arXiv:1502.03229v1 [math.DG] 11 Feb 2015

Why Use Sobolev Metrics on the Space of Curves

Martin Bauer, Martins Bruveris and Peter W. Michor

Abstract In this chapter we study reparametrization invariant Sobolev metrics on spaces of regular curves. We discuss their completeness properties and the resulting usability for applications in shape analysis. In particular, we will argue, that the development of efficient numerical methods for higher order Sobolev type metrics is an extremely desirable goal.

1 Introduction

Over the past decade Riemannian geometry on the infinite-dimensional spaces of parametrized and unparametrized curves has developed into an active research area. The interest has been fueled by the important role of these spaces in the areas of shape analysis and computer vision. Contributions in these fields include applications to medical image diagnostics [18], target and activity recognition in image analysis [38], plant leaf classification [23] and protein structure analysis [26] or human motion analysis in computer graphics [17]. In these research areas one is interested in studying the variability within a certain class of shapes. As a consequence an important goal is the development of statistical tools for these spaces.

Martin Bauer,
University of Vienna, Faculty for Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, AT, e-mail: `bauer.martin@univie.ac.at`

Martins Bruveris
Brunel Unversity London, College of Engineering Design and Physical Sciences, Department of Mathematics, John Crank Building, Brunel University London, Uxbridge, UB8 3PH, UK, e-mail: `martins.bruveris@brunel.ac.uk`

Peter W. Michor,
University of Vienna, Faculty for Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, AT, e-mail: `peter.michor@univie.ac.at`

Riemannian metrics provide the additional structure, that is needed to capture the nonlinearity of the space and at the same time linearize it sufficiently to enable computations. In this chapter we want to acquaint the reader with some of the metrics, that can be defined on the space of curves and discuss their properties with a view towards applications in shape analysis. We will concentrate particularly on completeness properties: do geodesics exist, when do they stop existing and how does it depend on the metric. For a more wide-ranging overview of Riemannian metrics on spaces of functions, see [9].

Parametrized Curves In this chapter we will discuss Riemannian metrics on two different spaces: first, the space of smooth, regular, closed curves in $\mathbb{R}^n$

$$\mathrm{Imm}(S^1, \mathbb{R}^d) = \left\{ c \in C^\infty(S^1, \mathbb{R}^d) : c'(\theta) \neq 0, \forall \theta \in S^1 \right\} ; \tag{1}$$

here Imm stands for *immersion*. This is an open set in the Fréchet space $C^\infty(S^1, \mathbb{R}^d)$ of all smooth functions and as such it is itself a Fréchet manifold. As an open subset of a vector space, its tangent space at any curve is the vector space itself, $T\,\mathrm{Imm}(S^1, \mathbb{R}^d) \cong \mathrm{Imm}(S^1, \mathbb{R}^d) \times C^\infty(S^1, \mathbb{R}^d)$. A Riemannian metric on $\mathrm{Imm}(S^1, \mathbb{R}^d)$ is a smooth map

$$G : \mathrm{Imm}(S^1, \mathbb{R}^d) \times C^\infty(S^1, \mathbb{R}^d) \times C^\infty(S^1, \mathbb{R}^d) \to \mathbb{R} ,$$

such that $G_c(\cdot, \cdot)$ is a symmetric, positive definite bilinear form for all curves c. An example of a Riemannian metric is the L^2-metric $G_c(h, k) = \int_{S^1} \langle h, k \rangle |c'| \, \mathrm{d}\theta$, which we will look at more closely in Sect. 2. When studying particular Riemannian metrics, it will be useful to consider larger spaces of less regular curves, but $\mathrm{Imm}(S^1, \mathbb{R}^d)$ will always be the common core.

Unparametrized Curves The other space, that we will consider, is the space of unparametrized curves, sometimes also denoted *shape space*. There are several closely related, but slightly differing ways to define this space mathematically. We will consider an *unparametrized curve* or *shape* to be an equivalence class of parametrized curves, that differ only by a reparametrization. In other words, c_1 and c_2 represent the same shape, if $c_1 = c_2 \circ \varphi$ for some reparametrization $\varphi \in \mathrm{Diff}(S^1)$; mathematically $\mathrm{Diff}(S^1)$ is the diffeomorphism group of the circle, that is, the set of all smooth invertible maps $\varphi : S^1 \to S^1$. With this definition the space of unparametrized curves is the quotient

$$B(S^1, \mathbb{R}^d) = \mathrm{Imm}(S^1, \mathbb{R}^d) / \mathrm{Diff}(S^1) .$$

Apart from isolated singular points, the space $B(S^1, \mathbb{R}^d)$ is also an infinite-dimensional manifold and the projection $p : c \to [c]$ assigning each curve its equivalence class is a submersion.[1]

---

[1] In applications one often wants to consider curves and shapes modulo Euclidean motions, leading to the spaces $\mathrm{Imm}(S^1, \mathbb{R}^d)/\mathrm{Mot}$ and $B(S^1, \mathbb{R}^d)/\mathrm{Mot}$, where $\mathrm{Mot} = SO(d) \ltimes \mathbb{R}^d$ denotes the Euclidean motion group. All metrics discussed in this chapter are invariant under the motion group and therefore induce a Riemannian metric on the quotients $\mathrm{Imm}(S^1, \mathbb{R}^d)/\mathrm{Mot}$ and $B(S^1, \mathbb{R}^d)/\mathrm{Mot}$.

Reparametrization Invariant Metrics To define a Riemannian metric on the space $B(S^1, \mathbb{R}^d)$ we will start with a Riemannian metric G on $\mathrm{Imm}(S^1, \mathbb{R}^d)$, that is invariant under the action of $\mathrm{Diff}(S^1)$; such metrics are called *reparametrization invariant*. This means, G has to satisfy

$$G_{c \circ \varphi}(h \circ \varphi, k \circ \varphi) = G_c(h, k) \,,$$

for all curves c, tangent vectors h, k and reparametrizations φ. Then we can use the formula

$$G_{[c]}(X, X) = \inf \left\{ G_c(h, h) \,:\, T_c p.h = X \right\} \,,$$

to define a Riemannian metric on shape space $B(S^1, \mathbb{R}^d)$, such that the projection p is a Riemannian submersion.

Geodesic Distance An important concept in shape analysis is the notion of distance between two curves or shapes. A Riemannian metric leads to a natural distance function, the *induced geodesic distance*. The distance measures the length of the shortest path between two curves. If $c_0, c_1 \in \mathrm{Imm}(S^1, \mathbb{R}^d)$ are two parametrized curves, the distance between them is defined as

$$\mathrm{dist}_I(c_0, c_1) = \inf_{\substack{\gamma(0) = c_0 \\ \gamma(1) = c_1}} \int_0^1 \sqrt{G_{\gamma(t)}(\gamma_t(t), \gamma_t(t))} \, \mathrm{d}t \,,$$

where the infimum is taken over all smooth paths γ, that connect the curves c_0 and c_1. Whether there exists a path, realizing this infimum, is an interesting and nontrivial question in Riemannian geometry.

   If we start with a reparametrization invariant metric G and it induces a Riemannian metric on shape space $B(S^1, \mathbb{R}^d)$, then we will be interested in computing the geodesic distance on $B(S^1, \mathbb{R}^d)$. The geodesic distances for parametrized and unparametrized curves are related by

$$\mathrm{dist}_B([c_0], [c_1]) = \inf_{\varphi \in \mathrm{Diff}(S^1)} \mathrm{dist}_I(c_0, c_1 \circ \varphi) \,. \tag{2}$$

For applications in shape analysis it is important to find a stable and fast method to numerically compute this quantity for arbitrary shapes $[c_0]$ and $[c_1]$. We will comment in the later sections, for which metrics a reparametrization φ, realizing the above infimum, exists, and what the obstructions to its existence are in other cases.

Organization We will look at three families of metrics: first, the L^2-metric in Sect. 2, which is among the simplest reparametrization invariant metrics, but unfortunately unsuitable for shape analysis; then, first order Sobolev metrics in Sect. 3, which are very well suited for numerical computations and therefore among the most widely used Riemannian metrics in applications; finally, we will look at higher order Sobolev metrics in Sect. 4 and argue, why their theoretical properties make

---

In Sect. 3 we will encounter metrics, that live naturally on the space $\mathrm{Imm}(S^1, \mathbb{R}^d)/\mathrm{Tra}$ of curves modulo translations.

them good candidates for use in shape analysis. At the end we will explain, how these metrics can be generalized to spaces of parametrized and unparametrized surfaces.

2 The L^2-Metric

The arguably simplest Riemannian metric on the space of smooth, regular curves, that is invariant under reparametrizations, is the L^2-metric

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

where we use $ds = |c'| \, d\theta$ to denote arc length integration. It is integration with respect to ds rather than $d\theta$, that makes this metric reparametrization invariant, as can be seen from the following calculation,

$$G_{c \circ \varphi}(h \circ \varphi, k \circ \varphi) = \int_{S^1} \langle h \circ \varphi, k \circ \varphi \rangle \, (|c'| \circ \varphi) \, \varphi' \, d\theta = G_c(h,k) \,.$$

Similarly, if we wanted to include derivatives of h,k in the metric and keep the metric reparametrization invariant, we would need to use the arc length differentiation $D_s h = \frac{1}{|c'|} h'$ rather that $h' = \partial_\theta h$.

Geodesic Equation The geodesic equation of the L^2-metric is a nonlinear, second order PDE for the path $c(t,\theta)$. It has the form

$$(|c_\theta| c_t)_t = -\frac{1}{2} \left(\frac{|c_t|^2}{|c_\theta|} c_\theta \right)_\theta \,. \tag{3}$$

where $c_\theta = \partial_\theta c = c'$ and $c_t = \partial_t c$ denote the partial derivatives. While the equation is as simple as one can hope for—the geodesic equations for higher order metrics have many more terms—there are currently no existence results available for it.

Open Question. Given a pair of an initial curve and an initial velocity $(c_0, u_0) \in T \operatorname{Imm}(S^1, \mathbb{R}^d)$, does the geodesic equation admit short time solutions with the given initial conditions?

We know, that we cannot hope for long time existence, since it is possible to shrink a circle along a geodesic path down to a point in finite time. Numerical evidence in [29, Sect. 5.3] suggest that geodesics should exist as long as the curvature of the curve remains bounded.

Geodesic Distance The lack of existence results for the geodesic equation is not the biggest problem of the L^2-metric. The crucial property, that makes it unsuitable for applications in shape analysis, is that the induced geodesic distance vanishes.

The *geodesic distance* between two curves $c_0, c_1 \in \operatorname{Imm}(S^1, \mathbb{R}^d)$ is defined as the infimum over the lengths of all paths γ, connecting the two curves, i.e.

$$\text{dist}_I(c_0, c_1) = \inf_{\substack{\gamma(0)=c_0 \\ \gamma(1)=c_1}} \int_0^1 \sqrt{G_{\gamma(t)}(\gamma_t(t), \gamma_t(t))}\, dt \;.$$

It was found [29, 28, 4] that for the L^2-metric the geodesic distance between any two curves is zero.[2] What does this mean? If γ is a smooth, non-constant path, then $\partial_t \gamma(t)$ cannot be identically zero and so the length $\int_0^1 \sqrt{G_\gamma(\gamma_t, \gamma_t)}\, dt$ will be strictly positive. The meaning of $\text{dist}_I(c_0, c_1) = 0$ is that we can find arbitrary short paths connecting c_0 and c_1. No path will have zero length, but given any $\varepsilon > 0$, we can find a path with length $< \varepsilon$. How do these paths look like? They are easier to visualize for the geodesic distance on the space of unparametrized curves, which we will describe next.

For the L^2-metric the geodesic distance between the unparametrized curves $[c_0], [c_1] \in B(S^1, \mathbb{R}^d)$, represented by c_0, c_1 can be computed as the following infimum,

$$\text{dist}_B([c_0], [c_1]) = \inf_\gamma \int_0^1 \sqrt{G_\gamma(\gamma_t^\perp, \gamma_t^\perp)}\, dt \;;$$

here $\gamma(t)$ is a path starting at c_0 and ending at any curve in the equivalence class $[c_1]$, that is $\gamma(1) = c_1 \circ \varphi$ for some $\varphi \in \text{Diff}(S^1)$. We denote by $\gamma_t^\perp = \gamma_t - \langle \gamma_t, v \rangle v$, with $v = D_s \gamma$, the projection of the vector $\gamma_t(t, \theta) \in \mathbb{R}^d$ to the subspace orthogonal to the curve $\gamma(t)$ at θ.
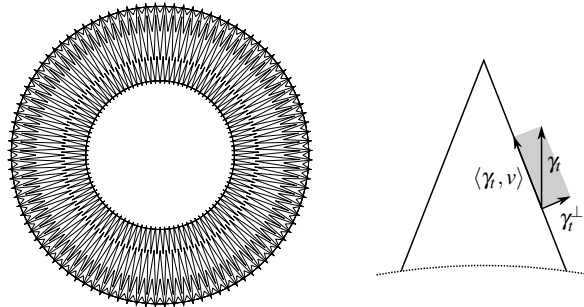
Fig. 1 Left side: a short curve with respect to the L^2-metric in the space $B(S^1, \mathbb{R}^2)$ of unparametrized curves connecting two concentric circles. We see that the intermediate curves are sawtooth-shaped. Right side: Along a sawtooth the tangential component $\langle \gamma_t, v \rangle v$ is large, while the normal component $\gamma^\perp$ becomes small, the steeper the slope of the sawtooth.

A short path connecting two concentric circles can be seen in Fig. 1. The key observation is that the sawtooth-shaped curves have a large tangential velocity, but only a small normal velocity. Since for the geodesic distance on $B(S^1, \mathbb{R}^d)$ we only

---

[2] We encounter the vanishing of the geodesic distance for L^2-metrics on several spaces: on the space $\text{Imm}(M, N)$ of immersions between two manifolds M, N of arbitrary dimension, M compact; on the Virasoro–Bott group [4]; and even for Sobolev metrics on the diffeomorphism group of a compact manifold, provided the order of the metric is $< \frac{1}{2}$ [5, 8].

measure the normal part of the velocity vector, these paths have a short length. The more teeth we use, the smaller the normal component of the velocity and the smaller the length of these paths.

The vanishing of the geodesic distance started the search for stronger metrics, that would be more useful to shape analysis.

Almost Local Metrics One class of metrics, designed to have non-vanishing distance, while being as simple as possible, is the class of almost local metrics. The motivating idea behind almost local metrics was the observation that for paths with short length in the L^2-metric, the intermediate curves are long and have large curvature. Thus one hopes that by adding weights, that depend on length and curvature, to the metric, these paths will be sufficiently penalized and the geodesic distance will become non-zero. *Almost local metrics*[3] are metrics of the form

$$G_c(h,k) = \int_{S^1} \Phi(\ell_c, \kappa) \langle h, k \rangle \, ds \;, \tag{4}$$

with Φ some function of the two variables $\ell_c = \int_{S^1} ds$ (length) and κ (curvature). If Φ depends only on ℓ_c, the resulting metric $G_c(h,k) = \Phi(\ell_c) \int_{S^1} \langle h, k \rangle \, ds$ is a conformal rescaling of the L^2-metric [41, 34]. Other choices for Φ include $\Phi(\kappa) = 1 + A\kappa^2$ with A a positive constant [29] or the scale invariant metric $\Phi(\ell_c, \kappa) = \frac{1}{\ell_c^3} + \kappa^2 \ell_c$ [30].

For all these metrics it has been shown that they induce a point-separating distance function[4] on the space $B(S^1, \mathbb{R}^d)$ of unparametrized curves. However, similarly to the L^2-metric, little is known about solutions of the geodesic equation and while the geodesic distance is point-separating on the space $B(S^1, \mathbb{R}^d)$, it is not point-separating on the space $\mathrm{Imm}(S^1, \mathbb{R}^d)$ of parametrized curves. In the next two sections we will discuss a different strategy to strengthen the L^2-metric, leading to the class of Sobolev metrics.

3 First Order Metrics and the Square Root Velocity Transform

One way to deal with the degeneracy of the L^2-metric is by adding terms, that involve first derivatives of the tangent vectors. Such metrics are called *first order Sobolev metrics* or, short, H^1-metrics. An example is the metric

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

---

[3] These metrics are not local, because the length ℓ_c is not a local quantity, however it is only a mild non-locality; hence the name "almost local" metrics.

[4] A distance function $d(\cdot, \cdot)$ is *point-separating*, if $d(x,y) > 0$ whenever $x \neq y$. This is stronger than non-vanishing, which would only require two points x, y with $d(x, y) \neq 0$.

with $D_s h = \frac{1}{|c'|} h'$ denoting the arc length derivative and $ds = |c'| d\theta$. If we omit the L^2-term, we arrive at $G_c(h,k) = \int_{S^1} \langle D_s h, D_s k \rangle ds$, which is a metric on the space $\text{Imm}(S^1, \mathbb{R}^d)/\text{Tra}$ of regular curves modulo translations. The scale-invariant version of this metric has been studied in [42, 43] and it has the remarkable property that one can find explicit formulas for minimizing geodesics between any two curves.

We will concentrate in this section on a related metric, obtained by using different weights for the tangential and normal components of $D_s h$,

$$G_c(h,k) = \int_{S^1} \langle D_s h^\perp, D_s k^\perp \rangle + \frac{1}{4} \langle D_s h, v \rangle \langle D_s k, v \rangle \, ds \, ; \tag{5}$$

here $v = D_s c = \frac{1}{|c'|} c'$ is the unit length tangent vector along c and $D_s h^\perp = D_s h - \langle D_s h, v \rangle v$ is the projection of $D_s h$ to the subspace $\{v\}^\perp$ orthogonal to the curve. This is a Riemannian metric on $\text{Imm}(S^1, \mathbb{R}^d)/\text{Tra}$ and it is the metric used in the *square root velocity (SRV)* framework [37]. The reason for singling out this metric is that the SRV framework has been used successfully in applications [38, 40, 23] and the SRV transform has a simple and accessible form. We will comment on other H^1-metrics at the end of the section.

The Square Root Velocity Transform The *square root velocity transform (SRVT)* is the map

$$R : \text{Imm}(S^1, \mathbb{R}^d) \to C^\infty(S^1, \mathbb{R}^d), \quad c \mapsto \frac{1}{\sqrt{|c'|}} c' \, ,$$

assigning each curve c a function $q = R(c)$.

Every vector space $(V, \langle \cdot, \cdot \rangle)$ with an inner product can be regarded as a Riemannian manifold: the Riemannian metric g at each point $x \in V$ is simply the inner product, $g_x(\cdot, \cdot) = \langle \cdot, \cdot \rangle$. The SRVT is an isometry between the Riemannian manifold $\big(\text{Imm}(S^1, \mathbb{R}^d)/\text{Tra}, G\big)$, where G is the Riemannian metric (5) and the space $C^\infty(S^1, \mathbb{R}^d)$ with the L^2-inner product $\langle u, v \rangle_{L^2} = \int_{S^1} \langle u, v \rangle d\theta$.

The SRVT for Open Curves Things are simple on the space of open curves $\text{Imm}([0, 2\pi], \mathbb{R}^d)/\text{Tra}$. The SRVT is a one-to-one mapping between the space $\text{Imm}([0, 2\pi], \mathbb{R}^d)/\text{Tra}$ and the set $C^\infty(S^1, \mathbb{R}^d \setminus \{0\})$ of functions that don't pass through the origin in $\mathbb{R}^d$. This is an open subset of all functions and thus geodesics with respect to the metric (5) correspond to straight lines under the SRVT: the path $c(t) = R^{-1}(q_0 + th)$ is a geodesic in $\text{Imm}([0, 2\pi], \mathbb{R}^2)/\text{Tra}$ and given two curves c_0, c_1, the geodesic connecting them is

$$c(t) = R^{-1}((1-t)q_0 + tq_1) \, ,$$

with $q_i = R(c_i)$.

The SRVT for Closed Curves Things are slightly more complicated for closed curves. The inverse of the SRVT is given by the formula

$$R^{-1}(q)(\theta) = \int_0^\theta q|q|\,\mathrm{d}\sigma\,,$$

and we see that if we want the curve $c = R^{-1}(q)$ to be closed, i.e., $c(0) = c(2\pi)$, then we need $\int_{S^1} q|q|\,\mathrm{d}\sigma = 0$. Indeed the image of $\mathrm{Imm}([0,2\pi],\mathbb{R}^d)/\mathrm{Tra}$ under the SRVT is the set

$$\mathrm{Im}(R) = \left\{ q \in C^\infty(S^1,\mathbb{R}^2): \ q(\theta) \neq 0 \text{ and } \int_{S^1} |q|q\,d\theta = 0 \right\}\,.$$

We have the condition $q'(\theta) \neq 0$ as before to ensure that $c'(\theta) \neq 0$ and an additional constraint so that the curves $c = R^{-1}(q)$ are closed. Even though we don't have a closed expression for the geodesics, it is still possible to compute the geodesics numerically without much difficulty.

Minimizing Geodesics for Parametrized Curves Let us first look at open curves. In the SRV representation, $q_i = R(c_i)$, the minimizing path between q_0 and q_1 is the straight line $q(t) = (1-t)q_0 + tq_1$. In particular the minimizing path always exists. It can happen, however, that the straight line between $q_0(\theta)$ and $q_1(\theta)$ passes through the origin; at points (t,θ), where this happens the derivative of the curve $c(t,\theta) = R^{-1}(q(t))(\theta)$ vanishes, i.e., $c'(t,\theta) = 0$ and thus the curve is not regular at those points. Apart from that, any two curves can be joined by a unique minimizing geodesic, which can be computed via an explicit formula, and we know when the intermediate curves will fail to be regular.

For closed curves the situation is less explicit, because now we also have to satisfy the nonlinear constraint $\int_{S^1} q|q|\,\mathrm{d}\theta = 0$. This is a d-dimensional constraint on an otherwise infinite-dimensional space and furthermore the function $q \mapsto \int_{S^1} q|q|\,\mathrm{d}\theta$ is continuous with respect to the L^2-topology. Numerical evidence suggest, that minimizing geodesics continue to exist between any two curves. In particular, computing minimizing geodesics between parametrized curves is a fast and stable operation; an example of a geodesic can be seen in Fig. 2.
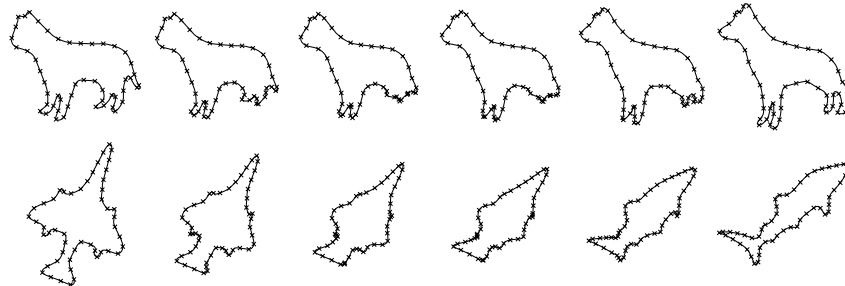
Fig. 2 Minimal geodesics between two pairs of parametrized curves. Images taken from [6].

Smoothness of the minimizing geodesics is another issue. The natural target space for the SRVT is the space $L^2(S^1, \mathbb{R}^d)$ of square-integrable functions. If the SRVT of a curve lies in $L^2(S^1, \mathbb{R}^d)$, the curve itself is only absolutely continuous. Unfortunately the Riemannian metric $G_c(h, k)$, given by (5), does not have to be finite for absolutely continuous curves c and tangent vectors h, k; the term $D_s h = \frac{1}{|c'|} h'$ may well become infinite. We are approaching the frontier of the Riemannian framework now: any two (open) curves can be joined by a minimizing path, however the space, where the path lives—the completion of the space of smooth curves, if one wants to use the term—is not a Riemannian manifold any more.

Minimizing Geodesics for Unparametrized Curves If we want to find minimizing geodesics between two unparametrized curves $C_0, C_1 \in B(S^1, \mathbb{R}^d)/\text{Tra}$, represented by the curves $c_0, c_1 \in \text{Imm}(S^1, \mathbb{R}^d)/\text{Tra}$, one way to do this is to minimize $\text{dist}_I(c_0, c_1 \circ \varphi)$ over $\varphi \in \text{Diff}(S^1)$ or equivalently over all parametrized curves $c_1 \circ \varphi$ representing the shape C_1; indeed, the geodesic distance on $B(S^1, \mathbb{R}^d)/\text{Tra}$ is given by

$$\text{dist}_B(C_0, C_1) = \inf_{\varphi \in \text{Diff}(S^1)} \text{dist}_I(c_0, c_1 \circ \varphi) . \tag{6}$$

If the infimum is attained for $\psi \in \text{Diff}(S^1)$ and if we denote by $c(t)$ the minimizing geodesic between c_0 and $c_1 \circ \psi$, then the curve $[c(t)]$ in $B(S^1, \mathbb{R}^d)/\text{Tra}$ is the minimizing geodesic between C_0 and C_1. Thus we are interested, whether the infimum (6) is attained and if it is, in what space.

Let us look at $\text{dist}_I(c_0, c_1 \circ \varphi)$, first for open curves in the SRV representation. Let $q_i = R(c_i)$. We have $R(c_1 \circ \varphi) = \sqrt{\varphi'} q_1 \circ \varphi$ and

$$\text{dist}_I(c_0, c_1 \circ \varphi)^2 = \int_0^{2\pi} |q_0(\theta) - \sqrt{\varphi'(\theta)} q_1(\varphi(\theta))|^2 \, d\theta .$$

Assume that ψ minimizes this expression, fix $\theta \in S^1$ and set $\tilde{\theta} = \psi(\theta)$. Even though finding ψ has to be done over the whole interval $[0, 2\pi]$ simultaneously, it is very instructive to look at just one point at a time. Consider the infimum

$$\inf_{\sqrt{\varphi'(\theta)} \geq 0} |q_0(\theta) - \sqrt{\varphi'(\theta)} q_1(\tilde{\theta})|^2 .$$

This is a d-dimensional minimization problem, that can be solved explicitly; its solution is visualized in Fig. 3. Denote by α the angle between $q_0(\theta)$ and $q_1(\tilde{\theta})$. If $\frac{\pi}{2} \leq \alpha \leq \frac{3\pi}{2}$, then the infimum is attained for $\sqrt{\varphi'(\theta)} = 0$. In other words, for $q_1(\tilde{\theta})$ lying in the half-plane "opposite" $q_0(\theta)$, the optimal reparametrization would scale it to 0. Next we look at close by points. If the optimal scaling at θ is $\sqrt{\varphi'(\theta)} = 0$ and $\theta + \Delta\theta$ is close enough, then the angle between $q_0(\theta + \Delta\theta)$ and $q_1(\tilde{\theta})$ will also lie inside $[\frac{\pi}{2}, \frac{3\pi}{2}]$ and so $\sqrt{\varphi'(\theta + \Delta\theta)} = 0$ as well. But this would lead to φ being constant on a whole subinterval of $[0, 2\pi]$.

The true situation is more complicated than that, in particular for closed curves, where we additionally have the nonlocal constraint $\int_{S^1} q|q| \, d\theta = 0$ to satisfy. But we
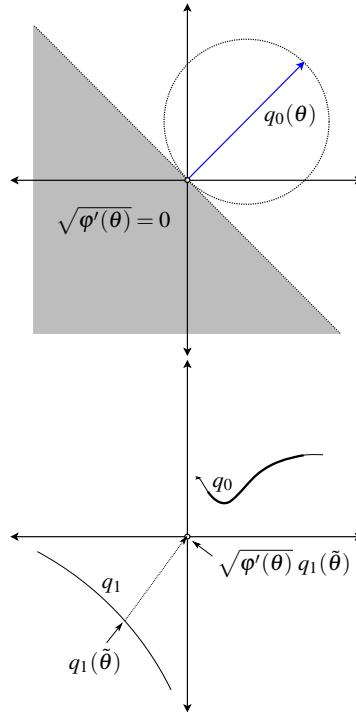
Fig. 3 Left side: solution to the finite dimensional minimization problem. In the halfplane below the dotted line, the solution is given by $\sqrt{\varphi'(\theta)} = 0$. On the halfplane above the dotted line the solution is given by the unique value $\sqrt{\varphi'(\theta)}$ such that $\sqrt{\varphi'(\theta)}q_1(\tilde{\theta})$ lies on the dotted circle. Right side: effect of the reparametrization action on the space of SRVTs.
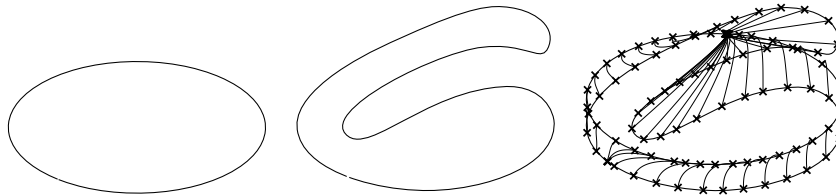
Fig. 4 Left side: initial curve. Middle figure: target curve. Right figure: minimal geodesic on shape $\mathscr{S}(S^1, \mathbb{R}^2)$ between an ellipse and an ellipse with a large fold. One can see, that the fold grows out of a singular point. Image taken from [6].

do observe the scaling-to-zero behavior in numerical computations; see for example Fig. 4.

Incompleteness The key conclusion is this: we should expect the solution ψ of the minimization problem $\inf_\varphi \text{dist}_I(c_1, c_2 \circ \varphi)$ to have intervals, where it is constant— that is true, even if we solve the problem on a finite-dimensional approximation space. If I is such an interval and $\psi|_I = \theta_I \in S^1$, then this means that the whole segment $c_1(I)$ of the first curve corresponds to the point $c_2(\theta_I)$ on the second curve.

Now we can switch c_1 and c_2. Then the optimal reparametrization is ψ^{-1}. However, since ψ is constant on the interval I, its inverse ψ^{-1} will have a jump at the point θ_I. What does this mean for minimizing geodesics? If $c(t)$ is a length-minimizing path between c_2 and $c_1 \circ \psi^{-1}$, then the point $c_2(\theta_I)$ will "open up" to the whole segment $c_1(I)$.

A geodesic is supposed to encode the differences between the shapes represented by c_1 and c_2 in its initial velocity $\partial_t \gamma(0)$. However the geodesic starting at c_2 sees only the parametrized curve $c_1 \circ \psi^{-1}$ and since ψ^{-1} has a jump at θ_I, jumping over the interval I, this interval is missing from the curve $c_1 \circ \psi^{-1}$. How then can the geodesic encode the shape differences, if it does not "see" them?

We understand that this is not a rigorous proof[5]. However there is numerical evidence pointing in the same direction. In Fig. 4 we see an attempt to numerically compute the minimizing geodesic between an ellipse and an ellipse with a large fold, both considered as unparametrized curves. The picture on the right shows the point-to-point correspondences after we have minimized over the reparametrization group. We can see that indeed one point on the ellipse wants to correspond to the part of the fold, where the tangent vectors points in the opposite direction to the tangent vector of the ellipse.

In the example of Fig. 4 we could have cleverly selected the shape with the fold as the initial curve and compute our geodesic starting from there. Then there would be no problem of one point wanting to become a whole segment. However, for two general curves c_1, c_2, we would expect to encounter mixed behavior: some segments of c_1 would collapse to points on c_2 and other points on c_1 would expand to segments of c_2.

This effect is not caused by the global "curvedness" of the manifold of shapes, it is rather a manifestation of the incompleteness. We expect to see this as soon as we match two curves whose tangent vectors point in opposite directions. Since in the SRV representation distances are measured in the L^2-norm, this behavior can occur for a pair of curves with arbitrary small (geodesic) distance. The fold in Fig. 4 can be arbitrary small, but the behavior will be the same.

Let us fix a representative curve c for the shape $[c]$ and let us look at the space of unparametrized curves through the lens of the exponential map, while standing at the curve c. We look at a shape $[c_1]$, represented by a curve c_1, by finding a reparametrization ψ_1, s.t. $\text{dist}_I(c, c_1 \circ \psi)$ is minimal, and then we compute $v_1 = \exp_c^{-1}(c_1 \circ \psi_1)$. This vector v_1 is what we see, when we look at the shape $[c_1]$. Now, if

---

[5] See [43, Sect. 4.2] for a rigorous proof, that this behavior indeed occurs for the metric $G_c(h,k) = \frac{1}{\ell_c} \int_0^{2\pi} \langle D_s h, D_s k \rangle \, ds$ on the space of unparametrized open curves modulo translations.

the reparametrization ψ has jumps, then the curve $c_1 \circ \psi_1$ wil miss parts of the shape $[c_1]$; furthermore, there will be several shapes $[c_2]$, distinct from $[c_1]$ only in the part that is missing from $c_1 \circ \psi_1$, such that the corresponding optimal representing curves $c_2 \circ \psi_2$ coincide with $c_1 \circ \psi_1$. This implies that $v_2 = \exp_c^{-1}(c_2 \circ \psi_2)$ coincides with v_1, while the shapes $[c_2]$ and $[c_1]$ differ. In other words we look at different shapes, but see the same thing. In fact there are many regions in shape space, that cannot be distinguished using the exponential map and these regions start arbitrary close to the starting shape $[c]$.

Joint Reparametrizations It is possible that searching for *one* reparametrization is the wrong problem. Mathematically an equivalent way to define the geodesic distance on $B(S^1, \mathbb{R}^d)/\mathrm{Tra}$ is via

$$\mathrm{dist}_B(C_1, C_2) = \inf_{\varphi_1, \varphi_2 \in \mathrm{Diff}(S^1)} \mathrm{dist}_I(c_1 \circ \varphi_1, c_2 \circ \varphi_2) \ .$$

Using the invariance of dist_I under reparametrizations we can recover (6). Now we are looking for reparametrizations of both curves, such that the infimum is attained. The advantage of this approach is that we can avoid jumps, that would necessarily appear in the one–reparametrization strategy, by instead setting the other reparametrization to be constant on the corresponding interval.

The underlying behavior does not change: points on one curve can be matched to intervals on the other and vice versa. If ψ_1, ψ_2 represent a pair of optimal reparametrizations for two curves c_1, c_2 and ψ_1 is constant on the interval I with $\psi_1|_I = \theta_I$, then the point $c_1(\theta_I)$ will correspond to the interval $c_2(\psi_2(I))$. Instead of jumping over the interval $\psi_2(I)$, we now reparametrize c_1 and the reparametrized curve $c_1 \circ \psi_1$ waits at θ_I until ψ_2 has moved past I.

The strategy of joint reparametrizations was proposed in [33, 24], where the authors consider only curves without the periodicity constraint. Even for open curves, it is not known, whether for any two absolutely continuous curves, there exists a pair of reparametrizations realizing the geodesic distance; in [24] this is shown only under the additional assumption, that one of the curves is piecewise linear.

Other H^1-metrics There are many different H^1-metrics to choose from. For a start, the constants 1 and $\frac{1}{4}$ are rather arbitrary and we could look at the full family of metrics of the form

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

with $a, b > 0$. This family has been given the name *elastic metrics* and has been studied for plane curves in [31, 6]. All metrics in this family are uniformly equivalent, i.e., if G and H are two elastic metrics with possibly different constants a, b, there exists a constant C, such that

$$C^{-1} H_c(h, h) \leq G_c(h, h) \leq C H_c(h, h) \ ,$$

holds for all curves c and all tangent vectors h.

All H^1-metrics can be made invariant with respect to scalings by multiplying them with an appropriate power of the length $\ell(c)$; for example the following metric is scale-invariant,

$$G_c(h,k) = \int_{S^1} \ell_c^{-3} \langle h,k \rangle + \ell_c^{-1} \langle D_s h, D_s k \rangle \, \mathrm{d}s \, .$$

A modified H^1-metric was introduced in [39] with the property that scalings, translations and general deformations of the curve are all orthogonal to each other; this metric was then applied to tracking moving objects.

We have looked at completeness properties only for the H^1-metric corresponding to the SRVT; a similar, more rigorous, discussion can be found for the scale-invariant version of elastic metric corresponding to the choice $a = b$ in [43] and we conjecture, that all H^1-metrics share the same qualitative behavior. We will see in the next section, what happens, if we additionally penalize second and higher derivatives of the tangent vectors.

4 Higher Order Sobolev Metrics

Riemannian metrics involving first derivatives of the tangent vectors can lead to very efficient computations, but some of their mathematical properties are less convenient. Now we will make the metric dependent on higher derivatives. It is not easy to give a definition of a general Sobolev-type Riemannian metric, that is both general enough to encompass all known examples and concrete enough so that we can work easily with it. We will approach this class by looking at families of examples instead, noting common features as well as differences.

A very useful family is that of *Sobolev metrics with constant coefficients*. These are metrics of the form

$$G_c(h,k) = \int_{S^1} a_0 \langle h,k \rangle + a_1 \langle D_s h, D_s k \rangle + \cdots + a_n \langle D_s^n h, D_s^n k \rangle \, \mathrm{d}s \, , \tag{7}$$

with constants $a_0, \ldots, a_n$. The largest n, such that $a_n \neq 0$ is called the *order* of the metric. We require $a_j \geq 0$ for the metric to be positive semi-definite, $a_n > 0$ for it to be a metric of order n and $a_0 > 0$ for it to be non-degenerate. If $a_0 = 0$, then constant tangent vectors are in the kernel of G and, provided there is at least one non-zero coefficient, G defines a non-degenerate metric on the quotient space $\mathrm{Imm}(S^1, \mathbb{R}^d)/\mathrm{Tra}$ of regular curves modulo translations. Most of the metrics encountered in Sect. 3 were of this type.

Using integration by parts we can rewrite (7) to obtain

$$G_c(h,k) = \int_{S^1} a_0 \langle h,k \rangle + a_1 \langle -D_s^2 h, k \rangle + \cdots + a_n \langle (-1)^n D_s^{2n} h, k \rangle \, \mathrm{d}s \, , \tag{8}$$

enabling us to write the metric in the form $G_c(h,k) = \int_{S^1} \langle L_c h, k \rangle \, ds$, with $L_c = \sum_{j=0}^{n} (-1)^j a_j D_s^{2j}$, a differential operator of order $2n$.

Metrics with Nonconstant Coefficients We could loosen our restrictions on the coefficients a_j and permit them to be functions, that depend on the curve c and quantities derived from it, e.g., $a_j = a_j(\ell_c, \kappa, D_s\kappa, \dots)$. In Sect. 2 we have considered such metrics of order zero, the almost local metrics. Several examples of higher order metrics with nonconstant coefficients have been investigated in the literature. The completeness properties of first and second order metrics with coefficients depending on the length are studied in [27]. The idea in [35] and [11] is to decompose a tangent vector $h = h^{\parallel} + h^{\perp}$ into a part tangent and a part normal to the curve and consider derivatives of these quantities. Some special examples of second order metrics can be found in [10].

Geodesic Equation The geodesic equation of a Sobolev metric with constant coefficients is a nonlinear PDE, second order in time and of order $2n$ in the space variable. It is given by

$$\partial_t \left(\sum_{j=0}^{n} (-1)^j a_j |c'| D_s^{2j} c_t \right) = -\frac{a_0}{2} |c'| D_s \left(\langle c_t, c_t \rangle D_s c \right)$$

$$+ \sum_{k=1}^{n} \sum_{j=1}^{2k-1} (-1)^{k+j} \frac{a_k}{2} |c'| D_s \left(\langle D_s^{2k-j} c_t, D_s^j c_t \rangle D_s c \right) .$$

We can see that if $a_j = 0$ for $j \geq 1$, then this equation reduces to the geodesic equation (3) of the L^2-metric. The left hand side of the geodesic equation is the time derivative of the momentum, $L_c c_t |c'|$. For metrics of order $n \geq 1$ the geodesic equation is locally well-posed [30].

Now we come to the main difference between Sobolev metrics of order one and metrics of higher order. In a nutshell, first order Sobolev metrics are only *weak* Riemannian metrics, while Sobolev metrics of higher order, if extended to a suitable, larger space, are *strong* Riemannian metrics.

Weak Sobolev Metrics Let G be a Sobolev metric of order one,

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

Fix the curve c and look at the inner product $G_c(\cdot, \cdot)$. The natural space to define $G_c(\cdot, \cdot)$ is the Sobolev space $H^1(S^1, \mathbb{R}^d)$ of functions with square-integrable derivatives, together with the inner product

$$\langle h, k \rangle_{H^1} = \int_{S^1} \langle h, k \rangle + \langle h', k' \rangle \, d\theta .$$

If c is smooth enough, say $c \in C^1$, then we see that $G_c(\cdot, \cdot)$ defines an inner product on H^1, which is equivalent to the standard inner product. Unfortunately we cannot allow c itself to be an H^1-function. We need uniform control on the derivative c'

to guarantee that the integral $\int_{S^1} \langle h', k' \rangle |c'|^{-1} \, d\theta$ is finite, but the H^1-norm does not provide that. The best we can do is to extend G to the space

$$G : C^1 \text{Imm}(S^1, \mathbb{R}^d) \times H^1(S^1, \mathbb{R}^d) \times H^1(S^1, \mathbb{R}^d) \to \mathbb{R} \, .$$

In this sense G is a *weak Riemannian metric*[6]; the topology induced by the inner product $G_c(\cdot, \cdot)$, in this case the H^1-topology, is weaker than the manifold topology, here the C^∞- or C^1-topology.

Strong Sobolev Metrics The situation is different for Sobolev metrics with constant coefficients of order 2 or higher. Let us look at an example:

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

with $D_s h = \frac{1}{|c'|} h'$ and $D_s^2 h = \frac{1}{|c'|} h'' - \frac{\langle c', c'' \rangle}{|c'|} h'$. Again the natural space for $G_c(\cdot, \cdot)$ is the Sobolev space $H^2(S^1, \mathbb{R}^d)$ and it would appear that we need $c \in C^2$ for the inner product to be well-defined. However a careful application of Sobolev embedding and multiplier theorems—see [16, Sect. 3.2]—shows that we can extend G to a smooth inner product on the space

$$G : \mathscr{I}^2(S^1, \mathbb{R}^d) \times H^2(S^1, \mathbb{R}^d) \times H^2(S^1, \mathbb{R}^d) \to \mathbb{R} \, ;$$

here we denote by $\mathscr{I}^2(S^1, \mathbb{R}^d) = \{c \in H^2 : c'(\theta) \neq 0 \; \forall \theta \in S^1\}$ the space of H^2-curves with non-vanishing tangent vectors. The crucial fact is the Sobolev embedding $H^2 \hookrightarrow C^1$, implying that the H^2-norm controls first derivatives uniformly. This also implies that $\mathscr{I}^2$ is an open set in H^2. Thus G becomes a *strong Riemannian metric* on $\mathscr{I}^2(S^1, \mathbb{R}^d)$; the topology induced by each inner product $G_c(\cdot, \cdot)$ coincides with the manifold topology.

Similarly Sobolev metrics of order n with constant coefficients induce strong metrics on the space $\mathscr{I}^n(S^1, \mathbb{R}^d)$ of regular Sobolev curves, provided $n \geq 2$.

Note however that Sobolev metrics of order 2 and higher are strong metrics only when considered on the larger space $\mathscr{I}^n(S^1, \mathbb{R}^d)$, not on the space $\text{Imm}(S^1, \mathbb{R}^d)$ of smooth curves. On $\text{Imm}(S^1, \mathbb{R}^d)$ the metric is still a weak metric. That said the difference between metrics of order 1 and those of higher order is that for higher order metrics we are able to pass to the larger space $\mathscr{I}^n(S^1, \mathbb{R}^d)$—one could say we are "completing" the space of smooth curves—on which it becomes a strong metric, while for first order metrics such a completion does not exist.

Properties of Strong Metrics The following is a list of properties we get "for free" simply by working with a smooth, strong Riemannian metric as opposed to a weak one:

---

[6] An infinite-dimensional Riemannian manifold (M, g) is called *strong*, if g induces the natural topology on each tangent space or equivalently, if the map $g : TM \to (TM)'$ is an isomorphism. If g is merely a smoothly varying nondegenerate bilinear form on TM we call (M, g) a *weak* Riemannian manifold, indicating that the topology induced by g can be weaker than the natural topology on TM or equivalently $g : TM \to (TM)'$ is only injective.

- The Levi-Civita covariant derivative exists and the geodesic equation has local solutions, which depend smoothly on the initial conditions.
- The exponential map exists and is a local diffeomorphism.
- The induced geodesic distance is point-separating and generates the same topology as the underlying manifold.

The theory of strong, infinite-dimensional Riemannian manifolds is described in [25] and [20]. For weak Riemannian manifolds all these properties have to be established by hand and there are examples, where they fail to hold. The geodesic distance for the L^2-metric discussed in Sect. 2 on the space of curves vanishes identically [29, 28], it is not known whether the geodesic equation for the L^2-metric has solutions and [7] presents a weak Riemannian manifold that has no Levi-Civita covariant derivative.

We would like to note that the distinction between weak and strong Riemannian metrics arises only in infinite-dimensions. Every Riemannian metric on a finite-dimensional manifold is strong. Therefore phenomena, like the vanishing of the geodesic distance or the failure of geodesics to exist, can only arise in infinite dimensions. For better or for worse, this is the setting, where the joy and pain of shape analysis occurs.

Completeness Properties What are the operations on the manifold of curves, that are used in shape analysis?

(1) We want to use the Riemannian exponential map $\exp_c : T_cM \to M$ to pass between the tangent space at one point and the manifold itself. Its inverse $\log_c = \exp_c^{-1}$ allows us to represent the nonlinear manifold or at least a part thereof in a vector space.

(2) Given two curves, we want to compute the geodesic distance between them, that is, the length of the shortest path joining them. Often we are also interested in the shortest path itself; it can be used to transfer information between two curves and the midpoint can serve as the average of its endpoints.

(3) Given a finite set $\{c_1, \ldots, c_n\}$ of curves, we are interested in the average of this set. On a Riemannian manifold this is usually the Fréchet or Karcher mean, i.e., a curve c^*, minimizing the sum of squared distances, $c^* = \operatorname{argmin}_c \sum_i d(c, c_i)^2$, where d is the geodesic distance.

This is by no means an exhaustive list, but describes rather the basic operations, one wants to perform. The ability to do so places conditions on the manifold of curves and the Riemannian metric. Let us look at these conditions.

(1) On a general Riemannian manifold the exponential map $\exp_c : U \to M$ is defined only on an open neighborhood $U \subseteq T_cM$ of 0 and it is rarely known exactly how U looks like. For us to be able to freely map tangent vectors to curves, we want a globally defined exponential map. Since the exponential map is defined as $\exp_c(h) = \gamma(1)$, where γ is the geodesic with initial conditions (c, h) and using the property $\exp_c(th) = \gamma(t)$, we can see that a globally defined exponential map is equivalent to requiring that geodesics exist for all time. This property is called *geodesic completeness*.

(1') Asking for the exponential map to be invertible is more difficult. On a strong Riemannian manifold this is always the case locally. Imposing it globally is a very restrictive condition. The Weil–Peterson metric [36] comes closest; it is a metric with negative sectional curvature, meaning that the derivative of the exponential map is everywhere invertible.

(2) Here we want to know, whether any two curves can be joined by a minimizing geodesic, i.e., a geodesic, whose length realizes the geodesic distance. On a finite-dimensional manifold geodesic completeness would imply this property; this is not the case in infinite dimensions [1]. This is not to say, that we cannot hope for minimizing geodesics to exist, but rather, that it will have to be proven independently of (1).

(3) Ensuring that the Fréchet mean exists for all finite collections of curves is difficult. But there is a theorem [3] stating that the mean exists and is unique on a dense subset of the n-fold product $M \times \ldots \times M$, provided the manifold is *metrically complete*. This means that the manifold (M, d) together with the induced geodesic distance is complete as a metric space.

The properties (1), (2) and (3) for Riemannian manifolds are called completeness properties. In finite dimensions the theorem of Hopf–Rinow states that (1) and (3) are equivalent and either one implies (2). For infinite-dimensional strong Riemannian manifolds the only remaining implication is that metric completeness implies geodesic completeness.[7].

Completeness for Sobolev Metrics   Let us look at the situation for Sobolev metrics on the space of parametrized curves. We have argued in Sect. 3 that we shouldn't expect H^1-metrics to be geodesically or metrically complete. Things look better for Sobolev metrics of higher order. In fact it is shown in [16] and [15] that these metrics satisfy all the above mentioned completeness properties. [8]

Theorem 1. *Let $n \geq 2$ and let G be a Sobolev metric of order n with constant coefficients. Then*

(1) $(\mathscr{I}^n(S^1, \mathbb{R}^d), G)$ is geodesically complete;
(2) Any two elements in the same connected component of $\mathscr{I}^n(S^1, \mathbb{R}^d)$ can be joined by a minimizing geodesic;
(3) $(\mathscr{I}^n(S^1, \mathbb{R}^d), \mathrm{dist}_{\mathscr{I}})$ is a complete metric space.

The geodesic equation for Sobolev metrics has a smoothness preserving property. If the initial conditions are smoother that H^n—let us say the initial curve and initial velocity are C^∞—then the whole geodesic will be as smooth as the initial conditions. Therefore the space $(\mathrm{Imm}(S^1, \mathbb{R}^d), G)$ of smooth immersions with a Sobolev metric of order n is also geodesically complete.

---

[7] A counterexample, showing that in infinite dimensions metric and geodesic completeness together do not imply existence of minimizing geodesics can be found in [1]; similarly, that geodesic completeness does not imply metric completeness is shown in [2]

[8] A related result holds for the space of curves of bounded second variation, together with a Finsler BV^2-metric. It is shown in [32] that any two curves in the same connected component of the space $BV^2(S^1, \mathbb{R}^2)$ can be joined by a length-minimizing path.

Completeness for Unparametrized Curves Similar completeness properties hold for unparametrized curves. The correct space, where to look for completeness, is the quotient

$$\mathscr{B}^n(S^1,\mathbb{R}^d) = \mathscr{I}^n(S^1,\mathbb{R}^d)/\mathscr{D}^n(S^1)\;,$$

of regular curves of Sobolev class H^n modulo reparametrizations of the same regularity; the group $\mathscr{D}^n(S^1) = \{\varphi \in H^n(S^1,S^1) : \varphi'(\theta) > 0\}$ is the group of H^n-diffeomorphisms. We have the following theorem [15]:

Theorem 2. *Let $n \geq 2$ and let G be a Sobolev metric of order n with constant coefficients. Then*

(1) $(\mathscr{B}^n(S^1,\mathbb{R}^d),\mathrm{dist}_{\mathscr{B}})$ with the quotient metric induced by the geodesic distance on $(\mathscr{I}^n,G)$ is a complete metric space;
(2) Given $C_1,C_2 \in \mathscr{B}^n$ in the same connected component, there exist $c_1,c_2 \in \mathscr{I}^n$ with $c_1 \in \pi^{-1}(C_1)$ and $c_2 \in \pi^{-1}(C_2)$, such that

$$\mathrm{dist}_{\mathscr{B}}(C_1,C_2) = \mathrm{dist}_{\mathscr{I}}(c_1,c_2)\;;$$

equivalently, the infimum in

$$\mathrm{dist}_{\mathscr{B}}(\pi(c_1),\pi(c_2)) = \inf_{\varphi \in \mathscr{D}^n(S^1)} \mathrm{dist}_{\mathscr{I}}(c_1,c_2 \circ \varphi)$$

is attained;
(3) Any two shapes in the same connected component of $\mathscr{B}^n(S^1,\mathbb{R}^d)$ can be joined by a minimizing geodesic.

The only drawback is that the space $\mathscr{B}^n(S^1,\mathbb{R}^d)$ of Sobolev shapes is not a manifold any more[9]. It is however a topological space and equipped with the geodesic distance function a metric space. We have to understand a minimizing geodesic in the sense of metric spaces, i.e., a curve $\gamma : I \to \mathscr{B}^n$ is a minimizing geodesic, if

$$\mathrm{dist}_{\mathscr{B}}(\gamma(t),\gamma(s)) = \lambda|t-s|$$

holds for some $\lambda > 0$ and all $t,s \in I$.

We would like to point out that part (2) of Thm. 2 is the counterpart of the incompleteness discussion in Sect. 3. This theorem states that given two shapes, represented by two parametrized curves, we can find an optimal reparametrization φ of the second curve. The fact that $\varphi \in \mathscr{D}^n(S^1)$ guarantees that φ is at least a C^1-diffeomorphism of the circle; thus no intervals are collapsed to single points and neither φ nor φ^{-1} has any jumps.

---

[9] This has to do with smoothness properties of the composition map in Sobolev spaces. While the smooth reparametrization group $\mathrm{Diff}(S^1)$ acts smoothly on the space $\mathrm{Imm}(S^1,\mathbb{R}^d)$ of smooth curves [21], the group of Sobolev reparametrizations $\mathscr{D}^n(S^1)$ acts only continuously on Sobolev curves $\mathscr{I}^n(S^1,\mathbb{R}^d)$. Moreover, the smooth shape space $B(S^1,\mathbb{R}^d)$ is (apart from isolated singularities) a manifold, but the Sobolev shape space $\mathscr{B}^n(S^1,\mathbb{R}^d)$ is only a topological space. This is the price we have to pay for completeness. See [15, Sect. 6] for details.

Numerical Computations We have argued that Sobolev metrics of sufficiently high order have nice mathematical properties, which are relevant to applications in shape analysis. What we have not done is present convincing applications, showcasing their superior performance. This is because the numerical computation of minimizing geodesics between parametrized or unparametrized curves is still an open problem. First attempts at discretizing special cases of H^2-metrics can be found in [11, 32]. While first order metrics have nice representations in terms of the SRVT or the basic mapping [43], which greatly simplifies the numerics, there is no such analogon for higher order metrics[10] Finding a robust and stable discretization of metrics of order 2 and higher remains a challenge.

5 Riemannian Metrics on the Space of Surfaces

In the previous sections we have presented several reparametrization invariant metrics on the space of curves. We want to conclude the exposition with a short excursion to the space of regularly parametrized surfaces, i.e., $\mathrm{Imm}(M,\mathbb{R}^d) = \{f \in C^\infty(M,\mathbb{R}^d) : T_x f \text{ injective } \forall x \in M\}$ with M a compact 2-dimensional manifold without boundary. Typical choices for M are the sphere S^2 and the torus $S^1 \times S^1$. In particular we will describe how the previously described metrics can be generalized from $\mathrm{Imm}(S^1,\mathbb{R}^2)$ to $\mathrm{Imm}(M,\mathbb{R}^d)$. We will follow the presentation of [12, 13].

Sobolev metrics To generalize Sobolev metrics (8) from the space of curves to the space of surfaces, we need the right replacements for the arc length derivative D_s and integration ds. For an immersion $f \in \mathrm{Imm}(M,\mathbb{R}^d)$, we denote by $g = g^f = f^*\langle\cdot,\cdot\rangle$ the pullback of the Euclidean metric to M. This makes (M,g) into a Riemannian manifold with a Laplace operator $\Delta^g = -\mathrm{div}^g \circ \mathrm{grad}^g$ and volume form vol^g. We will use Δ^g and vol^g as replacements for $-D_s^2$ and ds. A Sobolev metric of order n on $\mathrm{Imm}(M,\mathbb{R}^d)$ is given by

$$G_f(h,k) = \int_M a_0\langle h,k\rangle + a_1\langle\Delta^g h,k\rangle + \cdots + a_n\langle(\Delta^g)^n h,k\rangle\, \mathrm{vol}^g \; ; \tag{9}$$

here the tangent vectors h,k are seen as maps $h,k : M \to \mathbb{R}^d$ and the Laplace operator acts on each coordinate separately[11].

---

[10] In [10] a representation of second order metrics, similar to the SRVT was developed. However, image of the resulting transformations have infinite co-dimension, which, compared to the SRVT, complicates the situation.

[11] On $\mathrm{Imm}(M,\mathbb{R}^d)$ there is no canonical way to define Sobolev metrics. We could have also used the definition

$$G_f(h,k) = \int_M a_0\langle h,k\rangle + \cdots + a_n\langle(\nabla^g)^n h, (\nabla^g)^n k\rangle\, \mathrm{vol}^g \;,$$

where ∇^g is the covariant derivative on M. Then the differential operator associated to this metric is $L_f = \sum_{j=0}^n a_j\left((\nabla^g)^*\right)^n (\nabla^g)^n \otimes \mathrm{vol}^g$, with $(\nabla^g)^* = \mathrm{Tr}(g^{-1}\nabla^g)$ denoting the adjoint of ∇^g. When $n \geq 2$, the operator $\left((\nabla^g)^*\right)^n (\nabla^g)^n$ differs from $(\Delta^g)^n$ in the lower order terms—they are connected via the Weitzenböck formulas.

The associated operator L_f, allowing us to write $G_f(h,k) = \int_{S^1} \langle L_f h, k \rangle \, ds$, is given by $L_f = \sum_{j=0}^{n} a_j (\Delta^g)^j \otimes \text{vol}^g$. Every metric in this family is invariant under the action of the diffeomorphism group $\text{Diff}(M)$ and induces a Riemannian metric on the quotient space of unparametrized surfaces $B(M,\mathbb{R}^3) = \text{Imm}(M,\mathbb{R}^3)/\text{Diff}(M)$.

Similarly as in the previous section one can also allow the coefficients a_j to depend on the surface f. Coefficients depending on the total volume, the mean curvature and the Gauß curvature have been considered in [14]. The class of almost local metrics on surfaces has also been studied [13].

Geodesic Distance For the geodesic distance we obtain similar results as for curves: the geodesic distance vanishes for the L^2-metric on $\text{Imm}(M,\mathbb{R}^d)$. Both higher order Sobolev metrics and almost local metrics depending on mean curvature or total volume induce point-separating distances on the space of unparametrized surfaces. Whether higher order Sobolev metrics induce a point-separating distance function on $\text{Imm}(M,\mathbb{R}^d)$ itself is not known.

Geodesic Equation The formulas for the geodesic equations for Sobolev metrics (9) become very quickly very technical; see [12]. As an example we present the geodesic equation of the H^1-metric with coefficients $a_0 = a_1 = 1$; this is the metric induced by the operator field $L_f = 1 + \Delta^g$.

$$\partial_t \left(L_f f_t \otimes \text{vol}^g \right) = \left(\text{Tr} \left(g^{-1} S^f g^{-1} \langle \nabla^g f_t, \nabla^g f_t \rangle \right) - \frac{1}{2} \text{Tr} \left(g^{-1} \nabla^g \langle \nabla^g f_t, f_t \rangle \right).H^f \right.$$
$$\left. - \frac{1}{2} \langle L_f f_t, f_t \rangle.H^f - Tf.\langle L_f f_t, \nabla^g f_t \rangle^{\sharp} \right) \otimes \text{vol}^g \; ;$$

here S^f denotes the second fundamental form, $H^f = \text{Tr}(g^{-1}S^f)$ the vector valued mean curvature and ∇^g the covariant derivative of the surface f.

Outlook On spaces of surfaces the theory of Sobolev metrics is significantly less developed than on the space of curves. We conjecture that Sobolev metric of order $n \geq 3$ will again be strong Riemannian metrics on the space $\mathscr{I}^n(M,\mathbb{R}^d)$ of Sobolev surfaces. Nothing is known about completeness properties of these metrics.

An analogue of the SRVT transform has been developed for surfaces in [19, 22]. However questions regarding the invertibility of the transform and the characterization of its image remain open. So far no numerical experiments for higher order Sobolev metrics on the space of surfaces have been conducted.

References

1. C. J. Atkin. The Hopf-Rinow theorem is false in infinite dimensions. *Bull. London Math. Soc.*, 7(3):261–266, 1975.
2. C. J. Atkin. Geodesic and metric completeness in infinite dimensions. *Hokkaido Math. J.*, 26(1):1–61, 1997.
3. D. Azagra and J. Ferrera. Proximal calculus on Riemannian manifolds. *Mediterr. J. Math.*, 2(4):437–450, 2005.
4. M. Bauer, M. Bruveris, P. Harms, and P. W. Michor. Vanishing geodesic distance for the Riemannian metric with geodesic equation the KdV-equation. *Ann. Global Anal. Geom.*, 41(4):461–472, 2012.
5. M. Bauer, M. Bruveris, P. Harms, and P. W. Michor. Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group. *Ann. Global Anal. Geom.*, 44(1):5–21, 2013.
6. M. Bauer, M. Bruveris, S. Marsland, and P. W. Michor. Constructing reparameterization invariant metrics on spaces of plane curves. *Differential Geom. Appl.*, 34:139–165, 2014.
7. M. Bauer, M. Bruveris, and P. Michor. Homogeneous Sobolev metric of order one on diffeomorphism groups on real line. *Journal of Nonlinear Science*, pages 1–40, 2014.
8. M. Bauer, M. Bruveris, and P. W. Michor. Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group. II. *Ann. Global Anal. Geom.*, 44(4):361–368, 2013.
9. M. Bauer, M. Bruveris, and P. W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vis.*, 50:60–97, 2014.
10. M. Bauer, M. Bruveris, and P. W. Michor. R-transforms for Sobolev H^2-metrics on spaces of plane curves. *Geometry, Imaging and Computing*, 1(1):1–56, 2014.
11. M. Bauer and P. Harms. Metrics on spaces of surfaces where horizontality equals normality. *http://arxiv.org/abs/1403.1436*, 2014.
12. M. Bauer, P. Harms, and P. W. Michor. Sobolev metrics on shape space of surfaces. *J. Geom. Mech.*, 3(4):389–438, 2011.
13. M. Bauer, P. Harms, and P. W. Michor. Almost local metrics on shape space of hypersurfaces in n-space. *SIAM J. Imaging Sci.*, 5(1):244–310, 2012.
14. 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.
15. M. Bruveris. Completeness properties of Sobolev metrics on the space of curves. *J. Geom. Mech.*, 2014.
16. M. Bruveris, P. W. Michor, and D. Mumford. Geodesic completeness for Sobolev metrics on the space of immersed plane curves. *Forum Math. Sigma*, 2:e19 (38 pages), 2014.
17. M. Eslitzbichler. Modelling character motions on infinite-dimensional manifolds. *The Visual Computer*, pages 1–12, 2014.
18. J. Glaunès, A. Qiu, M. I. Miller, and L. Younes. Large deformation diffeomorphic metric curve mapping. *International Journal of Computer Vision*, 80(3):317–336, 2008.
19. I. H. Jermyn, S. Kurtek, E. Klassen, and A. Srivastava. Elastic shape matching of parameterized surfaces using square root normal fields. In *Proceedings of the 12th European conference on Computer Vision - Volume Part V*, ECCV'12, pages 804–817, Berlin, Heidelberg, 2012. Springer-Verlag.
20. W. P. A. Klingenberg. *Riemannian Geometry*, volume 1 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 1995.
21. A. Kriegl and P. W. Michor. *The Convenient Setting for Global Analysis*. AMS, Providence, 1997. 'Surveys and Monographs 53'.
22. S. Kurtek, E. Klassen, Z. Ding, and A. Srivastava. A novel riemannian framework for shape analysis of 3D objects. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:1625–1632, 2010.
23. H. Laga, S. Kurtek, A. Srivastava, and S. J. Miklavcic. Landmark-free statistical analysis of the shape of plant leaves. *J. Theoret. Biol.*, 363:41–52, 2014.

24. S. Lahiri, D. Robinson, and E. Klassen. Precise matching of PL curves in R^N in the square root velocity framework, 2015.
25. S. Lang. *Fundamentals of Differential Geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999.
26. W. Liu, A. Srivastava, and J. Zhang. A mathematical framework for protein structure comparison. *PLoS Comput Biol*, 7(2):e1001075, 02 2011.
27. A. Mennucci, A. Yezzi, and G. Sundaramoorthi. Properties of Sobolev-type metrics in the space of curves. *Interfaces Free Bound.*, 10(4):423–445, 2008.
28. P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math.*, 10:217–245 (electronic), 2005.
29. P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS) 8 (2006), 1-48*, 2006.
30. 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.
31. W. Mio, A. Srivastava, and S. Joshi. On shape of plane elastic curves. *Int. J. Comput. Vision*, 73(3):307–324, July 2007.
32. G. Nardi, G. Peyré, and F.-X. Vialard. Geodesics on shape spaces with bounded variation and Sobolev metrics. *http://arxiv.org/abs/1402.6504*, 2014.
33. D. T. Robinson. *Functional data analysis and partial shape matching in the square root velocity framework*. PhD thesis, Florida State University, 2012.
34. J. Shah. H^0-type Riemannian metrics on the space of planar curves. *Quart. Appl. Math.*, 66(1):123–137, 2008.
35. J. Shah. An H^2 Riemannian metric on the space of planar curves modulo similitudes. *Adv. in Appl. Math.*, 51(4):483–506, 2013.
36. E. Sharon and D. Mumford. 2D-shape analysis using conformal mapping. *International Journal of Computer Vision*, 70:55–75, 2006.
37. A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE T. Pattern Anal.*, 33(7):1415–1428, 2011.
38. J. Su, S. Kurtek, E. Klassen, and A. Srivastava. Statistical analysis of trajectories on Riemannian manifolds: bird migration, hurricane tracking and video surveillance. *Ann. Appl. Stat.*, 8(1):530–552, 2014.
39. 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 J. Imaging Sci.*, 4(1):109–145, 2011.
40. Q. Xie, S. Kurtek, and A. Srivastava. Analysis of AneuRisk65 data: elastic shape registration of curves. *Electron. J. Stat.*, 8:1920–1929, 2014.
41. A. Yezzi and A. Mennucci. Conformal metrics and true "gradient flows" for curves. In *Proceedings of the Tenth IEEE International Conference on Computer Vision*, volume 1, pages 913–919, Washington, 2005. IEEE Computer Society.
42. L. Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.*, 58(2):565–586 (electronic), 1998.
43. 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.