given by

$$\delta\theta_j = \arccos \mathrm{Re}\left(\frac{v_{j-1}\overline{v_j}}{|v_{j-1}\overline{v_j}|}\right). \tag{5.3}$$

To simplify notation later on, we set $\delta\theta_1 = \theta_1$. The coefficient $s_j \in \{-1, 1\}$ describes the *orientation* of the jth exterior angle and is given by

$$s_j = \mathrm{sign}\left(\mathrm{Im}\left(v_{j-1}\overline{v_j}\right)\right), \tag{5.4}$$

where we define the sign function for any real number a according to the convention

$$\mathrm{sign}(a) := \begin{cases} 1 & \text{if } a > 0, \\ -1 & \text{if } a \leqslant 0. \end{cases}$$

We set $s_1 = 1$.

Lemma 5.1. *Let c be a smooth immersion with a fixed representation of c' in polar coordinates, $c' = re^{i\theta}$, so that $\theta(0) = \arctan\frac{\mathrm{Im}(c'(0))}{\mathrm{Re}(c'(0))}$ and θ is continuous. Let $\{c^n\}$ be a sequence of PL immersions with vertices sampled from c (i.e., each c^n is a secant approximation of c). Assume that the sequence $\{c^n\}$ converges uniformly to c in C^1, and let (r^n, θ^n) be polar coordinates for c^n, obtained by formulas* (5.1) *and* (5.2)*. Then $r^n \to r$ and $\theta^n \to \theta$ uniformly.*

Proof. The assumption of uniform convergence $(c^n)'(t) \to c'(t)$ immediately implies $r^n(t) \to r(t)$ uniformly, and one only needs to check convergence of the angle functions. Let v_j^n denote the derivative vectors of c^n, and let $\hat{v}_j^n$ denote the derivative vectors of c sampled at

vertex parameters of c^n. Moreover, let $\hat{\theta}_j^n$ denote samples of the function θ at the vertex parameters of c^n.

Fix a small $\epsilon > 0$, and in particular assume that $\epsilon < \frac{\pi}{3}$. Our assumption that $(c^n)' \to c'$ uniformly implies that there exists N such that for all $n > N$ we may choose a collection of angles $\{\tilde{\theta}_j^n\}$ satisfying $v_j^n = r_j^n \exp(i\,\tilde{\theta}_j^n)$ and $|\tilde{\theta}_j^n - \hat{\theta}_j^n| < \epsilon$ for all j. Increasing N if necessary, we may also assume that the following estimates hold for all j:

- $|\theta_j^n - \theta_{j-1}^n| < \pi$ (by the definition of θ^n), and
- $|\hat{\theta}_j^n - \hat{\theta}_{j-1}^n| < \epsilon$ (by continuity of c).

We claim that $\tilde{\theta}_j^n = \theta_j^n$ for all j. By the definition of θ_j^n, the contrary would imply that for some j we have

$$\pi < |\tilde{\theta}_j^n - \tilde{\theta}_{j-1}^n| \leq |\tilde{\theta}_j^n - \hat{\theta}_j^n| + |\tilde{\theta}_{j-1}^n - \hat{\theta}_{j-1}^n| + |\hat{\theta}_j^n - \hat{\theta}_{j-1}^n| < \pi,$$

and we have arrived at a contradiction. Thus $|\theta_j^n - \hat{\theta}_j^n| < \epsilon$ for all j, and this completes the proof. ■

Remark 5.2. Sequences of approximating PL curves as described in Lemma 5.1 arise by linear spline approximation. For a smooth $c: [0, 1] \to \mathbb{R}^2$, we pick a mesh $t_0 < t_1 < \cdots < t_n$ and define a linear spline c_n interpolating between the $c(t_j)$. As the mesh is refined, the resulting sequence of PL curves converges uniformly to c (see, e.g., Theorems 6.1 and 6.15 of [42]).

Lemma 5.1 immediately implies the following convergence result.

Proposition 5.3. *Let c be a smooth immersed plane curve, and let $\{c^n\}$ be a sequence of PL immersions as in Lemma* 5.1*. Then $F_{a,b}(c^n) \to F_{a,b}(c)$ in L^2.*

Remark 5.4. This shows that if a PL curve approximates a smooth curve, then taking sufficiently many samples produces a transformed curve, which is close to the transformed smooth curve. On the other hand, if the PL curve is truly meant to contain jagged angles, then the discrete $F_{a,b}$ is still well-defined but may not be a faithful representation of the curve in transform space. In this scenario, ad hoc methods are necessary to extend the transformation.

5.3. Injectivity.

Theorem 3.2 says that $F_{a,b}$ induces a bijection

$$\mathrm{Imm}(I, \mathbb{C})/\{\mathrm{Tra}, \mathrm{Rot}\} \leftrightarrow C^\infty(I, \mathbb{C}^*)/\,\mathrm{Rot}\,.$$

Unfortunately, the same property is not enjoyed by the extended $F_{a,b}$-transform, as shown by the following example.

Example 5.5. Let c_1 and c_2 be the PL curves given by $c_1'(t) = 1$ and

$$c_2'(t) = \begin{cases} 1, & t \in [0, 1/2], \\ e^{i\theta}, & t \in [1/2, 1], \end{cases}$$

for some fixed choice of $\theta \in (0, \pi]$. Then $F_{a,b}(c_1)(t) = 2b$ and

$$F_{a,b}(c_2)(t) = \begin{cases} 2b, & t \in [1/2, 1], \\ 2b \exp\left(i\theta \frac{a}{2b}\right), & t \in [1/2, 1]. \end{cases}$$

Taking parameters a and b satisfying $\frac{a}{2b} = \frac{2\pi}{\theta}$ yields $F_{a,b}(c_1) = F_{a,b}(c_2)$, while it is clear that c_1 and c_2 do not differ by a rigid rotation.

This potentially causes a major problem when computing distances between PL curves: two PL curves which do not differ by a rotation can receive zero geodesic distance in $F_{a,b}$-transform space. Luckily, the next proposition shows that this situation is highly nongeneric and that it does not arise in most applications.

Proposition 5.6. *Let c_1 and c_2 be PL immersions with k segments whose derivatives are represented in polar coordinates by the functions (r^1, θ^1) and (r^2, θ^2), respectively, as defined by formulas* (5.1) *and* (5.2)*. The images of the curves under the extended $F_{a,b}$-transform are the same if and only if $r^1(t) = r^2(t)$ for all t and there exist integers ℓ_j such that*

$$s_j^1 \delta\theta_j^1 = s_j^2 \delta\theta_j^2 + \frac{4b}{a} \ell_j \pi$$

for all $j = 1, \ldots, k$.

Proof. Assume that $F_{a,b}(c_1) = F_{a,b}(c_2)$. Then $|F_{a,b}(c_1)(t)| = |F_{a,b}(c_2)(t)|$ holds for all t, and it follows that $r^1(t) = r^2(t)$ for all t. Since $r^j(t) \neq 0$, we have that $\exp\left(\frac{a\theta^1}{2b}\right) = \exp\left(\frac{a\theta^2}{2b}\right)$. This equality holds if and only if

$$\frac{a\theta_j^1}{2b} = \frac{a\theta_j^2}{2b} + 2\ell_j \pi$$

for all $j = 1, \ldots, k$ for some integers ℓ_j. Therefore

$$\theta_j^1 = \theta_j^2 + \frac{4b}{a} \ell_j \pi . \tag{5.5}$$

Setting $j = 1$ and recalling that we are using the convention $\delta\theta_1 = \theta_1$ and $s_1 = 1$ provides our first condition on the θ-functions. The $j = 2$ instance of (5.5) reads

$$\theta_2^1 = \theta_2^2 + \frac{4b}{a}\ell_2\pi \Leftrightarrow \theta_1^1 + s_2^1\delta\theta_2^1 = \theta_1^2 + s_2^2\delta\theta_2^2 + \frac{4b}{a}\ell_2\pi$$
$$\Leftrightarrow s_2^1\delta\theta_2^1 = s_2^2\delta\theta_2^2 + \frac{4b}{a}(\ell_2 - \ell_1)\pi\,.$$

The claim then follows in the $j = 2$ case after relabeling the integer coefficient of $\frac{4b}{a}\pi$. The general claim follows similarly by induction. ■

It follows that $F_{a,b}$ is *generically injective* on $\mathrm{PL}(I, \mathbb{C})/\{\mathrm{Tra}, \mathrm{Rot}\}$ in the sense that if c_1 and c_2 are PL curves with $F_{a,b}(c_1) = F_{a,b}(c_2)$, then there exists an arbitrarily small perturbation (in the C^1 sense) $\tilde{c}_2$ of c_2 such that the SO(2)-orbit of $F_{a,b}(\tilde{c}_2)$ is different from that of $F_{a,b}(c_1)$. We also have the following corollary, which provides injectivity of $F_{a,b}$ on PL curves with bounded exterior angles.

Corollary 5.7. *The extended map $F_{a,b}$ is injective when restricted to the set of PL immersions with exterior angles uniformly bounded by $\frac{2b}{a}\pi$.*

In particular, note that $F_{a,b}$ is always injective on PL curves when $\frac{a}{2b} \leqslant 1$.

Proof. Let c_1 and c_2 be PL immersions such that $\delta\theta_j^1, \delta\theta_j^2 < \frac{2b}{a}\pi$ for all j. If $F_{a,b}(c_1) = F_{a,b}(c_2)$, then Proposition 5.6 implies that there exist integers ℓ_j such that

$$\left|s_j^1\delta\theta_j^1 - s_j^2\delta\theta_j^2\right| = \left|\frac{4b}{a}\ell_j\pi\right|$$

for all j. We have

$$\left|s_j^1\delta\theta_j^1 - s_j^2\delta\theta_j^2\right| \leqslant \delta\theta_j^1 + \delta\theta_j^2 < \frac{4b}{a}\pi,$$

so that each $\ell_j = 0$. Therefore $\theta^1 = \theta^2$ and it follows that $c_1 = c_2$. ■

For PL curves obtained by sampling a smooth curve, dense enough sampling ensures that exterior angles can be bounded by an arbitrarily small number. It follows that $F_{a,b}$ can be guaranteed to be injective. To quantify this, consider the simple case of a secant approximation of a smooth curve c by an equilateral PL curve with edgelength s. At a vertex of the PL curve, let $\delta\theta$ denote the exterior angle, and let κ denote the curvature of c at that point. Then we have the Taylor approximation (see [1])

$$\delta\theta = s\kappa + O\left(s^3\right).$$

Assuming that the PL curve has N vertices, that c is normalized to have length 1, and that s is therefore roughly equal to $1/N$, it follows that there is an asymptotic bound

$$|\delta\theta| \leqslant \frac{\kappa_{max}}{N} + O\left(1/N^3\right),$$

where κ_{max} is the maximum curvature of c. We therefore obtain the desired bound on the turning angle as soon as

$$N > \frac{a}{2b\pi}\left(\kappa_{max} + O\left(1/N^2\right)\right).$$

5.4. Exact matching.

The algorithm for computing geodesic distances in the shape space $\mathrm{Imm}(I, \mathbb{C})/\mathrm{Diff}^+(I)$ outlined in section 4.3 calls for a solution of the optimization problem (4.1). We now wish to demonstrate the existence of solutions to the corresponding problem in the PL setting. In order to achieve a solution, we replace the smooth diffeomorphism group $\mathrm{Diff}^+(I)$ with the semigroup

$$\bar{\Gamma} = \{\gamma \in \mathrm{AC}(I, I) \mid \gamma(0) = 0, \gamma(1) = 1, \gamma'(t) \geqslant 0 \text{ when } \gamma'(t) \text{ exists } \}.$$

The PL optimization problem seeks the optimal reparameterization γ for PL curves c_1 and c_2 satisfying

$$\gamma = \mathrm{arginf}\left\{\mathrm{dist}_g{}^{a,b}(c_1, c_2(\gamma)) \mid \gamma \in \bar{\Gamma}\right\}. \tag{5.6}$$

Because the $F_{a,b}$-transform induces the same optimization problem for general elastic metrics $g^{a,b}$ as the one studied in the SRVF setting $g^{1,1/2}$, we are able to directly appeal to the recent work of Lahiri, Robinson, and Klassen [31].

Proposition 5.8. *The optimal reparameterization between two PL curves* $[c_1]$ *and* $[c_2]$ *with respect to* $g^{a,b}$ *is realized by a PL function in* $\bar{\Gamma}$.

Proof. The images of the PL curve c_j under $F_{a,b}$ are piecewise constant maps q_j into $\mathbb{C} \approx \mathbb{R}^2$. The reparameterization action of the semigroup $\bar{\Gamma}$ on c_j transforms into an action on the image curves by the formula

$$\gamma * q_j = \sqrt{\gamma'} \cdot q_j \circ \gamma,$$

defined pointwise almost everywhere. The optimization problem (5.6) becomes

$$\inf_{\gamma \in \bar{\Gamma}} \|q_1 - \gamma * q_2\|_{L^2}.$$

It follows from [31, Theorem 5] that this new optimization problem is solved by a PL element of $\bar{\Gamma}$. ■

It is observed in [31] that an optimal matching $\gamma \in \bar{\Gamma}$ between PL curves c_1 and c_2 may contain flat or vertical portions (i.e., the derivative may vanish or there may be discontinuities). Observe that in this case we could find reparameterizations γ_j of both curves c_j so that each γ_j only contains flat portions and no discontinuities. The reparameterized curves $c_j(\gamma_j)$ no longer lie in the space of PL curves, but rather in some slightly more general space where the parameterization can stop at vertices for positive time. In any case, we still obtain geodesics between these mildly singular curves, and the resulting geodesic distance is still a sensible metric on the subspace of PL curves. This behavior is thoroughly explored in [31] for the SRVF setting, and it will be interesting to study it in detail for our general setting in future work.

6. Numerical experiments.

6.1 Implementation issue.

In the case when $\frac{a}{2b} > 1$, the inverse mapping from the $F_{a,b}$-transform space to the space of curves is not guaranteed to produce valid angle functions when the curves are expressed in polar coordinates. This is due to large local differences between the angle functions for the shapes being compared. In these cases, we use a path-straightening algorithm to find the appropriate angle functions along the geodesic path. We omit the details of this algorithm for brevity, but the basic idea is as follows. Plane curves to be compared are first mapped via $F_{a,b}$ by converting to polar coordinates, as discussed above. Optimal alignment and reparameterization is performed on the transform side. For rotational alignment, we use the standard version of Procrustes analysis (a singular value decomposition problem). To find optimal reparameterizations of open curves, we use the dynamic programming algorithm of [40]. To find optimal reparameterizations of closed curves, we also use dynamic programming, but with an additional search for an optimal starting point on each curve (seed search). After optimal rotational alignment and reparameterization, we calculate the polar coordinate r-function along the geodesic explicitly. However, inverting the full $F_{a,b}$-transform along the geodesic, and specifically inverting the polar coordinate θ-function, proves numerically difficult. We thus compute the polar coordinate θ-functions along the geodesic using a gradient descent algorithm based on an objective function given by the elastic metric-based energy of the path in the space of curves. Since the r-functions along the geodesic path can be easily computed, the overall problem simplifies slightly to one where only the first term in (1.1) plays a role in the optimization problem; this results in a simple and computationally efficient path-straightening algorithm that iteratively updates the path according to the gradient of the path energy [44, 29]. We note that the $F_{a,b}$-transform still provides a significant numerical simplification in this case as it allows us to search for optimal rotations and reparameterizations in the transform space under the L^2 metric.

6.2 Examples.

We show several examples of geodesic paths and associated geodesic distances between shapes of open (Figures 3–5) and closed (Figures 6 and 7) curves for different parameter values in the elastic metric, computed via the methods described in sections 4 and 6.1. Each figure shows the curve evolution and the optimal reparameterization in blue with the identity

in red for comparison. While the example in Figure 3 considers two simulated open curves, the examples in Figures 4–7 use curves from the well-known MPEG-7 computer vision database.[2] Comparing within each figure shows the qualitative effect of parameter values on the curve evolutions, while comparing distances within shape classes across figures gives a sense of the range and distinguishing powers of the geodesic distance metrics. One can immediately make some qualitative observations. First, note that in each example the optimal reparameterizations for the $\frac{a}{2b} = 0.5$ and 0.17 parameter choices are quite close to the identity. This was experimentally observed to be typical and makes sense heuristically: as the value of $\frac{a}{2b}$ is decreased, the penalty for stretching deformations in the elastic metric dominates, and bending deformations along the geodesic becomes less costly than reparameterizing to avoid them. Studying this behavior more quantitatively will be the subject of future work. Second, observe that distances increase as the ratio $\frac{a}{2b}$ is decreased. This is explained both by the higher penalty on reparameterizations (hence on matching similar features in the curves), as well as the fact that the (length normalized) curves lie on the Hilbert sphere $H^{\Sigma}_{a,b}(2b)$ after applying $F_{a,b}$ (see Proposition 3.4). The figures were produced by fixing $a = 1$ and varying b, so that decreasing $\frac{a}{2b}$ corresponds to increasing the radius of the sphere containing their transformed images.

We also provide a short classification experiment that shows the benefits of using general weights for the stretching and bending terms in the context of differentiating forgeries from genuine signatures. The data used here consists of 40 different signatures, which are a subset of the SVC 2004 dataset [51]. Each signature class contains 20 genuine signatures and 20 skilled forgeries. We classified each signature as genuine or a forgery using leave-one-out nearest neighbor with respect to the geodesic distance under the $g^{a,b}$ metric for various parameter choices (a,b). As baselines, signatures were classified by two classical metric-based methods: L^2 distance between arclength parameterized curves, Hausdorff distance between the (discrete) arclength parameterized curves considered as unstructured point clouds, and Fréchet distance between the curves. Table 1 reports the overall classification rate (across all 40 signature classes) for each metric, as well as the number of signature classes where classification was perfect. The first group of three results contains parameter values covered by the $R_{a,b}$-transform of [5], with $\frac{a}{2b} = 1$ corresponding to the SRVF and $\frac{a}{2b} = \frac{1}{2}$ corresponding to the complex square-root transform. The last three results are for parameter values obtained using our new transform. Classification is more successful for the parameter values given by our new results, with $\frac{a}{2b} = 2$ matching the performance of the SRVF. Figure 8 shows the classification rates for $\frac{a}{2b} = 1$ and $\frac{a}{2b} = 2$ across the 40 individual signature classes. Although these parameter values have the same overall performance, we see that their performances differ drastically by signature class. This suggests that certain parameter values may be more sensitive to features appearing in particular signature classes, and that it is in general beneficial to vary the choice of parameters to suit a given task.

[2] http://www.dabi.temple.edu/~shape/MPEG7/dataset.html

6.3. Complexity and runtime.

The computational complexity for computing distances between curves in shape space with any of the elastic metrics lies in the reparameterization step. The particular algorithm we are using, outlined in [40], enforces artificial constraints on the slope of each segment in the discrete reparameterization in order to improve speed. To match curves with n samples, the complexity of our algorithm is $O(n^2)$ (without the slope constraints, the complexity is $O(n^4)$—see [40]). The runtime to compute geodesic distance between unparameterized open curves with $n = 100$ samples is on the order of hundredths of a second. For closed curves, the exhaustive seed search scales the runtime by a factor of n, but this can be improved for practical computations; e.g., one can compute L^2 distance between parameterized curves for all seed choices, then keep some smaller collection of best seeds to search over with dynamic programming. For $\frac{a}{2b} \leqslant 1$, computing geodesics between registered curves is essentially instantaneous due to their explicit form for open curves under the $F_{a,b}$-transform and the fast projection algorithm for closed curves. When $\frac{a}{2b} > 1$, the gradient-descent-based algorithm described in section 6.1 is employed to compute geodesics, and its runtime is highly dependent on parameter choices, which we have not optimized.

7. Future directions.

Our first direction for future work is to develop various statistical tools under this framework. These will include computation of summary statistics such as the average and covariance of a set of shapes, exploration of variability in various shape classes through PCA, building generative shape models, inference via hypothesis testing and confidence intervals, and finally regression analysis. Given the simplification of the metric under the proposed $F_{a,b}$-transform, and the simple geometry of the preshape space, we will be able to build upon existing work in this area based on the SRVF transform (i.e., $\frac{a}{2b} = 1$) We have seen in the presented examples that different choices of a and b produce different geodesic paths, thus resulting in different statistical analyses.

Second, we will build statistical models that allow the data to choose appropriate values of a and b for the given application and task. For example, in the context of classification, one can learn optimal weights on training data and then apply the proposed framework on a held-out set. Furthermore, one can build hierarchical Bayesian models for registration, comparison, and averaging of shapes of planar curves that include priors on the values of a and b. Such models can be developed in a similar manner to the functional data approaches in [28, 32]. In those works, the authors simply work with fixed values of a and b. We propose to extend those methods by additionally including the weights for stretching and bending in the posterior distribution.

In previous work, one of the authors extended the results of Younes et al. [55] to give a metric with explicit geodesics on the space of closed loops in $\mathbb{R}^3$ [38]. This is accomplished by replacing the complex squaring map with the Hopf map $S^3 \to S^2$. Using quaternionic arithmetic, we expect that the results of this paper can be extended to provide transforms which simplify metrics on space curves as well.

Funding:

This work was partially supported by National Science Foundation grants DMS-1613054, CCF-1740761, and CCF-1839252 and by National Institute of Health grant R37-CA214955.

REFERENCES

[1]. Anoshkina EV, Belyaev AG, and Seidel H-P, Asymptotic analysis of three-point approximations of vertex normals and curvatures, in Proceedings of Vision, Modeling, and Visualization, Aka GmbH, 2002, pp. 211–216.

[2]. Ayers B, Luders E, Cherbuin N, and Joshi SH, Corpus callosum thickness estimation using elastic shape matching, in International Symposium on Biomedical Imaging, 2015, pp. 1518–1521.

[3]. Bauer M, Bruveris M, Charon N, and MØller-Andersen J, A relaxed approach for curve matching with elastic metrics, ESAIM Control Optim. Calc. Var, 25 (2019), 72.

[4]. Bauer M, Bruveris M, Harms P, and MØller-Andersen J, A numerical framework for Sobolev metrics on the space of curves, SIAM J. Imaging Sci, 10 (2017), pp. 47–73, 10.1137/16M1066282.

[5]. Bauer M, Bruveris M, Marsland S, and Michor PW, Constructing reparameterization invariant metrics on spaces of plane curves, Differential Geom. Appl, 34 (2014), pp. 139–165.

[6]. Bauer M, Bruveris M, and Michor PW, Why use Sobolev metrics on the space of curves, in Riemannian Computing in Computer Vision, Springer, 2016, pp. 233–255.

[7]. Bauer M, Harms P, and Michor PW, Almost local metrics on shape space of hypersurfaces in n-space, SIAM J. Imaging Sci, 5 (2012), pp. 244–310, 10.1137/100807983.

[8]. Bauer M, Harms P, and Michor PW, Curvature weighted metrics on shape space of hypersurfaces in n-space, Differential Geom. Appl, 30 (2012), pp. 33–41.

[9]. Boyer DM, Lipman Y, Clair ES, Puente J, Patel BA, Funkhouser T, Jernvall J, and Daubechies I, Algorithms to automatically quantify the geometric similarity of anatomical surfaces, Proc. Natl. Acad. Sci. USA, 108 (2011), pp. 18221–18226. [PubMed: 22025685]

[10]. Bronstein AM, Bronstein MM, Kimmel R, Mahmoudi M, and Sapiro G, A Gromov-Hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching, Int. J. Comput. Vis, 89 (2010), pp. 266–286.

[11]. Bruveris M, Optimal reparametrizations in the square root velocity framework, SIAM J. Math. Anal, 48 (2016), pp. 4335–4354, 10.1137/15M1014693.

[12]. Bruveris M, Michor PW, and Mumford D, Geodesic completeness for Sobolev metrics on the space of immersed plane curves, in Forum of Mathematics, Sigma, Vol. 2, Cambridge University Press, 2014, e19.

[13]. Celledoni E, Eslitzbichler M, and Schmeding A, Shape analysis on Lie groups with applications in computer animation, J. Geom. Mech, 8 (2016), pp. 273–304.

[14]. Charon N and Trouvé A, The varifold representation of nonoriented shapes for diffeomorphic registration, SIAM J. Imaging Sci, 6 (2013), pp. 2547–2580, 10.1137/130918885.

[15]. Charpiat G, Faugeras O, and Keriven R, Shape statistics for image segmentation with prior, in IEEE Computer Vision and Pattern Recognition, 2007, pp. 1–6.

[16]. Chazal F, Cohen-Steiner D, Guibas LJ, Mémoli F, and Oudot SY, Gromov-Hausdorff stable signatures for shapes using persistence, Computer Graphics Forum, 28 (2009), pp. 1393–1403.

[17]. Dryden IL and Mardia KV, Statistical Shape Analysis: With Applications in R, 2nd ed., Wiley, New York, 2016.

[18]. Duncan A, Klassen E, and Srivastava A, Statistical shape analysis of simplified neuronal trees, Ann. Appl. Stat, 12 (2018), pp. 1385–1421.

[19]. Glaunés J, Qiu A, Miller MI, and Younes L, Large deformation diffeomorphic metric curve mapping, Int. J. Comput. Vis, 80 (2008), pp. 317–336. [PubMed: 20419045]

[20]. Gronwald F, On non-Riemannian parallel transport in Regge calculus, Classical Quantum Gravity, 12 (1995), pp. 1181–1189.

[21]. Haker S, Zhu L, Tannenbaum A, and Angenent S, Optimal mass transport for registration and warping, Int. J. Comput. Vis, 60 (2004), pp. 225–240.

[22]. Hamilton RS, The inverse function theorem of Nash and Moser, Bull. Amer. Math. Soc, 7 (1982), pp. 65–222.

[23]. Jermyn IH, Kurtek S, Klassen E, and Srivastava A, Elastic Shape Matching of Parameterized Surfaces Using Square Root Normal Fields, in European Conference on Computer Vision, Springer, 2012, pp. 804–817.

[24]. Joshi SH, Klassen E, Srivastava A, and Jermyn IH, A novel representation for Riemannian analysis of elastic curves in $\mathbb{R}^n$, in IEEE Computer Vision and Pattern Recognition, 2007, pp. 1–7.

[25]. Kendall DG, Shape manifolds, Procrustean metrics, and complex projective shapes, Bull. London Math. Soc, 16 (1984), pp. 81–121.

[26]. Klassen E, Srivastava A, Mio W, and Joshi SH, Analysis of planar shapes using geodesic paths on shape spaces, IEEE Trans. Pattern Anal. Mach. Intell, 26 (2004), pp. 372–383. [PubMed: 15376883]

[27]. Kriegl A and Michor PW, The Convenient Setting of Global Analysis, Math. Surveys Monogr 53, AMS, 1997.

[28]. Kurtek S, A geometric approach to pairwise Bayesian alignment of functional data using importance sampling, Electron. J. Stat, 11 (2017), pp. 502–531.

[29]. Kurtek S, Klassen E, Gore JC, Ding Z, and Srivastava A, Elastic geodesic paths in shape space of parametrized surfaces, IEEE Trans. Pattern Anal. Mach. Intell, 34 (2012), pp. 1717–1730. [PubMed: 22144521]

[30]. Laga H, Kurtek S, Srivastava A, Golzarian M, and Miklavcic SJ, A Riemannian elastic metric for shape-based plant leaf classification, in 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), 2012, pp. 1–7.

[31]. Lahiri S, Robinson D, and Klassen E, Precise matching of PL curves in $\mathbb{R}^n$ in the square root velocity framework, Geom. Imaging Comput, 2 (2015), pp. 133–186.

[32]. Lu Y, Herbei R, and Kurtek S, Bayesian registration of functions with a Gaussian process prior, J. Comput. Graph. Statist, 26 (2017), pp. 894–904.

[33]. Mémoli F, Gromov–Wasserstein distances and the metric approach to object matching, Found. Comput. Math, 11 (2011), pp. 417–487.

[34]. Michor PW and Mumford D, Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms, Doc. Math, 10 (2005), pp. 217–245.

[35]. Michor PW and Mumford D, An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach, Appl. Comput. Harmon. Anal, 23 (2007), pp. 74–113.

[36]. Miller MI and Younes L, Group actions, homeomorphisms, and matching: A general framework, Int. J. Comput. Vis, 41 (2001), pp. 61–84.

[37]. Mio W, Srivastava A, and Joshi S, On shape of plane elastic curves, Int. J. Comput. Vis, 73 (2007), pp. 307–324.

[38]. Needham T, Kähler structures on spaces of framed curves, Ann. Global Anal. Geom, 54 (2018), pp. 123–153.

[39]. Regge T, General relativity without coordinates, Nuovo Cimento (10), 19 (1961), pp. 558–571.

[40]. Robinson DT, Functional Data Analysis and Partial Shape Matching in the Square Root Velocity Framework, Ph.D. thesis, Florida State University, 2012.

[41]. Royden HL and Fitzpatrick P, Real Analysis, Vol. 2, Macmillan, New York, 1968.

[42]. Schumaker L, Spline Functions: Basic Theory, Cambridge University Press, 2007.

[43]. Small CG, The Statistical Theory of Shape, Springer, 1996.

[44]. Srivastava A, Klassen E, Joshi SH, and Jermyn IH, Shape analysis of elastic curves in Euclidean spaces, IEEE Trans. Pattern Anal. Mach. Intell, 33 (2011), pp. 1415–1428. [PubMed: 20921581]

[45]. Srivastava S, Lal SB, Mishra D, Angadi U, Chaturvedi K, Rai SN, and Rai A, An efficient algorithm for protein structure comparison using elastic shape analysis, Algorithms Molecular Biol, 11 (2016), 27.

[46]. Sundaramoorthi G, Yezzi A, and Mennucci AC, Sobolev active contours, Int. J. Comput. Vis, 73 (2007), pp. 345–366.

[47]. Trouvé A and Younes L, Diffeomorphic matching problems in one dimension: Designing and minimizing matching functionals, in European Conference on Computer Vision, 2000, pp. 573–587.

[48]. Trouvé A and Younes L, On a class of diffeomorphic matching problems in one dimension, SIAM J. Control Optim, 39 (2000), pp. 1112–1135, 10.1137/S036301299934864X.

[49]. Tumpach AB and Preston SC, Quotient elastic metrics on the manifold of arc-length parameterized plane curves, J. Geom. Mech, 9 (2017), pp. 227–256.

[50]. Whitney H, On regular closed curves in the plane, Compositio Math, 4 (1937), pp. 276–284.

[51]. Yeung D, Chang H, Xiong Y, George S, Kashi R, Matsumoto T, and Rigoll G, SVC2004: First International Signature Verification Competition, in Biometric Authentication, 2004, pp. 16–22.

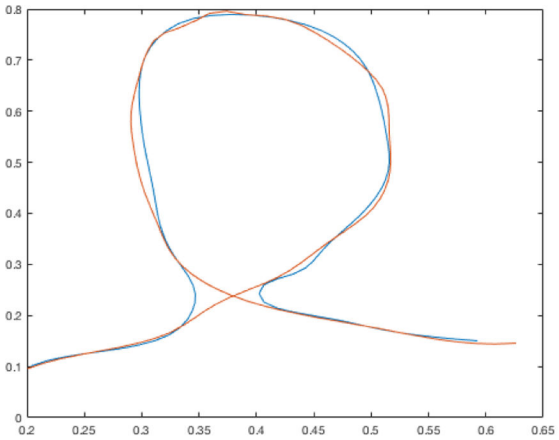
[52]. You Y, Huang W, Gallivan KA, and Absil P-A, A Riemannian approach for computing geodesies in elastic shape analysis, in Signal and Information Processing, 2015, pp. 727–731.

[53]. Younes L, Computable elastic distances between shapes, SIAM J. Appl. Math, 58 (1998), pp. 565–586, 10.1137/S0036139995287685.

[54]. Younes L, Elastic Distance between Curves under the Metamorphosis Viewpoint, preprint, https://arxiv.org/abs/1804.10155, 2018.

[55]. Younes L, Michor PW, Shah JM, and Mumford DB, A metric on shape space with explicit geodesics, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl, 19 (2008), pp. 25–57.

[56]. Zahn CT and Roskies RZ, Fourier descriptors for plane closed curves, IEEE Trans. Comput, 21 (1972), pp. 269–281.



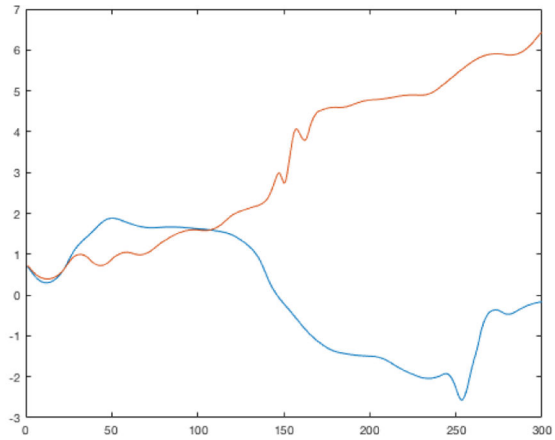


Figure 1.

Left: A pair of curves which are close in Hausdorff distance. Right: The continuous polar angle functions of the curves are quite different, resulting in images under $F_{a,b}$ which are far apart in Hausdorff distance.

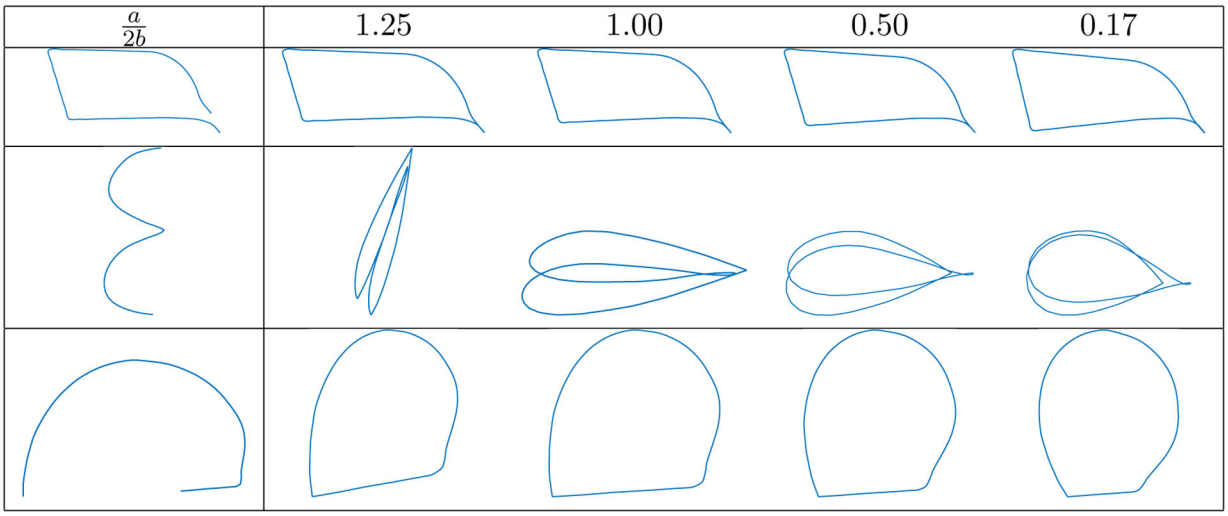


Figure 2.
Projections of an open curve into the preshape space of closed curves under different $F_{a,b}$-transforms.

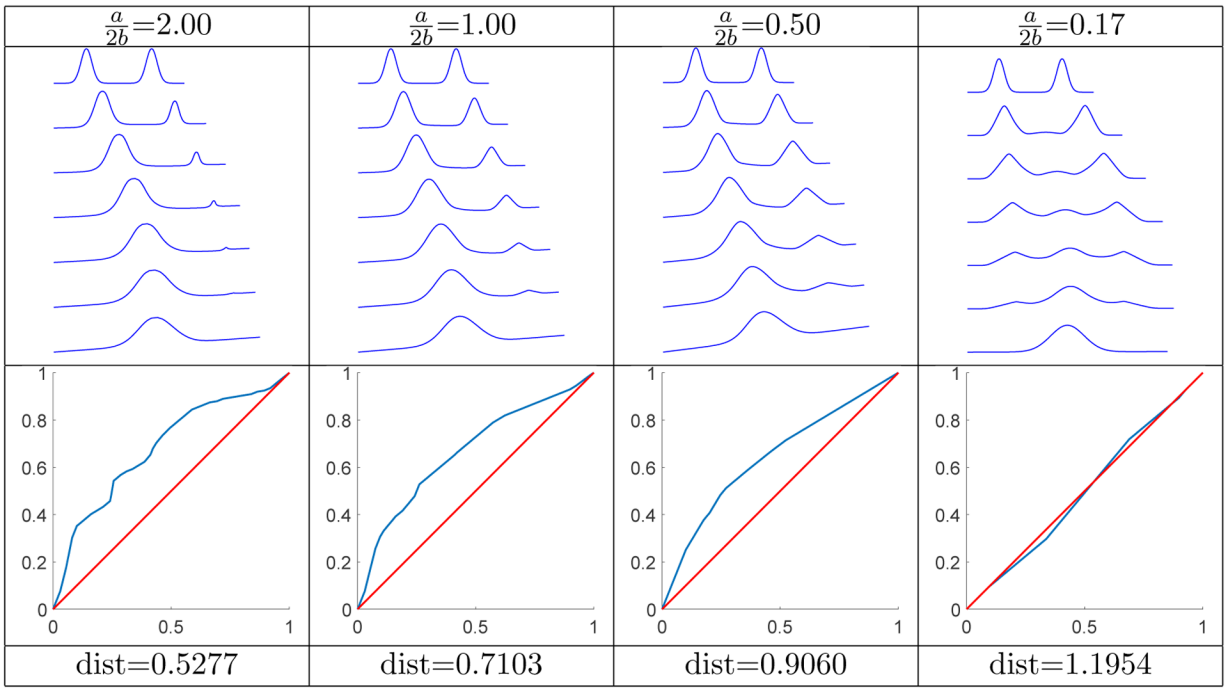


Figure 3.
Geodesic path and distance between shapes of two artificial open curves.

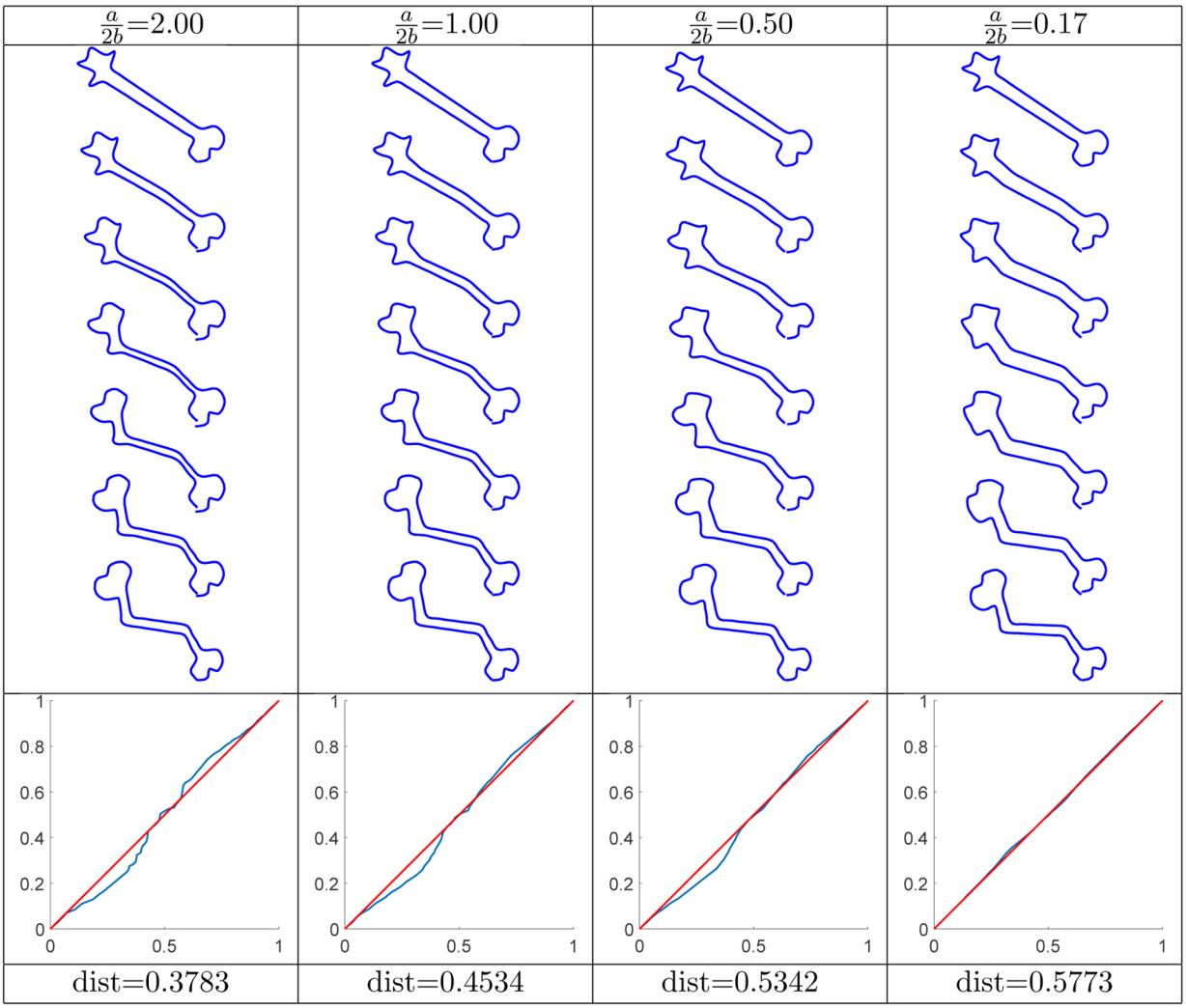


Figure 4.
Geodesic path and distance between two structurally different bone shapes.

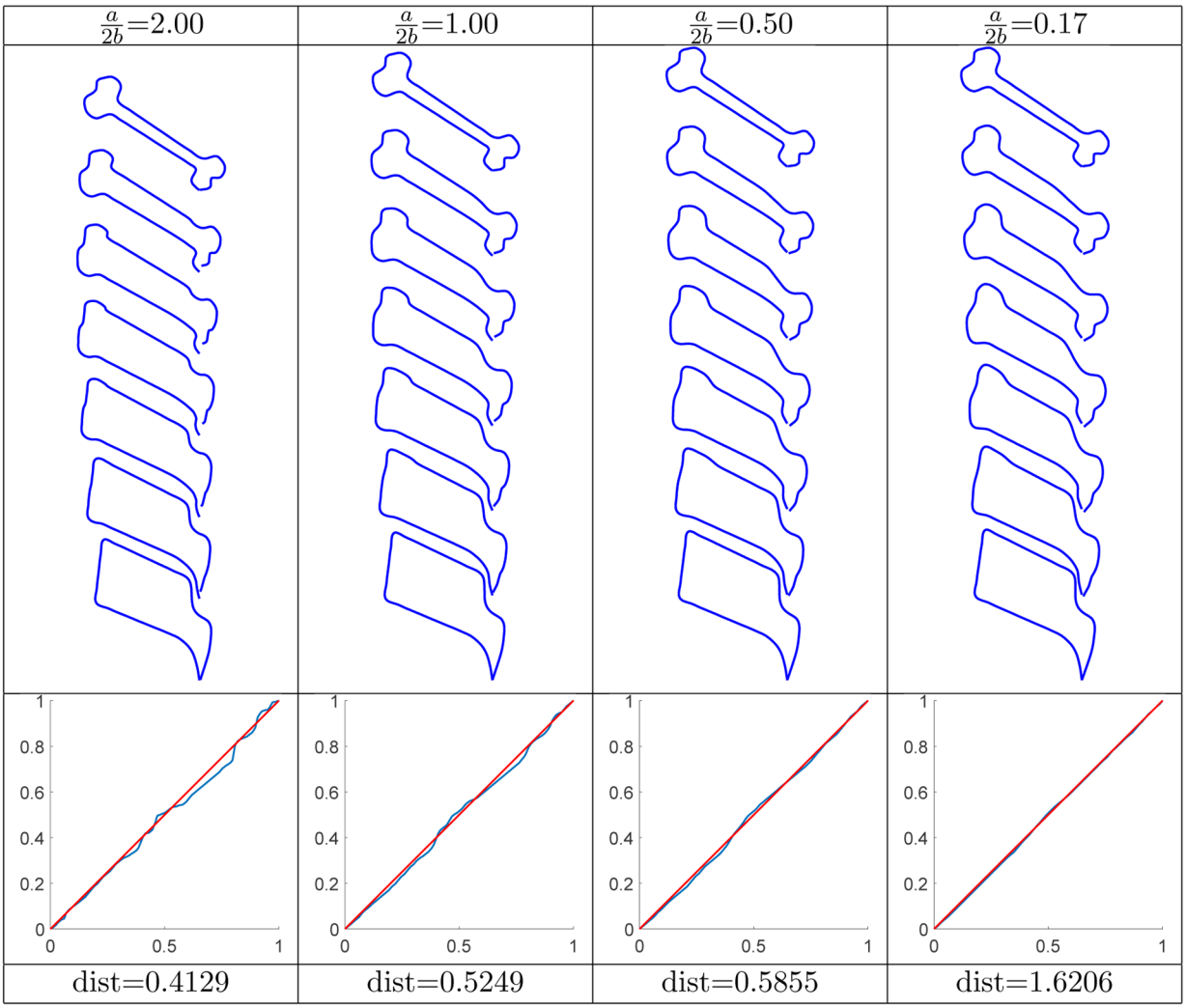


Figure 5.
Geodesic path and distance between two very different shapes.

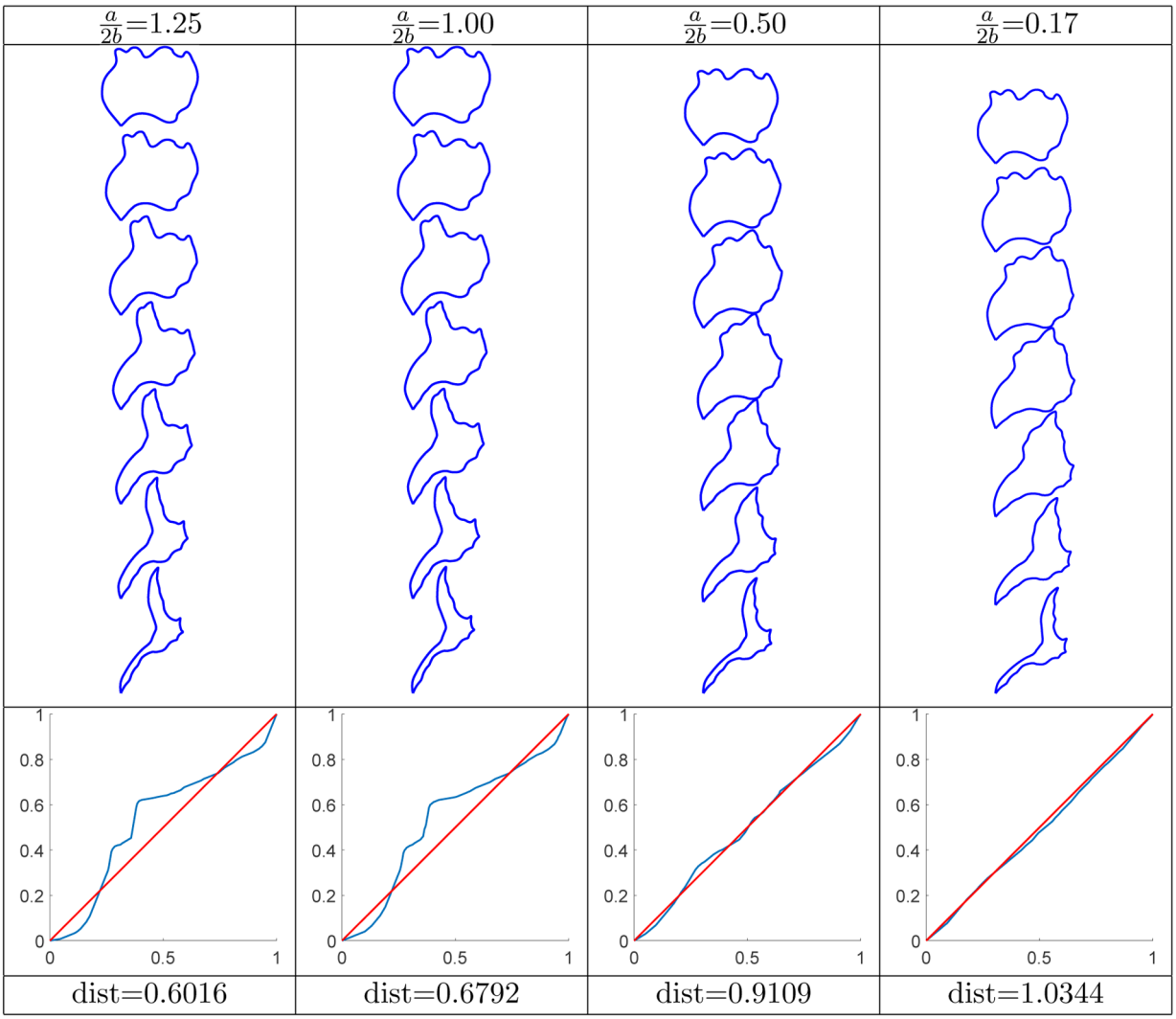


Figure 6.
Geodesic path and distance between two very different shapes.

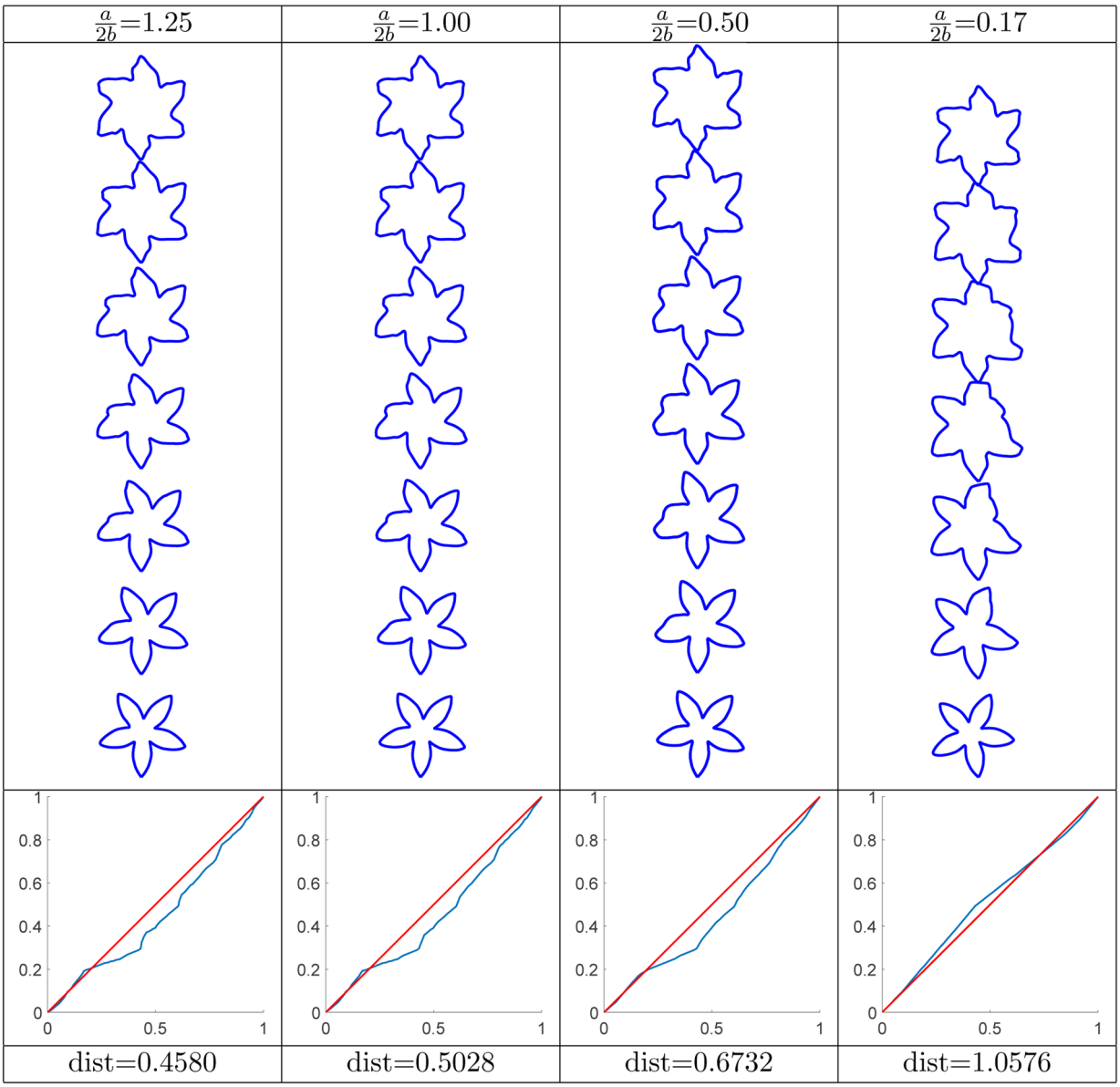


Figure 7.
Geodesic path and distance between two shapes of flowers with a different number of petals.

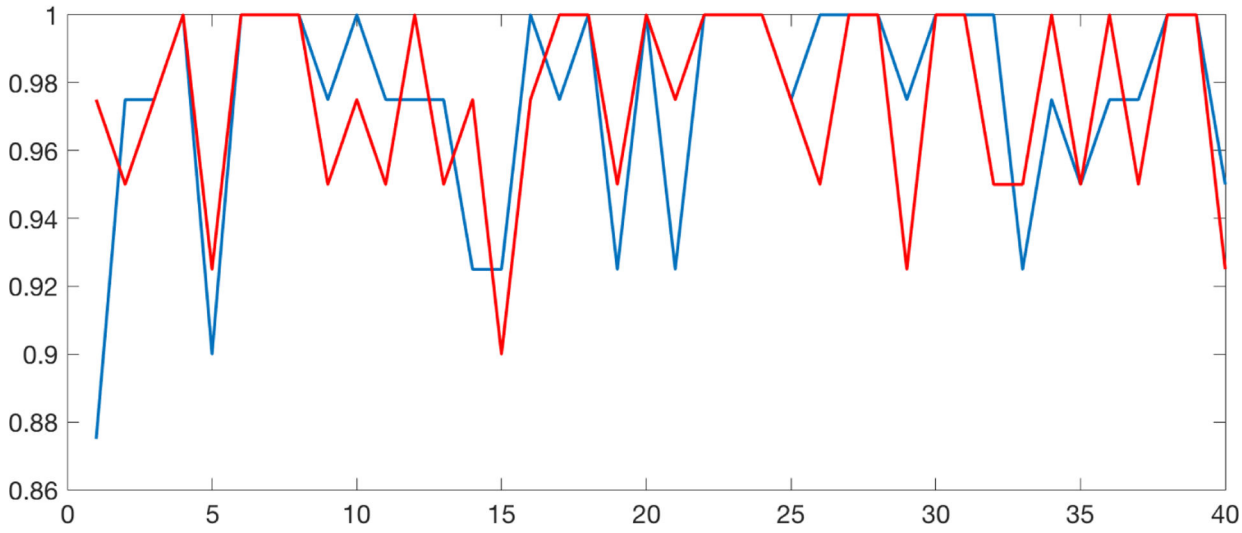


Figure 8.

Classification rates across 40 individual signature classes using the geodesic distance for parameters $\frac{a}{2b} = 1$ (in blue) and $\frac{a}{2b} = 2$ (in red).

Table 1

Overall classification rates and number of perfect classifications for the signature experiment.

Method	Classification rate	Perfect matches
Arclength	86.44	1
Hausdorff	90.81	3
Fréchet	91.75	3
$\frac{a}{2b} = \frac{1}{4}$	86.50	0
$\frac{a}{2b} = \frac{1}{2}$	93.81	6
$\frac{a}{2b} = 1$	97.50	19
$\frac{a}{2b} = 2$	97.50	19
$\frac{a}{2b} = 3$	96.69	16
$\frac{a}{2b} = 4$	97.00	16