Second order elastic metrics on the shape space of curves

Martin Bauer  
bauer.martin@univie.ac.at  
Technical University of Vienna

Martins Bruveris  
martins.bruveris@brunel.ac.uk  
Brunel University London

Philipp Harms  
philipp.harms@math.ethz.ch  
ETH Zürich

Jakob Møller-Andersen  
jakmo@dtu.dk  
Technical University of Denmark

Abstract

Second order Sobolev metrics on the space of regular unparametrized planar curves have several desirable completeness properties not present in lower order metrics, but numerics are still largely missing. In this paper, we present algorithms to numerically solve the initial and boundary value problems for geodesics. The combination of these algorithms allows to compute Karcher means in a Riemannian gradient-based optimization scheme. Our framework has the advantage that the constants determining the weights of the zero, first, and second order terms of the metric can be chosen freely. Moreover, due to its generality, it could be applied to more general spaces of mapping. We demonstrate the effectiveness of our approach by analyzing a collection of shapes representing physical objects.

1 Introduction

Unparametrized curves arise naturally in shape analysis and its many applications, including medical imaging [24, 32, 34], object tracking [30, 31], computer animation [5, 12], computer aided design [13], speech recognition [29], analysis of bird migration patterns and hurricane paths [28], biology [11, 18], and many other fields [3, 17]. In many instances, the rationale for identifying curves differing only by a reparametrization is that the curves represent the boundaries of physical shapes. Shapes can be analyzed mathematically by endowing the space of shapes with a Riemannian metric. Riemannian shape analysis has developed into an active field of research by now.

To do statistics on shape space, robust and efficient implementations of the boundary and initial value problems for geodesics are needed. Unfortunately, the arguably simplest metric on shape space is degenerate: it is well-known that the Riemannian distance induced by the $L^2$-metric vanishes on spaces of parametrized and unparametrized curves [19]. The discovery of this degeneracy led to an investigation of Sobolev metrics of higher order [20, 26, 33].

© 2015. The copyright of this document resides with its authors.  
It may be distributed unchanged freely in print or electronic forms.  
†All authors contributed equally to the article  
DOI: https://dx.doi.org/10.5244/C.29.DIFFCV.9
For the important class of first order metrics numerics are well developed by now, but a restriction on the parameters of the metric must be imposed [15, 27]. The situation is drastically different for second order metrics, despite the fact that they enjoy better completeness properties [8, 9]. On the positive side, the geodesic boundary value problem under second order Finsler metrics on the space of $BV^2$ curves was implemented numerically in [22]. Moreover, on the space of parametrized curves, there are numerics for the boundary and initial value problems for geodesics under second order Sobolev metrics [4,6]. On spaces of unparametrized curves numerics for second order Sobolev metrics are, however, still lacking. This is an important challenge and the topic of this paper.

We present a numerical implementation of the initial and boundary value problem for geodesics of planar, unparametrized curves under second order Sobolev metrics. Our implementation is based on a discretization of the Riemannian energy functional using B-splines. The boundary value problem for geodesics is solved by a gradient descent on the set of discretized paths and the initial value problem by discrete geodesic calculus [25]. Our approach is general in that it involves no restriction on the parameters of the metric and allows to factor out rigid transformations.

We tested our implementation on a dataset of shapes representing various groups of similar physical objects [14]. We were able to reconstruct the groups using agglomerative clustering based on the matrix of pairwise geodesic distances between the shapes. We also studied within-group shape variations by performing nonlinear principal component analysis using the Karcher means of the groups as base points. We visualized the results by solving the geodesic equation forward and backward in time along the principal directions.

Our algorithms performed well on the shapes in the database. Some care was needed to avoid singularities in the spline representations of the contours extracted from the binary images and of the paths initializing the gradient descent of the energy functional.

In future work, our framework could be applied to other spaces of mappings like manifold-valued curves, embedded surfaces, or more general spaces of immersions (see [1, 2] for details and [3] for a general overview). A rigorous convergence analysis of the proposed discretization remains to be done.

2 Mathematical background

We extend the exposition of [6] to unparametrized curves. The space of smooth, regular, parametrized curves is defined as

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

where Imm stands for immersion. $\text{Imm}(S^1, \mathbb{R}^d)$ is an open subset of the Fréchet space $C^\infty(S^1, \mathbb{R}^d)$ and as such itself a Fréchet manifold. The tangent space $T_c \text{Imm}(S^1, \mathbb{R}^d)$ at any curve $c$ is the vector space $C^\infty(S^1, \mathbb{R}^d)$. Geometrically, tangent vectors at $c$ are $\mathbb{R}^d$-valued vector fields along $c$.

We denote the Euclidean inner product on $\mathbb{R}^d$ by $\langle \cdot, \cdot \rangle$. Moreover, for any fixed curve $c$, we denote differentiation and integration with respect to arc length by $D_s = \frac{1}{|c_\theta|} \partial_\theta$ and $ds = |c_\theta| d\theta$, respectively.

**Definition 2.1.** Second order Sobolev metrics on $\text{Imm}(S^1, \mathbb{R}^d)$ are Riemannian metrics of the form

$$G_c(h,k) = \int_{S^1} a_0(h,k) + a_1(D_sh, D_sk) + a_2(D_s^2h, D_s^2k) \, ds,$$
where $a_0, a_2 > 0$, $a_1 \geq 0$ are constants and $h, k \in T_c \text{Imm}(S^1, \mathbb{R}^d)$ tangent vectors.

**Remark 2.2** (Choice of the constants). In the definition of Sobolev metrics, there is freedom in the choice of the relative weighting of the $L^2$-, $H^1$- and $H^2$-parts. This choice influences the geometry of the space of curves and should, in each application, be informed by the data at hand.

The reparametrization group is the diffeomorphism group of the circle,

$$\text{Diff}(S^1) = \{ \varphi \in C^\infty(S^1, S^1) : \varphi \text{ bijective} \},$$

which is an infinite-dimensional regular Fréchet Lie group $[16]$. Reparametrizations act on curves by composition from the right, $(c, \varphi) \rightarrow c \circ \varphi$. The shape space of unparametrized curves is the orbit space $\mathcal{S}(S^1, \mathbb{R}^d) = \text{Imm}(S^1, \mathbb{R}^d)/\text{Diff}(S^1)$ of this group action.¹

**Theorem 2.3** ([16], Sect. 1.5). The space $\mathcal{S}(S^1, \mathbb{R}^d)$ is a Fréchet manifold and the base space of the principal fibre bundle

$$\pi : \text{Imm}(S^1, \mathbb{R}^d) \rightarrow \mathcal{S}(S^1, \mathbb{R}^d), \quad c \mapsto c \circ \text{Diff}(S^1),$$

with structure group $\text{Diff}(S^1)$. A Sobolev metric $G$ on $\text{Imm}(S^1, \mathbb{R}^d)$ induces a metric on $\mathcal{S}(S^1, \mathbb{R}^d)$, such that the projection $\pi$ is a Riemannian submersion.

**Remark 2.4** (Invariance of the metric). Thm. 2.3 hinges on the invariance of Sobolev metrics with respect to reparametrizations. Sobolev metrics are also invariant with respect to translations and rotations, but in general not to scalings. The lack of scale invariance can be addressed by introducing weights depending on the length $\ell_c$ of the curve $c$ as follows:

$$\bar{G}_c(h, k) = \int_{S^1} \frac{a_0}{\ell_c^3}(h, k) + \frac{a_1}{\ell_c}(D_2 h, D_2 k) + a_2 \ell_c (D_2 h, D_2 k) \, ds.$$

Deformations of curves are smooth paths $c : [0, 1] \rightarrow \text{Imm}(S^1, \mathbb{R}^d)$. Their velocity is $c_t$, the subscript $t$ denoting differentiation. The length of a path $c$ is

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

The distance between two shapes $\pi(c_0)$ and $\pi(c_1)$ in $\mathcal{S}(S^1, \mathbb{R}^d)$ with respect to the metric induced by $G$ is the infimum over all paths between $c_0$ and the orbit $c_1 \circ \text{Diff}(S^1)$, i.e.,

$$\text{dist}(c_0, c_1) = \inf_{c(0) = c_0, \ c(1) \in c_1 \circ \text{Diff}(S^1)} L(c).$$

Geodesics on $\text{Imm}(S^1, \mathbb{R}^d)$ are critical points of the energy functional

$$E(c) = \frac{1}{2} \int_0^1 G_{c(t)}(c_t(t), c_t(t)) \, dt.$$  

The spaces $\mathcal{S}(S^1, \mathbb{R}^d)$ and $\text{Imm}(S^1, \mathbb{R}^d)$ are related by a Riemannian submersion and geodesics on $\mathcal{S}(S^1, \mathbb{R}^d)$ can be lifted to horizontal geodesics on $\text{Imm}(S^1, \mathbb{R}^d)$. Conversely, the projection of a length-minimizing path between $c_0$ and $c_1 \circ \text{Diff}(S^1)$ is a geodesic on $\mathcal{S}(S^1, \mathbb{R}^d)$.

The space $\mathcal{S}(S^1, \mathbb{R}^d)$ equipped with a Sobolev metric possesses some nice completeness properties, which are summarized in the following theorem.

¹More precisely, we define $\mathcal{S}(S^1, \mathbb{R}^d) = \text{Imm}_0(S^1, \mathbb{R}^d)/\text{Diff}(S^1)$, where $\text{Imm}_0(S^1, \mathbb{R}^d)$ is the subset of free immersions, i.e., those upon which $\text{Diff}(S^1)$ acts freely. This is an open and dense subset of $\text{Imm}(S^1, \mathbb{R}^d)$. While important for theoretical reasons, this restriction has no influence on the practical applications of Sobolev metrics.
Theorem 2.5 (Minimizing geodesics, [8, 3]). Let $a_0, a_2 > 0$ and $a_1 \geq 0$. Then, given two curves $c_0, c_1$ in the same connected component of $\text{Imm}(S^1, \mathbb{R}^d)$, there exists a minimizing geodesic connecting them. Furthermore, there exists a minimizing geodesic connecting the shapes $\pi(c_0)$ and $\pi(c_1)$ in $S(S^1, \mathbb{R}^d)$.

Remark 2.6 (Elastic metrics). Closely related to the Sobolev metrics described here is the family of elastic metrics [21], which in the planar case is given by

$$G_c(h,k) = \int_s a^2 \langle D_sh, n \rangle \langle D_sk, n \rangle + b^2 \langle D_sh, v \rangle \langle D_sk, v \rangle ds.$$ 

Here $a, b$ are constants and $v, n$ denote the unit tangent and normal vectors to $c$. Two special cases deserve to be highlighted: for $a = 1$, $b = \frac{1}{2}$ [21] and $a = b$ [3] there exist nonlinear transforms, the square root velocity transform and the basic mapping, that greatly simplify the numerical computation of geodesics. Both of these metrics have been applied to a variety of problems in shape analysis. We note that the elastic metric with $a = b$ corresponds to a first order Sobolev metric as in Def. 2.1 with $a_0 = a_2 = 0$ and $a_1 = a^2 = b^2$. As it has no $L^2$-part, it is a Riemannian metric only on the space of curves modulo translations.

3 Numerical implementation

3.1 Discretization

We discretize curves using B-splines; $c = \sum_{j=1}^{N_\theta} d_j C_j(\theta)$, where $C_j$ are the B-splines of degree $n_\theta$, defined on a uniform periodic knot sequence, with all knots of multiplicity one. Observe that $C_j \in C^{n_\theta-1}([0, 2\pi])$. A path of curves can then be represented using tensor product B-splines, i.e.,

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

(1)

Here $B_i$ are the B-splines defined on the interval $[0, 1]$, with uniform knots and full multiplicity at the end points, $B_i \in C^{n_\theta-1}([0, 1])$. This implies that the boundary curves are given by $c(0, \theta) = \sum_{j=1}^{N_\theta} d_{1,j} C_j(\theta)$ and $c(1, \theta) = \sum_{j=1}^{N_\theta} d_{N_t,j} C_j(\theta)$.

Under the identification of $S^1$ with $\mathbb{R}/[0, 2\pi]$, diffeomorphisms $\psi : S^1 \to S^1$ can be written as $\psi = \text{id} + \phi$, where $\phi$ is a periodic function. We choose to discretize $\phi(\theta) = \sum_{i=1}^{N_\phi} \phi_i D_i(\theta)$, where $D_i$ are B-splines of degree $n_\phi$, defined on a uniform periodic knot sequence, similarly to $C_j$. The identity can be written in a B-spline basis using the Greville abscissas $\xi_i$, i.e.,

$$\text{id} = \sum_{i=1}^{N_\phi} \xi_i D_i.$$ 

Due to the positivity of B-splines, the condition $\psi' > 0$ ensuring that $\psi$ is a diffeomorphism takes the form

$$\phi_{i-1} - \phi_i < \xi_i - \xi_{i-1}.$$ 

(2)

To speed up convergence, we introduce an additional variable $\alpha \in \mathbb{R}$ representing constant shifts of the reparametrization. The resulting redundancy is eliminated by the constraint

$$\sum_{i=1}^{N_\phi} \phi_i = 0.$$ 

(3)

To compute the energy of paths (1), we use Gaussian quadrature on each interval between subsequent knots. Evaluating paths (and their derivatives) at the quadrature points is a simple multiplication of the spline collocation matrix with the vector of control points.
3.2 Geodesics and Karcher means

From now on we work with plane curves ($d = 2$). The boundary value problem for geodesics consists of minimizing the discretized energy (2) over all paths $c: [0, 1] \times [0, 2\pi] \rightarrow \mathbb{R}^2$, mappings $\psi = \text{id} + \phi: S^1 \rightarrow S^1$, shifts of the reparametrization $\alpha \in \mathbb{R}$, and rotations $R_\beta$ around the origin by an angle $\beta \in [0, 2\pi)$, subject to the constraints (2), (3), and

$$c(0, \cdot) = c_0(\cdot), \quad c(1, \cdot) = R_\beta(c_1(\psi(\cdot) - \alpha) + v),$$

where $c_0, c_1: [0, 2\pi] \rightarrow \mathbb{R}^2$ are given boundary curves. This is a finite dimensional constrained optimization problem, which we solve using Matlab’s interior point method \texttt{fmincon}. We achieved major performance improvements by fine-tuning the implementations of the gradient and hessian of the energy functional.

In the initial value problem for geodesics, the initial value and initial velocity of the sought geodesic are given. As described in Sect. 2, geodesics on $S(S^1, \mathbb{R}^d)$ can be lifted to horizontal geodesics on $\text{Imm}(S^1, \mathbb{R}^d)$. Thus, solving the geodesic initial value problem for unparametrized curves reduces to solving the problem for parametrized curves with horizontal initial velocities. To solve the latter problem, we use the time-discrete variational geodesic calculus of [25] as described in [6].

The Karcher mean $\bar{c}$ of a set $\{c_1, \ldots, c_n\}$ of curves is defined as the minimizer of

$$F(c) = \frac{1}{n} \sum_{j=1}^{n} \text{dist}(c, c_j)^2.$$  

(4)

It can be computed by iteratively solving initial and boundary value problems for geodesics. We refer to [23] for more details.

4 Numerical examples

4.1 Data acquisition and setup

We tested our implementation on a dataset of shapes collected by the Computer Vision Group at Brown university [14]. The dataset consists of black and white images of physical objects. It is natural to represent the objects by their boundaries using unparametrized curves. In addition to factoring out reparametrizations, we also factor out translations and rotations because we are not interested in the position of the curves in space. Some of the resulting
curves are depicted in Fig. 1. We used splines of degree $n_\theta = n_\phi = 3$ with $N_\theta = 60$ and $N_\phi = 20$ controls in space and of degree $n_t = 2$ with $N_t = 20$ controls in time.

To solve the boundary value for geodesics as described in Sect. 3.2, one has to construct an initial homotopy between the given boundary curves. If the curves differ only by a small deformation, this is unproblematic because the linear path connecting the curves can be used. For larger deformations, however, the linear path often has self-intersections, changing the winding number, which can lead to convergence problems during the energy minimization. Our solution was to construct homotopies by deforming the initial curve to a circle and the circle to the target curve using linear paths. This worked well for all examples presented in this article. It also led to the same optima in all cases where the linear path was free of self-intersections. In future work, we plan to investigate further ways of creating good and natural initial paths.

The choice of parameters $a_0$, $a_1$, and $a_2$ of the Riemannian metric can have a large influence on the resulting optimal deformations, as can be seen in Fig. 2. In this article, we used the following ad-hoc strategy for choosing the constants: we computed the average $L^2$-, $H^1$- and $H^2$-contributions $\bar{E}_0$, $\bar{E}_1$, $\bar{E}_2$ to the energy of linear paths between each pair of curves in the dataset. Then we normalized $a_0$ to 1 and chose $a_1$ and $a_2$ such that

$$a_0 \bar{E}_0 : a_1 \bar{E}_1 : a_2 \bar{E}_2 = 1 : 1 : 1$$

and

$$\bar{E} = a_0 \bar{E}_0 + a_1 \bar{E}_1 + a_2 \bar{E}_2 = 100.$$  

This resulted in the parameter values $a_1 = 250$ and $a_2 = 0.004$. It remains open how to best choose the constants depending on the data under consideration.

### 4.2 Comparison to non-elastic metrics

Riemannian metrics on spaces of unparametrized curves are often called elastic metrics, as they allow both for bending and stretching of the curve. In the elastic case, solving the boundary value problem for geodesics involves optimizing over the Diff($S^1$)-orbit of the initial or final shape. This is a computationally expensive and difficult task since the diffeomorphism group is infinite-dimensional.

An alternative and simpler approach is to parametrize the curves by unit-speed and to calculate geodesics in the space of parametrized curves. This could be called a non-elastic approach. (Of course, one might wish to also factor out rigid transformations and constant shifts of the parametrization, but this is much simpler because these groups are finite-dimensional.)

Figure 2: Geodesics between a fish and a tool in the space of unparametrized curves. The metric parameter $a_2$ is increased by a factor 10 in the second, a factor 100 in the third, and a factor 1000 in the fourth column. The corresponding geodesic distances are 138.02, 162.55, 246.46 and 468.74. Note that since we also optimize over translations and rotations of the target curve, the position in space varies.
Figure 3: Geodesics in the space of unparametrized (1st, 3rd) versus parametrized (2nd, 4th) curves modulo rotations and translations. Note that since we also optimize over translations and rotations of the target curve, the curves in the first two panels are aligned differently.

Our experiments suggest that using the more involved elastic approach pays off. The two approaches yield different results, and in particular in cases involving large amounts of stretching the geodesics found using the first approach appear more natural. However, in cases involving mainly bending of the curves the results are very similar (see Fig. 3).

4.3 Clustering and principal component analysis

We investigated if pairwise geodesic distances can be used to cluster shapes into meaningful groups. To this aim, we calculated all pairwise distances between the 33 shapes presented in Fig. 1. Solving the corresponding 528 boundary value problems took about two hours on a 3GHz processor with four cores. The resulting distance matrix is visualized in Fig. 4 using multi-dimensional scaling; the plot suggests that objects of the same group lie close together. Indeed, agglomerative clustering with 4 clusters reproduces exactly the subgroups of the database.

Finally, we studied within-group variations using non-linear principal component analysis. To this aim, we first computed the Karcher mean of each group. The corresponding optimization problem (4) was solved using a conjugate gradient method, as implemented in the Manopt library [7], on the finite-dimensional spline approximation of the Riemannian manifold of curves. The mean shapes for the groups of fish and humans can be seen in Fig. 5.
Next, we represented each shape in the group by the initial velocity from the mean \( \bar{c} \) using the inverse of the Riemannian exponential map. We then performed a principal component analysis with respect to the inner product \( G_{\bar{c}} \) on the set of initial velocities. In the group of human figures, the first three eigenvalues capture 67%, 22%, and 6% of within-group variation. In the group of fish, the first three eigenvalues capture only 40%, 25%, and 16% of within-group variation. Geodesics from the mean in the directions of the first two principal directions can be seen in Fig. 5. In the group of humans the first principal direction encodes bending of the arms and legs, whereas the second direction reflects stretching in the extremities.

5 Conclusions

In this article we developed a numerical framework for solving the initial and boundary value problem for geodesics of planar, unparametrized curves under second order Sobolev metrics. We tested our implementation on a dataset of shapes representing various groups of similar physical objects and obtained good experimental results. In future work, we plan to apply our algorithms to datasets of real-world medical data, prove rigorous convergence results for the discretization, and extend the framework to other spaces of mappings like manifold-valued curves and embedded surfaces.

Acknowledgements

We want to thank Peter W. Michor and Jens Gravesen for helpful discussions and valuable comments. All authors have been supported by the programme “Infinite-Dimensional Riemannian Geometry with Applications to Image Matching and Shape Analysis” held at the Erwin Schrödinger Institute. M. Bauer was supported by the European Research Council (ERC), within the project 306445 (Isoperimetric Inequalities and Integral Geometry) and by the FWF-project P24625 (Geometry of Shape spaces).
References


