Curve Matching with Applications in Medical Imaging

Martin Bauer\textsuperscript{1}, Martins Bruveris\textsuperscript{2}, Philipp Harms\textsuperscript{3}, Jakob Møller-Andersen\textsuperscript{4}**

\textsuperscript{1} Institut für Diskrete Mathematik und Geometrie, TU Wien
\textsuperscript{2} Department of Mathematics, Brunel University London
\textsuperscript{3} Department of Mathematics, ETH Zürich
\textsuperscript{4} Department of Applied Mathematics and Computer Science, Technical University of Denmark

Abstract. In the recent years, Riemannian shape analysis of curves and surfaces has found several applications in medical image analysis. In this paper we present a numerical discretization of second order Sobolev metrics on the space of regular curves in Euclidean space. This class of metrics has several desirable mathematical properties. We propose numerical solutions for the initial and boundary value problems of finding geodesics. These two methods are combined in a Riemannian gradient-based optimization scheme to compute the Karcher mean. We apply this to a study of the shape variation in HeLa cell nuclei and cycles of cardiac deformations, by computing means and principal modes of variations.

Keywords: Curve matching, Sobolev metrics, Riemannian shape analysis, discrete geodesics, minimizing geodesics

1 Introduction

The comparison and analysis of geometric shapes plays an important role in medical imaging [27], biology [9] as well as many other fields [11]. Spaces of geometric shapes are inherently nonlinear. To make standard methods of statistical analysis applicable, one can linearize the space locally around each shape. This can be achieved by introducing a Riemannian structure, which is able to describe jointly the global nonlinearity of the space as well as its local linearity.

Over the past decade, Riemannian shape analysis has become an active area of research in pure and applied mathematics. Driven by applications, a variety of spaces, equipped with different Riemannian metrics, have been used to represent geometrical shapes and their attributes. Ideally, the particular choice of metric should be dictated by the data at hand rather than by mathematical or numerical

\* All authors were partially supported by the Erwin Schrödinger Institute programme: Infinite-Dimensional Riemannian Geometry with Applications to Image Matching and Shape Analysis. 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).

\** All authors contributed equally to the article.
convenience. This leads to the task of developing efficient numerical methods for the statistical analysis of shapes for general and flexible classes of metrics.

The topic of this paper are second order Sobolev metrics on the space of regular, planar curves. This space and its quotients by translations, rotations, scalings, and reparametrizations are important and widely used spaces in shape analysis. Second order Sobolev metrics are mathematically well-behaved: the geodesic equation is globally well-posed, any two curves in the same connected component can be connected by a minimizing geodesic, the metric completion consists of all \( H^2 \)-immersions, and the metric extends to a strong Riemannian metric on the metric completion [7,8].

We provide the first numerical implementation of the initial and boundary value problems for geodesics with respect to these metrics and apply them to medical data. Our implementation allows us to compute Karcher means, principal components, and clusters of curves in reasonable time. The parameters in the metric can be chosen freely because we do not rely on transforms which exist only for special choices of parameters [23, 26]. We are also able to factor out the action of the finite-dimensional translation and rotation groups.

We illustrate the behaviour of the metrics in two medical applications. In Sect. 4.1 we use images of HeLa cell nuclei from [5] and compute the mean shape of the nucleus and the principal modes of shape variation; in Sect. 4.2 we use traces of cardiac image sequences from [12,24] and compare the mean shape of patients with Tetralogy of Fallot with the mean shape of a control group.

The paper is structured as follows: Sect. 2 contains the definition of second order Sobolev metrics and some relevant mathematical background. Section 3 describes our discretization of the geodesic equation and the Riemannian energy functional. Finally, in Sect. 4, we apply the metrics to medical imaging data.

2 Mathematical Background

In this article we center our attention on the space of smooth, regular curves with values in \( \mathbb{R}^d \).

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

where \( \text{Imm} \) stands for immersion. As an open subset of the Fréchet space \( C^\infty(S^1, \mathbb{R}^d) \), it is a Fréchet manifold. Its tangent space, \( T_c \text{Imm}(S^1, \mathbb{R}^d) \), at any curve \( c \) is the vector space \( C^\infty(S^1, \mathbb{R}^d) \).

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 \( D_s = \frac{1}{|c|} \partial_0 \) and \( ds = |c| d\theta \) respectively.

**Definition 1.** A second order Sobolev metric on \( \text{Imm}(S^1, \mathbb{R}^d) \) is given by

\[
G_c(h,k) = \int_{S^1} a_0 \langle h, k \rangle + a_1 \langle D_s h, D_s k \rangle + a_2 \langle D^2_s h, D^2_s k \rangle \, ds ,
\]

Implementations of second order Sobolev metrics can be found in [15] and [3], but none of the previous implementations seem to have been used in applications.
where $a_0, a_2 > 0$, $a_1 \geq 0$, and $h, k \in T_c \text{Imm}(S^1, \mathbb{R}^d)$ are tangent vectors.

**Remark 2 (Invariance of the metric).** The metric $G$ is invariant with respect to translations, rotations, and reparametrizations by diffeomorphisms of $S^1$, but not with respect to scalings. It can be made scale-invariant by introducing weights that depend on the length $\ell_c$ of the curve $c$. A scale-invariant metric is given by

$$\tilde{G}_c(h, k) = \int_{S^1} \frac{a_0}{\ell_c} \langle h, k \rangle + \frac{a_1}{\ell_c} \langle D_s h, D_s k \rangle + a_2 \ell_c \langle D_s^2 h, D_s^2 k \rangle ds .$$  \hspace{1cm} (3)

Depending on the application, it can be desirable to factor out some of these isometry groups. For example, we consider the HeLa cell nuclei in Sect. 4.1 as curves modulo translations and rotations. For the traces of cardiac images in Sect. 4.2, however, both the absolute position of the curve and its orientation have intrinsic meaning and should therefore not be factored out.

**Remark 3 (Choosing the constants).** The constants $a_0, a_1$ and $a_2$ in the definition of the metric determine the relative weight of the $L^2$, $H^1$, and $H^2$-parts. Their choice is a non-trivial and important task, and it should, in each application, be informed by the data at hand. The influence of different parameter values on geodesics and Karcher means can be seen in Fig. 1 and Fig. 2.

Fig. 1: Geodesics between a cat and a cow. The metric parameter $a_2$ in the first row is increased by a factor 10 in the second and a factor 100 in the third row.

**Remark 4 (Generalizations to higher-dimensions).** Sobolev metrics have natural generalizations to manifold-valued curves, embedded surfaces, and more generally spaces of immersions of a compact manifold $M$ into a Riemannian manifold $N$; see [1, 4] for details and [2] for a general overview.

Deformations of curves can be seen as smooth paths $c: [0, 1] \to \text{Imm}(S^1, \mathbb{R}^d)$. Their velocity is $c_t$, the subscript $t$ denoting differentiation.
Definition 5 (Geodesic distance). The length of a path of curves $c$ is

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

(4)

The distance between two curves in $\text{Imm}(S^1, \mathbb{R}^2)$ (with respect to the metric $G$) is the infimum of the lengths of all paths connecting these curves.

$$\text{dist}(c_0, c_1) = \inf_{c(0)=c_0, c(1)=c_1} L(c) .$$

Geodesics are locally distance-minimizing paths.

Geodesics can be described by a partial differential equation, called the geodesic equation. It is the first order condition for minima of the energy

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

(5)

Recently, some local and global existence results for geodesics of Sobolev metrics were shown in [7,8,13]. Since they provide the theoretical underpinnings for the numerical methods presented in this paper, we summarize them here.

Theorem 6 (Geodesic equation). Let $a_0, a_2 > 0$ and $a_1 \geq 0$. The geodesic equation of the metric $G$, written in terms of the momentum $p = |c'| (a_0 c_t - a_1 D^2 c_t + a_2 D^4 c_t)$, is given by

$$\partial_t p = -\frac{a_0}{2} \langle \phi | D_s (\langle c_t, c_t \rangle) D_s c \rangle + \frac{a_1}{2} \langle \phi | D_s (\langle D_s c_t, D_s c_t \rangle) D_s c \rangle - \frac{a_2}{2} \langle \phi | D_s (\langle D^2_s c_t, D^2_s c_t \rangle) D_s c \rangle .$$

(6)

Given any initial condition $(c_0, u_0) \in T \text{Imm}(S^1, \mathbb{R}^2)$, the solution of the geodesic equation exists for all time.

If, however, $a_0, a_1 > 0$ and $a_2 = 0$, then $G$ is a first order Sobolev metric. Its geodesic equation is locally, but not globally, well-posed.

Remark 7 (Comparison to elastic metrics). Closely related are elastic metrics [14], which in the planar case are given by

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

(7)
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 \( a = 1, b = \frac{1}{2} \) [23] and \( a = b \) [26] there exist nonlinear transforms, the square root velocity transform and the basic mapping, that greatly simplify numerics. 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. 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 Discretization

We propose to represent paths of curves as tensor product B-splines

\[
c(t, \theta) = \sum_{i=1}^{N_t} \sum_{j=1}^{N_\theta} c_{i,j} B_i(t) C_j(\theta)
\]

of degree \( n_t \) in time \((t)\) and \( n_\theta \) in space \((\theta)\), respectively, using \( N_t \) and \( N_\theta \) basis functions and controls \( c_{i,j} \in \mathbb{R}^d \). The spatial basis functions \( C_j \) are defined on the uniform knot sequence \( \theta_j = \frac{j-n}{N_\theta}, 0 \leq j \leq 2n + N_\theta \), and satisfy periodic boundary conditions. The time basis functions \( B_i \) are defined using uniform knots on the interior of the interval \([0,1]\) and with full multiplicity on the boundary. Full multiplicity of the boundary knots ensures that fixing the initial and final curves is equivalent to fixing the control points \( c_{1,j} \) and \( c_{N_t,j} \). Regarding smoothness, we have \( B_i \in C_{n_t-1}([0,1]) \) and \( C_j \in C_{n_\theta-1}([0,2\pi]) \).

The proposed discretization scheme is independent of the special form of the metric, and can be applied equally to the scale-invariant versions as well as to the family of elastic metrics.

3.1 Boundary Value Problem for Geodesics

To evaluate the energy (5) on discretized paths (8), we use Gaussian quadrature on each interval between subsequent knots. The resulting discretized energy \( E_{\text{discr}}(c_{1,1}, \ldots, c_{N_t,N_\theta}) \) is minimized over all discrete paths (8) with fixed initial and final control points \( c_{1,j} \) and \( c_{N_t,j} \). This is a finite-dimensional, nonlinear, unconstrained optimization problem.

We solved this problem using either Matlab’s \texttt{fminunc} function with gradients computed by finite differences, or using AMPL [10] and the Ipopt solver [25] with gradients computed by automatic differentiation. By classical approximation results [22], the discrete energy of discrete paths converges to the continuous energy of smooth paths as the number of control points tends to infinity. Hence we expect to obtain close approximations of geodesics. Establishing rigorous convergence results for our discretization will be the subject of future work.

\(^6\) We neglect the condition \( c_\theta(t, \theta) \neq 0 \). All smooth paths violating this condition have infinite energy because of the completeness of the metric. By approximation, spline paths violating this condition have very high energy and will, in practice, be avoided during optimization.
3.2 Computing the Karcher Mean

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

$$F(c) = \frac{1}{n} \sum_{j=1}^{n} \text{dist}(c, c_j)^2 . \quad (9)$$

It can be calculated by a gradient descent on $(\text{Imm}(S^1, \mathbb{R}^d), G)$. Letting $\text{Log}_c c_j$ denote the Riemannian logarithm, the gradient of $F$ with respect to $G$ is $\{17\}$

$$\text{grad}^G F(c) = \frac{1}{n} \sum_{j=1}^{n} \text{Log}_c c_j . \quad (10)$$

3.3 Initial Value Problem for Geodesics

To calculate the Karcher mean by gradient descent, one has to repeatedly solve the geodesic equation (6). To this aim, we use the time-discrete variational geodesic calculus $[21]$. Given three curves $c_0, c_1, c_2$, one defines the discrete energy $E_2(c_0, c_1, c_2)$ as follows:

$$E_2(c_0, c_1, c_2) = G_{c_0}(c_1 - c_0, c_1 - c_0) + G_{c_1}(c_2 - c_1, c_2 - c_1) . \quad (11)$$

A 3-tuple $(c_0, c_1, c_2)$ is a discrete geodesic if it is a minimizer of the discrete energy $E_2$ with fixed endpoints $c_0, c_2$. The discrete exponential map is defined as follows: $c_2 = \text{Exp}_{c_0} c_1$, if $(c_0, c_1, c_2)$ is a discrete geodesic, in other words, if $c_1 = \text{argmin} E_2(c_0, \cdot, c_2)$. To find $c_2$, we differentiate the discrete energy $E_2$ with respect to $c_1$ and solve the resulting system of nonlinear equations,

$$2G_{c_0}(c_1 - c_0, \cdot) - 2G_{c_1}(c_2 - c_1, \cdot) + D_{c_1} G(c_2 - c_1, c_2 - c_1) = 0 . \quad (12)$$

We use the solver `fsolve` in Matlab to solve this system of equations.

Given a time resolution $\Delta t$ and an initial velocity $v$, we choose $c_1 = c_0 + v \Delta t$ and compute $c_2 = \text{Exp}_{c_0} c_1$. The procedure is repeated with $c_1, c_2$ in order to compute $c_3 = \text{Exp}_{c_1} c_2$, which is iterated until we reach the required final time.

While the convergence results in $[21]$ do not apply to our setting, we found good experimental agreement with the solutions of the boundary value problem.

4 Applications

4.1 Hela Cells

In our first example we want to characterize nuclear shape variation in HeLa cells. We use fluorescence microscope images of HeLa cell nuclei$^7$ (87 images in total). The acquisition of the cells is described in $[5]$. A similar study on this dataset was performed using the LDDMM framework in $[19,20]$; applying intrinsically

$^7$ The dataset was downloaded from http://murphylab.web.cmu.edu/data.
defined Sobolev metrics to the same problem will provide a complimentary point of view.

To extract the boundary of the nucleus, we apply a thresholding method [16] to obtain a binary image and then fit – using least squares – a spline with $N_\theta = 12$ and $n_\theta = 4$ to the longest 4-connected component of the thresholded image. Then we reparametrize the boundary to approximately constant speed. The remaining degree of freedom is the starting point of the parametrization; when computing minimizing geodesics, we minimize over this parameter as well. The shapes of cell nuclei are thus represented by curves modulo translations, rotations, and constant shifts of the parametrisation, i.e. the two curves $c$ and $e^{\alpha\theta}c(-\alpha) + \lambda$ are considered equivalent. Examples of the extracted curves are depicted in Fig. 3.

Fig. 3: Eight boundaries of HeLa cell nuclei, the Karcher mean of all nuclei (enlarged), and six randomly sampled cells using a Gaussian distribution in normal coordinates with the same covariance as the data.

The parameters $a_0$, $a_1$, and $a_2$ of the Riemannian metric are chosen as follows: we compute the average $L^2$, $H^1$- and $H^2$-contributions $E_{L^2}$, $E_{H^1}$, $E_{H^2}$ to the energy of linear paths between each pair of curves in the dataset. As the $L^2$- and $H^2$-contributions scale differently, we rescale all curves such that $E_{L^2} = E_{H^2}$. Then we choose constants $a_0$, $a_1$, and $a_2$ such that

$$a_0 E_{L^2} : a_1 E_{H^1} : a_2 E_{H^2} = 3 : 1 : 6 \quad \text{and} \quad E = a_0 E_{L^2} + a_1 E_{H^1} + a_2 E_{H^2} = 100,$$

resulting in an average length $\bar{\ell}_c = 12.45$ and $a_0 = 3.36$, $a_1 = 2.20$, and $a_2 = 6.73$.

Fig. 4: Geodesics from the mean in the first four principal directions. The curves show the geodesic at times $-3, -2, \ldots, 2, 3$; the bold curve is the mean. One can see four different characteristic deformations of the cell: bending, stretching along the long axis, stretching along the short axis, and a combination of stretching with partial bending.
The average shape of the nucleus can be captured by the Karcher mean $\tau$. To compute the Karcher mean of the 87 nuclei, we use a conjugate gradient method on the Riemannian manifold of curves, as implemented in the Manopt library [6], to solve the minimization problem (9). We obtain convergence in 28 steps; the final value of the objective function (9) is $F(\tau) = 10.55$ and the norm of the gradient is $\|\text{grad}^G F(\tau)\|_\tau < 10^{-3}$. The mean shape can be seen in Fig. 3.

Having computed the mean $\tau$, we represent each nuclear shape $c_j$ by the initial velocity $v_j = \text{Log}_{c_\tau} c_j$ of the minimal geodesic connecting $c_\tau$ and $c_j$. We perform principal component analysis with respect to the inner product $G_\tau$ on the set of initial velocities $\{v_j : j = 1, \ldots, 87\}$. The first four eigenvalues are $4.10, 2.39, 1.68$ and $1.00$, and they explain $38.04\%, 60.21\%, 75.78\%$, and $85.07\%$ of the total variance. Geodesics from the mean in the directions of the first four principal directions can be seen in Fig. 4. Fig. 5 shows the data projected to the subspace spanned by the first two eigenvectors.

Finally, we sample from a normal distribution with the same covariance matrix as the data and project the sampled velocities $\hat{v}$ back to the space of curves using the exponential map $\hat{c} = \text{Exp}_c \hat{v}$; some examples can be seen in Fig. 3.

![Fig. 5: All 87 cell nuclei projected to the plane in the tangent space which is spanned by the first two principal directions. The mean (in blue) is situated at the origin. The units on the coordinate axes are standard deviations. We see that the first coordinate is related to the bending of the nucleus, while the second coordinate is related to its elongation.](image)

4.2 Traces of Cardiac Images

In our second application we study curves that are obtained from images of the cardiac cycle. More precisely, we consider a sequence of 30 cardiac images,
taken at equispaced time points along the cardiac cycle. Each image is projected to a barycentric subspace of dimension two, yielding a closed curve in the two-dimensional space of barycentric coordinates. After normalizing the coordinates [18, Sect. 3] we obtain a closed, plane curve – with the curve parameter representing time – to which we can apply the methods presented in Sect. 2. Details regarding the acquisition and projection of the images can be found in [12,24]; barycentric subspaces on manifolds are described in [18].

The data consists of 10 cardiac cycles of patients with Tetralogy of Fallot and 9 patients from a control group. Each cardiac cycle is originally represented by three-dimensional homogeneous coordinates $x_1 : x_2 : x_3$, sampled at 30 time points. We project the homogeneous coordinates onto the plane $x_1 + x_2 + x_3 = 1$ and choose a two-dimensional coordinate system for this plane. Then we use spline interpolation with degree $n_\theta = 3$ and $N_\theta = 30$ control points to reconstruct the planar curves from the data points; see Fig. 6.

![Fig. 6: Projections to a two-dimensional barycentric subspace of 30 images from the cardiac cycles of three patients. Cubic splines interpolation of degree $n_\theta = 3$ with $N_\theta = 30$ control points is used.](image)

The parameters $a_0$, $a_1$, and $a_2$ in the metric are chosen similarly to Sect. 4.1; however, the scale of the curves is not changed and we use equal weighting between the $L^2$-, $H^1$- and $H^2$-parts of the average energy for linear paths. This leads to parameters $a_0 = 1$, $a_1 = 0.1$, and $a_2 = 10^{-9}$. To see if the metric structure derived from the Sobolev metric enables us to distinguish between diseased patients and the control group, we compute all 171 pairwise distances between the 19 curves; this takes about 15 minutes on a 2 GHz single core processor. Multi-dimensional scaling of the distance matrix shows that the metric separates healthy and diseased patients quite well (Fig. 7a). Indeed, a cluster analysis based on the distance matrix recovers exactly – with exception of one outlier (patient 4) – the subgroups of healthy and diseased patients (Fig. 7b).

The Karcher means of the healthy and diseased subgroups as well as of the entire population are depicted in Fig. 8. The mean was computed using a gradient descent method as described in Sect. 3.2 with a threshold of $10^{-4}$ for the norm of the gradient. The average distance from the mean for the diseased group is 0.6853 with a variance of 0.0149, and for the control group the distance is 0.7708 with a variance of 0.0083.

To investigate the variability of the observed data, we perform principal component analysis on the initial velocities of the minimal geodesics connecting
Fig. 7: (a) Two dimensional representation of the data using multi dimensional scaling of the pairwise distance matrix. (b) A dendrogram of clusters computed from the pairwise distance matrix using the single linkage criterion. Healthy patients are labelled 1–9 and diseased ones 10–19.

Fig. 8: First row: Karcher means of pathological cardiac cycles (left), all cycles (middle), and healthy cycles (right). Second row: geodesic connecting the Karcher mean of pathological cycles to the Karcher mean of healthy cycles. The crosses denote the position of images, with respect to whom the barycentric projection was computed.

curves to the respective means (c.f. Sect. 4.1). Fig. 9 shows the initial velocities projected to the subspace spanned by the first two principal directions. Within the healthy and sick subgroups, less than 30% of the principal components are needed to explain 90% of the shape variation. If, in contrast, principal components are analyzed for the entire dataset based on the global Karcher mean, then 35% of the principal components are needed to explain 90% of the shape variation.

5 Conclusions

In this article we numerically solved the initial and boundary value problems for geodesics on the space of regular curves under second order Sobolev metrics. We analyzed two medical datasets using our approach. In future work we plan to...
prove rigorous convergence results for our discretizations, further investigate the impact of the constants in the metric, and treat unparametrized curves.

Acknowledgments

We would like to thank Xavier Pennec and Marc-Michel Rohé for providing us the cardiac image data and Hermann Schichl for his invaluable help with AMPL.

References