Geodesic Finite Elements on Simplicial Grids

Oliver Sander*

June 30, 2011

Abstract

We introduce geodesic finite elements as a conforming way to discretize partial differential equations for functions $v : \Omega \to M$, where Ω is an open subset of $\mathbb{R}^d$ and M is a Riemannian manifold. These geodesic finite elements naturally generalize standard first-order finite elements for Euclidean spaces. They also generalize the geodesic finite elements proposed for $d = 1$ in [24]. Our formulation is equivariant under isometries of M, and hence preserves objectivity of continuous problem formulations. We concentrate on partial differential equations that can be formulated as minimization problems. Discretization leads to algebraic minimization problems on product manifolds M^n. These can be solved efficiently using a Riemannian trust-region method. We propose a monotone multigrid method to solve the constrained inner problems with linear multigrid speed. As an example we numerically compute harmonic maps from a domain in $\mathbb{R}^3$ to S^2.

1 Introduction

Many problems in mathematical physics can be written as partial differential equations (PDEs) for functions

$$f : N \to M,$$

where N and M are Riemannian manifolds. In the vast majority of cases these manifolds will be open subsets of Euclidean spaces. Such problems are treated successfully using finite elements.

In this article we focus on the case that N is an open subset of $\mathbb{R}^d$, but M is a nonlinear Riemannian manifold. Such problems cannot be discretized using finite elements, because the standard definition of finite element functions presupposes a vector space structure on M. We give a more general definition which also encompasses the nonlinear case.[1]

Instances of such problems are, for example, the simulation of nematic liquid crystals. There, depending on the symmetry of the substance in question, functions describing crystal configurations take values in the unit sphere S^2, the projective plane $\mathbb{RP}^2$, or the special orthogonal group SO(3) [13]. In the

*This work was supported by the DFG research center Matheon

[1] The case that N is nonlinear is treated by the theory of geometric partial differential equations [14]. The problems due to nonlinearity of the domain N and the range space M are unrelated and need different techniques for their solution.

simulation of superfluid helium-3 more exotic spaces such as certain quotients of $S^2 \times \mathbb{RP}^3$ occur [20].

A different set of applications is the numerical treatment of Cosserat materials. These are mechanical models of continua, where, in addition to a position, each point has an associated orientation [23]. The value space M is hence the special Euclidean group $\mathrm{SE}(3) = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$. A treatment of one-dimensional Cosserat models by geodesic finite elements has been presented in [24]. Finally, PDEs for manifold-valued functions also occur in some image processing applications. For example, the diffusion equation for functions with values in S^2 can be used to denoise color images [26].

Various ways have been proposed in the literature to discretize problems for functions with values in a manifold M. The central problem is how to generalize linear interpolation, i.e., how to interpolate between a set of values $v_i \in M$, $i = 1, \dots, d+1$. One simple way is to use an embedding of M in a Euclidean space $\mathbb{R}^m$ [3]. The values are then interpolated in $\mathbb{R}^m$, and possibly projected back onto M. While this is simple, and in many cases cheap and even objective, the result depends on the embedding.

Alternatively, one can single out a tangent space T_pM of M, and retract the values v_i onto T_pM using the exponential map. The retracted values are then interpolated on T_pM, and projected back onto M (see, e.g., [22]). This works only as long as the v_i stay away from the cut locus of p. Also, the dependence on a fixed tangent space T_pM breaks objectivity.

A third approach notes that the interpolation function is needed only at a fixed set of quadrature points. The values there can be treated as additional variables. In the framework of an iterative solver they are initialized with known values, and only corrections—which live in linear spaces—are ever interpolated. This method was used by Simo and Vu-Quoc [25] to simulate Cosserat rods. However, as was shown later [12], the method introduces spurious dependencies of the solution on the initial iterate and the parameters of the path-following mechanism.

In this work we follow a new approach by using the Riemannian center of mass to generalize linear interpolation. The Riemannian center of mass is well-known in the differential geometry literature (see, e.g., [5] and the references it contains). Its usefulness for data interpolation on manifolds has been recognized by Buss and Fillmore [10] for $M = S^m$ and by Moakher [21] for $M = \mathrm{SO}(3)$. We extend this idea to obtain conforming first-order finite elements for general M, which we call geodesic finite elements. The result is an elegant theory which is completely intrinsic and covariant. The price is an increased algorithmic difficulty of handling the new interpolation functions.

We focus on PDEs that are the Euler–Lagrange equations of some energy functional $\mathcal{J}$.[2] Numerically, this means that one has to find minimizers of $\mathcal{J}$ over the discrete nonlinear finite element space V_h^M. The corresponding algebraic problem is the minimization of a functional J over a product manifold M^n, with n the number of grid vertices. For such problems, Absil et al. [1] have proposed the Riemannian trust-region method, which generalizes the standard trust-region approach to nonlinear configuration spaces, while retaining the convergence properties of regular trust-region methods. At each step a quadratic minimization problem with convex inequality constraints has to be solved. In

[2] The discretization of nonconservative equations will be the subject of a separate paper.

the case of finite elements, these problems are sparse, but possibly very large. We propose a monotone multigrid method as a way to solve the quadratic problems with the speed of a linear multigrid method.

The implicit definition of geodesic finite element functions by means of the Riemannian center of mass makes their numerical handling more involved than the handling of regular finite element functions. To evaluate a function value, a small minimization problem has to be solved. Function gradients can be obtained by additionally solving a small linear system of equations. This makes geodesic finite elements more expensive than other approaches. Still, we believe that the good theoretical properties and general elegance of our approach can outweigh the additional costs and complexity of the implementation.

The content of this article is as follows. Chapter 2 introduces simplicial geodesic interpolation, a generalization of linear interpolation to functions with values in a Riemannian manifold. This new concept is then used in Chapter 3 to construct geodesic finite element spaces. Chapters 4 and 5 deal with the numerical solution of the algebraic minimization problems. Chapter 4 introduces the Riemannian trust-region method and the monotone multigrid inner solver, while Chapter 5 shows how geodesic finite element functions and various of their derivatives can be evaluated algorithmically. Chapter 6 brings a numerical example, computing harmonic maps from a domain in $\mathbb{R}^3$ to the unit sphere S^2. Various formulae for unit spheres needed for the implementation are collected in an appendix.

2 Simplicial Geodesic Interpolation

The definition of a finite element space consists of two parts, namely an interpolation rule on the reference element and a way to piece together the local functions at the element boundaries to form a global space. In this chapter we treat the local problem by introducing a generalization of linear interpolation on the reference element to functions with values in Riemannian manifolds. The next chapter will then introduce the global spaces.

2.1 Definition

To motivate our definition we briefly review linear interpolation, which is at the heart of first-order finite elements. Let

$$\widetilde{\Delta}^d = \Big\{ \xi \in \mathbb{R}^d \;\Big|\; \xi_i \geq 0,\ i = 1, \dots, d,\ \sum_{i=1}^{d} \xi_i \leq 1 \Big\}$$

be the d-dimensional reference simplex. We omit the superscript d when it is clear from the context. In the finite element literature, an affine function $\tilde{v} : \widetilde{\Delta} \to \mathbb{R}$ is defined by its values $v_1, \dots, v_{d+1}$ at the simplex corners and the interpolation formula

$$\tilde{v}(\xi) = \sum_{i=1}^{d+1} v_i \varphi_i(\xi), \qquad \xi \in \widetilde{\Delta}, \tag{1}$$

where the functions

$$\varphi_i : \widetilde{\Delta} \to \mathbb{R}, \qquad \varphi_i(\xi) = \begin{cases} (1 - \sum_{j=1}^d \xi_j) & \text{if } i = 1, \\ \xi_{i-1} & \text{else,} \end{cases}$$

are the first-order Lagrangian shape functions on the reference simplex.

It is unclear how an interpolation formula like (1) would look like if $\tilde{v}$ were to take its values in a nonlinear manifold M, because the sum in (1) assumes a vector space structure on M. We therefore rewrite (1) to obtain a form that is more easily generalized to nonlinear spaces. For this, we introduce the d-dimensional standard simplex

$$\Delta^d = \Big\{ w \in \mathbb{R}^{d+1} \;\Big|\; w_i \geq 0,\, i = 1, \dots, d+1,\, \sum_{i=1}^{d+1} w_i = 1 \Big\},$$

whose corners are the canonical basis vectors in $\mathbb{R}^{d+1}$, which we denote by $e_1, \dots, e_{d+1}$. Coordinates $w = \{w_1, \dots, w_{d+1}\}$ on the standard simplex are called barycentric coordinates and can be computed from coordinates ξ on the reference simplex by

$$w = \mathcal{B}(\xi) = B\xi + c = \begin{pmatrix} -\mathbf{1}^T \\ \mathrm{Id}_{d\times d} \end{pmatrix} \xi + \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \tag{2}$$

with $-\mathbf{1}^T$ the row vector consisting of d entries of -1, and $\mathbf{0}$ the column vector consisting of d entries of 0. In barycentric coordinates w, the interpolation formula (1) reads

$$\tilde{v}(\xi) = v(w) = \sum_{i=1}^{d+1} v_i w_i. \tag{3}$$

We still need a sum to compute $v(w)$, but the shape functions are now simply the coordinate functions w_i. To also get rid of the sum in (3) we note that

$$\sum_{i=1}^{d+1} v_i w_i = \operatorname*{arg\,min}_{q\in\mathbb{R}} \sum_{i=1}^{d+1} w_i \|v_i - q\|^2.$$

Now the sum adds norms instead of values, and we can generalize this expression to nonlinear spaces.

Definition 2.1. *Let M be a connected Riemannian manifold and* $\operatorname{dist}(\cdot,\cdot) : M \times M \to \mathbb{R}$ *a distance metric on M. For values $v_1, \dots, v_{d+1} \in M$ we call*

$$\begin{aligned} \Upsilon \;:\; & M^{d+1} \times \Delta \to M \\ \Upsilon(v_1, \dots, v_{d+1}; w) &= \operatorname*{arg\,min}_{q\in M} \sum_{i=1}^{d+1} w_i \operatorname{dist}(v_i, q)^2 \end{aligned} \tag{4}$$

simplicial geodesic interpolation on M.

The construction is illustrated in Figure 1. For brevity we will frequently write $\Upsilon(v, w)$ instead of $\Upsilon(v_1, \dots, v_{d+1}; w)$.

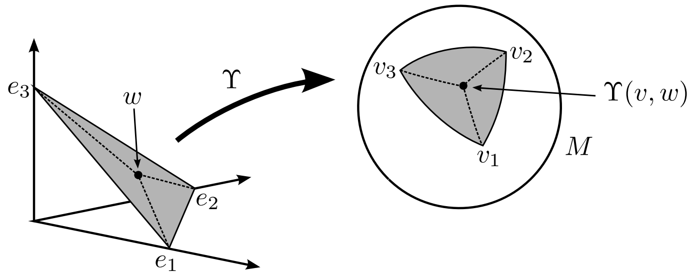

Figure 1: Geodesic simplicial interpolation from Δ^2 into a sphere. The simplex corners $e_1, \dots, e_3$ are mapped to the values $v_1, \dots, v_3$.

Remark 2.1. Strictly speaking we have not used the manifold structure of M, and Definition 2.1 would also make sense in more general metric spaces. The value of such a generalization is unclear.

As it turns out, simplicial geodesic interpolation is a special case of a more general concept known as the Riemannian center of mass [5].

Definition 2.2. *Let M be a Riemannian manifold and da a measure of unit weight. Then the Riemannian center of mass is defined as*

$$\mathcal{C}(da) = \operatorname*{arg\,min}_{q \in M} \int_M \operatorname{dist}(a,q)^2 \, da.$$

To see the relationship to the simplicial geodesic interpolation between $d+1$ values v_i on M, let $w \in \Delta$ be coordinates and define the discrete measure

$$d\mu_w = \sum_{i=1}^{d+1} w_i \delta_{v_i}, \qquad \delta_{v_i}(X) = \begin{cases} 1 & \text{if } v_i \in X, \\ 0 & \text{else,} \end{cases}$$

which has unit weight. Then $\mathcal{C}(d\mu_w) = \Upsilon(v,w)$.

2.2 Well-Posedness of the Definition

Using the Riemannian center of mass for the definition of a function $\Upsilon(v,\cdot) : \Delta \to M$ is only meaningful if the minimizer in (4) is well-defined. However it is not immediately clear from Definition 2.1 whether this is the case. Indeed, it is easy to construct a case where the function

$$f_{v,w} : M \to \mathbb{R}, \qquad f_{v,w}(q) = \sum_{i=1}^{d+1} w_i \operatorname{dist}(v_i, q)^2 \tag{5}$$

in (4) does not have a minimizer. An example is $M = \mathbb{R}^2 \setminus \{0\}$, $d = 2$, $w_i = 1/3$, $i = 1,2,3$, and values v_i at equal mutual distances on the unit circle. We avoid such cases in the following by requiring M to be complete. If M is compact then $f_{v,w}$ must have a minimum. However, this minimum need not be unique. For example, set M the unit sphere S^2, $d = 2$, $w_i = 1/3$, $i = 1,2,3$, and v_1, v_2, v_3 placed on the equator at equal distances. Then both poles minimize $f_{v,w}$.

Existence and uniqueness of a minimizer can be obtained, however, if the corner values are "close together" in a certain sense. The precise conditions have been given by Karcher [16]. Remember that a set $D \subset M$ is called convex if for each $p, q \in D$ the minimizing geodesic from p to q is entirely contained in D.

Theorem 2.1 (Karcher [16])**.** *Let M be complete, B_ρ an open geodesic ball of radius ρ in M, and $v_1, \dots, v_{d+1} \in M$. Assume that $v_i \in B_\rho$ for all $i = 1, \dots, d+1$.*

1. *If the sectional curvatures of M in B_ρ are bounded by a positive constant K, and $\rho < \frac{1}{4}\pi K^{-1/2}$, then the function*

$$f_{v,w}(q) = \sum_{i=1}^{d+1} w_i \operatorname{dist}(v_i, q)^2$$

has a unique minimizer in B_ρ for all $w \in \Delta^d$.

2. *If the sectional curvatures of M in B_ρ are at most 0, then $f_{v,w}$ has a unique minimizer in B_ρ for all $w \in \Delta^d$.*

Note that under the given assumptions the balls B_ρ are convex [9, Thm. II.1.4]. If M has nonpositive curvature and is simply connected we know from the theorem of Hadamard that geodesic balls of arbitrary size exist. Hence, together with the second part of Theorem 2.1 we get the following corollary.

Corollary 2.1. *Let M be complete, simply connected and have nonpositive sectional curvatures. Then for all $v_1, \dots, v_{d+1} \in M$, $w \in \Delta$, the functional $f_{v,w}$ has a unique minimum in M.*

The requirement for the v_i to be "close together" for Υ to be well-defined is not a serious restriction in a finite element context. There, many properties are expected to hold on sufficiently fine grids only, anyways. Theorem 3.2 will show that the conditions of Theorem 2.1 are fulfilled if the grid is fine enough. The reader should compare this to the one-dimensional theory in [24].

Karcher's proof of Theorem 2.1 contains a subresult which we will state separately in order to be able to refer to it later.

Lemma 2.1. *Under the assumptions of Theorem 2.1 the Hessian of $f_{v,w}$ exists and is invertible on B_ρ.*

2.3 Geometrical Properties

Since $\operatorname{dist}(p, q) = \|p - q\|$ in the Euclidean spaces $\mathbb{R}^m$, Definition 2.1 reproduces linear interpolation if $M = \mathbb{R}^m$ for any $m \in \mathbb{N}$. It also generalizes the interpolation along geodesics introduced in [24]. This justifies the name *simplicial geodesic interpolation* for Υ.

Lemma 2.2. *Let $p_1, p_2 \in M$ and let $\gamma : [0, 1] \to M$ be the minimizing geodesic from p_1 to p_2, assumed to be unique and parametrized by arc length. Then, for any $w = (w_1, w_2) \in \Delta^1$ we have $\Upsilon(p_1, p_2; w) = \gamma(w_2)$.*

Proof. Let $w \in \Delta^1$ be fixed and assume without loss of generality that $\operatorname{dist}(p_1, p_2) = 1$. We will compute a minimizer for the functional

$$f_w : M \to \mathbb{R}, \qquad f_w(q) = w_1 \operatorname{dist}(p_1, q)^2 + w_2 \operatorname{dist}(p_2, q)^2.$$

Introducing the new variables $\alpha = \operatorname{dist}(p_1, q)$ and $\beta = \operatorname{dist}(p_2, q)$ turns this into a minimization problem for a functional

$$\tilde{f}_w : \mathbb{R}^2 \to \mathbb{R}, \qquad \tilde{f}_w(\alpha, \beta) = w_1 \alpha^2 + w_2 \beta^2,$$

subject to the constraints

$$\alpha, \beta \geq 0, \qquad \text{and} \qquad \alpha + \beta \geq 1. \tag{6}$$

The constraints represent the fact that α and β are distances. Note that not all pairs α, β can actually be realized by a $q \in M$. The graph of $\tilde{f}_w$ is a paraboloid centered at the origin of $\mathbb{R}^2$, hence its minimizer subject to (6) will satisfy $\alpha + \beta = 1$. To compute the minimizer exactly note that the gradient of $\tilde{f}_w$ at the minimizer will be perpendicular to the line $\{\alpha, \beta \mid \alpha + \beta = 1\}$, or, in other words, a multiple of $(1, 1)^T$. Since $\nabla \tilde{f}_w = (2w_1\alpha, 2w_2\beta)^T$ this is the case if $\alpha = w_2$ and $\beta = w_1$. But such α, β can be realized, namely by a point q on the minimizing geodesic from p_1 to p_2 with $\operatorname{dist}(a, q) = w_2$ and $\operatorname{dist}(q, b) = w_1$. This proves the assertion. □

Our next result characterizes the images of the faces of Δ under a geodesic simplicial interpolation mapping.

Lemma 2.3. *Let*

$$\Delta_i^{d-1} = \Delta^d \cap \{w \in \mathbb{R}^{d+1} \mid w_i = 0\}$$

be the i-th $d-1$-dimensional face of the standard simplex Δ^d, and let $\Upsilon_v(w) : \Delta^d \to M$ be the geodesic interpolation of the fixed values $v_1, \dots, v_{d+1} \in M$. Then its restriction $\Upsilon_v|_{\Delta_i^{d-1}}$ is also a simplicial geodesic interpolation, namely of the values of the corners of Δ_i^{d-1}.

Proof. From the definition of geodesic simplicial interpolation we know that

$$\Upsilon_v(w) = \operatorname*{arg\,min}_{q \in M} \sum_{j=1}^{d+1} w_j \operatorname{dist}(v_j, q)^2.$$

The restriction to Δ_i^{d-1} is given by

$$\Upsilon_v|_{\Delta_i^{d-1}}(w) = \operatorname*{arg\,min}_{q \in M} \sum_{\substack{j=1 \\ j \neq i}}^{d+1} w_j \operatorname{dist}(v_j, q)^2, \tag{7}$$

and we have $\sum_{\substack{j=1 \\ j \neq i}}^{d+1} w_j = 1$. Hence (7) is the definition of a geodesic interpolation function for the coordinates $(w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_{d+1})$ and values $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{d+1})$. □

Applying this argument recursively we get the following generalization.

Corollary 2.2. *Let δ be any face of any dimension of Δ. Then $\Upsilon_v|_\delta$ is a function defined by simplicial geodesic interpolation between the values at the corners of δ.*

Together with Lemma 2.2 we find in particular that the images of the edges of the standard simplex under a geodesic simplicial interpolation mapping are geodesics in M.

Next we discuss the smoothness of the map Υ. For the finite element method we need derivatives of Υ with respect to w and the coefficients v_i. Their existence is based on the smoothness of the squared Riemannian distance. Remember that the injectivity radius $\operatorname{inj} p$ at a point p of a manifold M is the largest number r such that minimizing geodesics from p to q are unique if $\operatorname{dist}(p, q) < r$.

Lemma 2.4. *Let M be a complete Riemannian manifold and $p \in M$. Then the squared Riemannian distance $\operatorname{dist}(p, \cdot)^2 : M \to \mathbb{R}$ is infinitely differentiable at all $q \in M$ with $\operatorname{dist}(p, q) < \operatorname{inj} p$.*

Proof. It suffices to show this in one coordinate system. Under the assumptions there exists a normal coordinate chart centered at p and containing q. Let $(q_1, \dots, q_m)$ be the coordinates of q. Then by Corollary 6.11 from [18] we get $\operatorname{dist}(p, q)^2 = \sum_{i=1}^m q_i^2$, which is C^∞. □

Differentiability of Υ now follows from the implicit function theorem.

Theorem 2.2. *Let M be complete and let B_ρ be an open geodesic ball with $v_1, \dots, v_{d+1} \in B_\rho$. We assume that ρ is small enough such that the curvature assumptions of Theorem 2.1 hold, and that $\rho < 1/2 \operatorname{inj} v_i$ for all $i = 1, \dots, d+1$. Then the function*

$$\Upsilon(v_1, \dots, v_{d+1}; w) \ : \ M^{d+1} \times \Delta \to M$$

is infinitely differentiable with respect to the v_i and w.

Proof. By definition, the interpolation values $\Upsilon(v_1, \dots, v_{d+1}; w)$ are given as minimizers of the functional

$$f_{v,w}(q) = \sum_{i=1}^{d+1} w_i \operatorname{dist}(v_i, q)^2.$$

By Theorem 2.1 minimizers exist and are unique. By the assumptions and Lemma 2.4, $\operatorname{dist}(v_i, \cdot)^2$ is smooth on B_ρ for all v_i, and so is $f_{v,w}$. The function value of $\Upsilon(v, w)$ is alternatively described as a zero of the gradient vector field

$$F(v, w, q) := \nabla f_{v,w}(q).$$

The function $F(v, w, q)$ is smooth in w and the v_i. Also, its derivative with respect to q exists and is invertible. To see this note that $\frac{\partial F}{\partial q} = \operatorname{Hess} f_{v,w}$, and use Lemma 2.1. Hence by the implicit function theorem the function $\Upsilon(v, w)$, for which

$$F(v, w, \Upsilon(v, w)) = 0$$

holds, is also smooth with respect to w and the v_i. □

2.4 Symmetry Properties

Simplicial geodesic interpolation enjoys various symmetry properties expected from a finite element interpolation procedure. The first is a symmetry of the domain, the proof of which is evident. For finite element applications this symmetry means that it is irrelevant how the corners of an element in a grid are numbered.

Lemma 2.5. *Simplicial geodesic interpolation is invariant under permutation of the vertices of the standard simplex, i.e.,*

$$\Upsilon(\pi(v_1), \dots, \pi(v_{d+1}); \pi(w_1), \dots, \pi(w_{d+1})) = \Upsilon(v_1, \dots, v_{d+1}; w_1, \dots, w_{d+1})$$

for all permutations π of the vertices of Δ^d.

Remark 2.2. The desirability of this simple invariance forces us to disregard an alternative interpolation method that appears in [19]. There, in a different context, an interpolation procedure based on successive interpolation along geodesics is proposed. While this is much easier to evaluate, it lacks the above invariance, as follows from [7].

The second important symmetry is equivariance of the interpolation under isometries of M. In mechanics this property is known as frame-invariance or objectivity. The result for the general Riemannian center of mass has already been stated in [16]. We give a simple proof for the readers' convenience.

Lemma 2.6. *Let M be a complete Riemannian manifold and G a group that acts on M by isometries. Let $v_1, \dots, v_{d+1} \in M$ be such that the assumptions of Theorem 2.1 hold. Then*

$$Q\Upsilon(v_1, \dots, v_{d+1}; w) = \Upsilon(Qv_1, \dots, Qv_{d+1}; w)$$

for all $w \in \Delta$ and $Q \in G$.

Proof. Set $v^* = \Upsilon(Qv_1, \dots, Qv_{d+1}; w)$ for ease of notation and assume that

$$v^* \neq Q\Upsilon(v_1, \dots, v_{d+1}; w). \tag{8}$$

By the definition of Υ this means that

$$\sum_{i=1}^{d+1} w_i \operatorname{dist}(Qv_i, v^*)^2 < \sum_{i=1}^{d+1} w_i \operatorname{dist}\big(Qv_i, Q\Upsilon(v_1, \dots, v_{d+1}; w)\big)^2,$$

because, by Theorem 2.1, the minimum v^* is unique in a geodesic ball containing the Qv_i. However, since Q is an isometry, we get

$$\sum_{i=1}^{d+1} w_i \operatorname{dist}(v_i, Q^{-1}v^*)^2 < \sum_{i=1}^{d+1} w_i \operatorname{dist}\big(v_i, \Upsilon(v_1, \dots, v_{d+1}; w)\big)^2.$$

This is a contradiction, because by definition $\Upsilon(v_1, \dots, v_{d+1}; w)$ minimizes $f_{v,w}(\cdot) = \sum_i w_i \operatorname{dist}(v_i, \cdot)^2$ in a geodesic ball containing the v_i, and, again by Theorem 2.1, it is unique there. Hence the assumption (8) must be wrong and the lemma is proved. $\square$

As an example, note that this invariance is enjoyed by standard scalar finite elements. Indeed, let $M = \mathbb{R}$ with the canonical metric. Then $G = \mathbb{R}$ acts isometrically on $\mathbb{R}$ by addition. Using formula (3) for linear interpolation we see that for any group element Q_a acting by addition of the scalar a we have

$$\begin{aligned}\Upsilon(Q_a v, w) &= \Upsilon(a + v_1, \dots, a + v_{d+1}; w) \\ &= \sum_{i=1}^{d+1} w_i(a + v_i) = a + \sum_{i=1}^{d+1} w_i v_i = Q_a \Upsilon(v, w).\end{aligned}$$

In Cosserat mechanics, where $M = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$, the group G of isometries is $\mathbb{R}^3 \rtimes \mathrm{SO}(3)$ itself. Hence Lemma 2.6 states equivariance under translations and rotations of three-space. This means that discretization by geodesic finite elements will not destroy the frame-invariance of a continuous mechanics model.

3 Geodesic Finite Elements

In this section we use the interpolation method presented above to construct global finite element spaces. We show that these spaces are conforming in the sense that they are subsets of $H^1(\Omega, M)$, and we discuss the relationship between geodesic finite element functions and coefficient vectors. Finally we extend the equivariance result of the previous section (Lemma 2.6) to global geodesic finite element functions.

From now on let Ω be an open bounded subset of $\mathbb{R}^d$, $d \geq 1$. For simplicity we assume that Ω has a polygonal boundary. Let $\mathcal{G}$ be a conforming grid for Ω consisting of simplices only. The number of grid vertices shall be denoted by n.

Definition 3.1 (Geodesic Finite Elements)**.** *Let $\mathcal{G}$ be a simplicial grid on Ω, and let M be a complete Riemannian manifold. We call $v_h : \Omega \to M$ a geodesic finite element function for M if it is continuous, and for each element $T \in \mathcal{G}$ the restriction $v_h|_T$ is a geodesic simplicial interpolation in the sense that*

$$v_h|_T(x) = \Upsilon\big(v_{T,1}, \dots, v_{T,d+1}; \mathcal{F}_T(x)\big),$$

where $\mathcal{F}_T : T \to \Delta$ is affine and the $v_{T,i}$ are values in M. The space of all such functions v_h will be denoted by V_h^M.

With the help of Lemma 2.2 it is seen directly that this definition provides a generalization of the one-dimensional geodesic finite elements proposed in [24]. Setting $M = \mathbb{R}^m$ for $m \in \mathbb{N}$ we also recover the definition of standard first-order finite elements. On the other hand, the well-posedness of Definition 3.1 is again unclear, as we inherit the corresponding well-posedness problems from the definition of geodesic simplicial interpolation. We will see below, when we discuss the relationship between geodesic finite element functions and coefficient vectors, that the spaces V_h^M do contain sufficiently many functions for finite element analysis.

We begin our investigations by showing that geodesic finite element functions are conforming. We first introduce Sobolev spaces for manifold-valued functions (see, e.g., [6]).

Definition 3.2. *Let M be a Riemannian manifold isometrically embedded in $\mathbb{R}^m$ for some $m \in \mathbb{N}$. Then*

$$H^k(\Omega, M) := \{v \in H^k(\Omega, \mathbb{R}^m) \mid v(x) \in M \text{ a.e.}\}$$

is called the k-th order Sobolev space for functions with values in M.

It is now easy to show that geodesic finite element functions are indeed Sobolev functions.

Theorem 3.1. $V_h^M(\Omega) \subset H^1(\Omega, M)$.

The proof uses the following simple lemma, which can be shown similarly to Theorem 5.2 in [8].

Lemma 3.1. *Let $k \geq 1$, $m \geq 1$, and Ω open and bounded. A function $v : \Omega \to \mathbb{R}^m$ that is piecewise arbitrarily smooth is in $H^k(\Omega, \mathbb{R}^m)$ if and only if $v \in C^{k-1}(\Omega, \mathbb{R}^m)$.*

Proof of Theorem 3.1. Let $v_h \in V_h^M(\Omega)$. By Lemma 2.2, v_h is differentiable on each element T of $\mathcal{G}$ as a function $T \to M$. By Nash's theorem there exists a smooth isometric embedding of M into an $\mathbb{R}^m$, and then v_h is also differentiable on each element T as a function from T to $\mathbb{R}^m$. By definition, v_h is continuous and with Lemma 3.1 we can conclude that $v_h \in H^1(\Omega, \mathbb{R}^m)$. But also by definition we have $v_h(x) \in M$ for all $x \in \Omega$, and hence $v_h \in H^1(\Omega, M)$. $\square$

The classical linear finite element method distinguishes the discrete problem, which deals with finite element functions, from the algebraic problem, which deals with vectors of coefficients. The latter is used to implement numerical algorithms. Both formulations are equivalent, because a classical finite element function uniquely corresponds to a coefficient vector once a basis of the finite element space has been chosen. In the simplest case the basis is the nodal basis and the coefficients are the function values at the grid vertices.

The distinction between discrete and algebraic formulations persists in the theory of geodesic finite elements. However, the relationship between geodesic finite element functions $v_h \in V_h^M$ and sets of coefficients $\bar{v} \in M^n$ is more subtle. Since any $v_h \in V_h^M$ is continuous we can associate to it the coefficient set consisting of the values of v_h at the grid vertices. However, given a set of coefficients $\bar{v} \in M^n$ it is not clear whether there is a corresponding geodesic finite element function, and whether this function is unique, if there is one. The difficulty stems mainly from the corresponding problem for geodesic simplicial interpolation, but it is also not obvious whether individual geodesic interpolation functions can be stitched together continuously.

To formally investigate the relationship between geodesic finite element functions and sets of coefficients we define the nodal evaluation operator

$$\begin{aligned} \mathcal{E} \;:\; V_h^M &\to M^n \\ (\mathcal{E}(v_h))_i &= v_h(x_i), \qquad x_i \text{ the } i\text{-th vertex of } \mathcal{G}. \end{aligned}$$

To each geodesic finite element function $v_h \in V_h^M$ it associates the set of function values at the grid vertices. Since functions in V_h^M are continuous the operator $\mathcal{E}$ is well-defined and single-valued for all $v_h \in V_h^M$.

We are now interested in the inverse operator $\mathcal{E}^{-1}$, which associates geodesic finite element functions to a given set of coefficients. Since M is supposed to be geodesically complete, the inverse $\mathcal{E}^{-1}$ is defined for all $\bar{v} \in M^n$; however for arbitrary $\bar{v}$ it may be multi-valued.[3] Using Theorem 2.1 the sets of coefficients for which $\mathcal{E}^{-1}$ is single-valued can be characterized.

Lemma 3.2. *Let $\mathcal{G}$ be a simplicial grid with n vertices, and let $\bar{v} \in M^n$ be a set of coefficients. Set $\rho = \frac{1}{4}\pi K^{-1/2}$ if the sectional curvatures of M are bounded from above by a positive constant K, and set $\rho = \infty$ if the sectional curvatures are bounded from above by 0. If for each element T of $\mathcal{G}$ with associated coefficients $\bar{v}_{T,1}, \dots, \bar{v}_{T,d+1}$ there exists an open geodesic ball B_ρ of radius ρ such that $\bar{v}_{T,i} \in B_\rho$ for each $i = 1, \dots, d+1$, then there is a unique function $v_h \in V_h^M$ with $v_h = \mathcal{E}^{-1}(\bar{v})$.*

Proof. For each $x \in \Omega$ let T be an element of $\mathcal{G}$ with $x \in T$ and corner values $\bar{v}_{T,i}$, $i = 1, \dots, d+1$. Set $\mathcal{F}_T$ an affine mapping from T onto the standard simplex Δ^d and define the function $v_h : \Omega \to M$ by

$$v_h(x) = \Upsilon(\bar{v}_{T,1}, \dots, \bar{v}_{T,d+1}; \mathcal{F}_T(x)).$$

By the assumptions on the grid and the coefficients we can invoke Theorem 2.1 to get that v_h is single-valued on each T. Since

$$\Upsilon(\bar{v}_{T,1}, \dots, \bar{v}_{T,d+1}; e_i) = \bar{v}_{T,i} \qquad e_i = (\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{d+1-i})$$

we have that $\mathcal{E}(v_h) = \bar{v}$. Also, by Lemma 2.3, v_h varies continuously between adjacent elements. Hence $v_h \in V_h^M$. □

Corollary 2.1 has the following global analogon.

Corollary 3.1. *Let M be complete, simply connected, and have nonpositive sectional curvatures. Then $\mathcal{E} : V_h^M \to M^n$ is a bijection.*

If M has positive, but bounded sectional curvatures, the unique correspondence between geodesic finite element functions and sets of coefficients can still be obtained, but only in the following asymptotic sense.

Theorem 3.2. *Let M be a complete Riemannian manifold with sectional curvatures bounded from above by a constant $K > 0$, and let $v : \Omega \to M$ be Lipschitz continuous in the sense that there exists a constant L such that*

$$\operatorname{dist}(v(x), v(y)) \leq L|x - y|$$

for all $x, y \in \Omega$. Let $\mathcal{G}$ be a simplicial grid of Ω and h the length of the longest edge of $\mathcal{G}$. Set $\bar{v} = \mathcal{E}(v)$, tacitly extending the definition of $\mathcal{E}$ to all continuous functions $\Omega \to M$. If

$$h < \frac{\rho}{L} = \frac{\pi K^{-1/2}}{4L},$$

then the inverse of $\mathcal{E}$ has only a single value in V_h^M for each $\tilde{v}$ in a neighborhood of $\bar{v}$.

[3] An example of this is given in Figure 3 in [24].

Proof. Let T be an element of $\mathcal{G}$ with corners $x_{T,1}, \dots, x_{T,d+1} \in \overline{\Omega}$, and let $v_{T,1}, \dots, v_{T,d+1} \in M$ be the values of v at these corners. Set $\rho = \frac{1}{4}\pi K^{-1/2}$ and let $B_\rho(v_{T,1})$ be the geodesic open ball of radius ρ around $v_{T,1}$. Then for each $i = 1, \dots, d+1$ we have $v_{T,i} \in B_\rho(v_{T,1})$ because

$$\operatorname{dist}(v_{T,i}, v_{T,1}) \leq L|x_{T,i} - x_{T,1}| \leq Lh < \rho.$$

Hence we can use Lemma 3.2 to conclude that there is a unique function $v_h \in V_h^M$ with $\mathcal{E}(v_h) = \bar{v}$.

Since h is strictly less than ρ/L there is an $\epsilon \in (0,1)$ such that even $h < \epsilon\rho/L$. Then $\operatorname{dist}(v_{T,i}, v_{T,1}) \leq \epsilon\rho$ for all $1 \leq i \leq d+1$. Define $\epsilon^* = (1-\epsilon)\rho$ and set $B_{\epsilon^*}(v_{T,i})$ the geodesic ball of radius ϵ^* around $v_{T,i}$. Then for any $\tilde{v}_{T,i} \in B_{\epsilon^*}(v_{T,i})$, $i = 1, \dots, d+1$ we have, by the triangle inequality,

$$\operatorname{dist}(\tilde{v}_{T,i}, v_{T,1}) \leq \operatorname{dist}(\tilde{v}_{T,i}, v_{T,i}) + \operatorname{dist}(v_{T,i}, v_{T,1}) \leq (1-\epsilon)\rho + \epsilon\rho = \rho.$$

Hence

$$B_{\epsilon^*,\infty}(\bar{v}) := \{\tilde{v} \in M^n \mid \tilde{v}_i \in B_{\epsilon^*}(\bar{v}_i),\ i = 1, \dots, n\}$$

is a neighborhood of $\bar{v}$ in M^n, and for each $\tilde{v} \in B_{\epsilon^*,\infty}(\bar{v})$ there is a unique function $\tilde{v}_h$ in V_h^M with $\mathcal{E}(\tilde{v}_h) = \tilde{v}$. □

This lemma implies that for a given problem with a Lipschitz-continuous solution we can always find a grid fine enough such that we can disregard the distinction between V_h^M and M^n in the vicinity of the solution. Hence locally a geodesic finite element problem can be represented by a corresponding algebraic problem on the product manifold M^n. Locally, the function space V_h^M inherits the manifold structure of M^n, because functions defined by simplicial geodesic interpolation depend differentiably on their corner values (Lemma 2.2).

We close the chapter with the result that geodesic finite element functions are equivariant under isometries of M. It is a central result of geodesic finite element theory, because it implies that the invariance of a continuous model under isometries will not be lost by discretization. In mechanics, e.g., it allows for discrete approximations that are exactly frame-indifferent. The proof is trivial with the help of Lemma 2.6.

Theorem 3.3. *Let M be a Riemannian manifold and G a group that acts on M by isometries. Extend the action of G to M^n by setting*

$$(Q\bar{v})_i = Q\bar{v}_i \qquad \forall \bar{v} \in M^n, \quad i = 1, \dots, n,$$

and to V_h^M by setting

$$(Qv_h)(x) = Q(v_h(x)) \qquad \forall v_h \in V_h^M, \quad x \in \Omega.$$

Then

$$\mathcal{E}(Qv_h) = Q\mathcal{E}(v_h)$$

for all $Q \in G$ and $v_h \in V_h^M$, whenever these expressions are well defined.

4 Numerical Solution of Minimization Problems in Geodesic Finite Element Spaces

In the previous two chapters we have introduced geodesic finite element spaces as a conforming approximation for $H^1(\Omega, M)$, where Ω is an open subset of $\mathbb{R}^d$, $d \geq 1$, and M is a Riemannian manifold. We will now use these geodesic finite elements to solve partial differential equations for functions in $H^1(\Omega, M)$. We restrict our attention to time-independent PDEs that have a minimization formulation. That is, we assume that there is an energy functional

$$\mathcal{J} : H^1(\Omega, M) \to \mathbb{R}$$

such that the (stable) problem solutions are minima of $\mathcal{J}$ subject to suitable boundary conditions. For simplicity we consider Dirichlet boundary conditions only. More formally, our continuous problem is then to find a function $u \in H^1(\Omega, M)$ such that

$$\mathcal{J}(u) \leq \mathcal{J}(v) \tag{9}$$

for all v in a neighborhood of u in $H^1(\Omega, M)$, and

$$u = u_D \qquad \text{on } \partial\Omega, \tag{10}$$

with $u_D : \partial\Omega \to M$ sufficiently smooth. We leave out any discussion of the well-posedness of this kind of problems, as that would go beyond the scope of this article.

4.1 Discrete and Algebraic Problem Formulations

We state the discrete problem corresponding to (9)–(10). Remember that Ω is assumed to be polygonally bounded and let $\mathcal{G}$ be a conforming simplicial grid of Ω. We have shown in Theorem 3.1 that the geodesic finite element space V_h^M is a subspace of $H^1(\Omega, M)$. Consequently, the energy functional $\mathcal{J}$ is well-defined on V_h^M. We can formulate a discrete version of Problem (9)–(10) by restricting the ansatz space to V_h^M: Find a function $u_h \in V_h^M(\Omega)$ such that

$$\mathcal{J}(u_h) \leq \mathcal{J}(v_h) \tag{11}$$

for all v_h in a neighborhood $U \in V_h^M$ of u_h, and with

$$u_h = u_{h,D} \qquad \text{on } \partial\Omega,$$

where $u_{h,D}$ is a suitable approximation of the Dirichlet data u_D.

For a numerical treatment of (11) we need the corresponding algebraic formulation. Let u be a solution of the continuous problem (9) and assume that it is Lipschitz continuous. By Theorem 3.2 there is a number $h_0 > 0$ and a neighborhood V of $\mathcal{E}(u)$ in M^n such that for all coefficient sets $\bar{v} \in V$ there is a unique discrete function $\mathcal{E}^{-1}(\bar{v}) \in V_h^M$ if the maximum grid edge length is less than h_0. We assume that $\mathcal{E}(u_h) \in V$, which allows us to formulate the algebraic minimization problem corresponding to (11): Find $\bar{u} \in M^n$ such that

$$\mathcal{J}(\mathcal{E}^{-1}(\bar{u})) \leq \mathcal{J}(\mathcal{E}^{-1}(\bar{v})) \tag{12}$$

for all $\bar{v} \in M^n$ such that $\bar{v} \in \mathcal{E}(U) \cap V$, and subject to the boundary conditions

$$\bar{u}_i = (\mathcal{E}(u_{h,D}))_i \qquad \text{for all vertices } x_i \text{ on } \partial\Omega.$$

For simplicity of notation we define the algebraic energy functional

$$J : M^n \to \mathbb{R}, \qquad J(\bar{v}) := \mathcal{J}(\mathcal{E}^{-1}\bar{v}).$$

With this functional we can rewrite the algebraic problem (12) as: Find $\bar{u} \in M^n$ such that

$$J(\bar{u}) \leq J(\bar{v}) \tag{13}$$

for all $\bar{v} \in M^n$ such that $\bar{v} \in \mathcal{E}(U) \cap V$, and subject to the boundary conditions

$$\bar{u}_i = (\mathcal{E}(u_{h,D}))_i$$

for all vertices x_i on $\partial\Omega$.

4.2 Riemannian Trust-Region Methods

The algebraic problem (13) is a minimization problem on the manifold M^n. For such problems Absil et al. [2] introduced the Riemannian trust-region method. We briefly describe their method and then present an extension that allows to solve the quadratic subproblems that occur at each iteration very efficiently in a finite element context.

Let N be a Riemannian manifold with metric g, and $j : N \to \mathbb{R}$ twice differentiable. The basic idea of the Riemannian trust-region algorithm is that in a neighborhood of a point $q \in N$ the objective function can be lifted onto the tangent space $T_q N$. There, a vector space trust-region subproblem can be solved and the result transported back onto N.

More formally, let $\nu \in \mathbb{N}$ be an iteration number and let $q_\nu \in N$ be the current iterate. We obtain the lifted functional by setting

$$\begin{aligned} \hat{j}_\nu \ &: \ T_{q_\nu} N \to \mathbb{R} \\ \hat{j}_\nu(s) &= j(\exp_{q_\nu} s). \end{aligned}$$

Let $\rho_\nu > 0$ be the current trust-region radius. The Riemannian metric g turns $T_{q_\nu} N$ into a Banach space with the norm $\|\cdot\|_{q_\nu} = \sqrt{g_{q_\nu}(\cdot,\cdot)}$. There, the trust-region subproblem reads

$$s_\nu = \operatorname*{arg\,min}_{s \in T_{q_\nu} N} m_\nu(s), \qquad \|s\|_{q_\nu} \leq \rho_\nu, \tag{14}$$

with the quadratic, but not necessarily convex model

$$m_\nu(s) = \hat{j}_\nu(0) + g_{q_\nu}(\nabla \hat{j}_\nu(0), s) + \frac{1}{2} g_{q_\nu}(\operatorname{Hess} \hat{j}_\nu(0)s, s). \tag{15}$$

Here $\nabla \hat{j}_\nu$ is the gradient and $\operatorname{Hess} \hat{j}_\nu$ the Hessian of $\hat{j}_\nu$, and both are evaluated at $0 \in T_{q_\nu} N$. Note that (15) is independent of a specific coordinate system on $T_{q_\nu} N$. As a minimization problem of a continuous function on a compact set, (14) has at least one solution s_ν, which generates the new iterate by

$$q_{\nu+1} = \exp_{q_\nu} s_\nu.$$

As in trust-region methods in linear spaces, the quality of a correction step s_ν is estimated by comparing the functional decrease and the model decrease. If the quotient

$$\kappa_\nu = \frac{j(q_\nu) - j(\exp_{q_\nu} s_\nu)}{m_\nu(0) - m_\nu(s_\nu)} \tag{16}$$

is smaller than a fixed value η_1, then the step is rejected, and s_ν is recomputed for a smaller trust-region radius ρ. Otherwise the step is accepted. If κ_ν is larger than a second value η_2, the trust-region radius is enlarged for the next step. Common values are $\eta_1 = 0.01$ and $\eta_2 = 0.9$ [11].

For this method, Absil et al. proved global convergence to first-order stationary points, and, depending on the exactness of the inner solver, locally superlinear or even locally quadratic convergence. The interested reader should consult the monograph [2] for details.

4.3 Using Monotone Multigrid for the Quadratic Problems

Various solvers have been proposed for the inner quadratic problems (14). Absil et al. used the Steihaug–Toint algorithm [2], and their local convergence result relies on properties of this solver. As an alternative for finite element problems, we propose a multigrid method which can solve the subproblems (14) with linear complexity.

We focus again on the finite element case, which means that N is a product manifold M^n, with n the number of grid vertices. Our construction is based on the fact that the trust-region convergence theory allows norms other than $\|\cdot\|_{q_\nu}$ for the definition of the trust-region [11]. We therefore introduce a maximum norm

$$\|\cdot\|_{\infty,TM^n} : TM^n \to \mathbb{R}_0^+$$

on the tangent bundle of M^n by defining for each $q = (q^1, \dots, q^n) \in M^n$, $s = (s^1, \dots, s^n) \in T_q M^n$

$$\|s\|_{\infty,TM^n} = \max_{i=1,\dots,n} \|s^i\|_{\infty,T_{q^i}M}. \tag{17}$$

The maximum norms $\|\cdot\|_{\infty,T_{q^i}M} : T_{q^i}M \to \mathbb{R}_0^+$ on the tangent spaces of M are to be understood with respect to fixed but arbitrary bases. Together with this choice of bases, the norm $\|\cdot\|_{\infty,TM^n}$ turns M^n into a Finsler manifold.

Using $\|\cdot\|_{\infty,TM^n}$ for the definition of the trust region, the quadratic subproblem (14) reads

$$s_\nu = \operatorname*{arg\,min}_{s \in T_{q_\nu}M^n} m_\nu(s), \qquad \|s\|_{\infty,T_{q_\nu}M^n} \le \rho_\nu, \tag{18}$$

with m_ν given again by (15). The key observation now is that the new trust region

$$K^{\mathrm{tr}}_{\infty,\nu} := \big\{ s \in T_{q_\nu}M^n \;\big|\; \|s\|_{\infty,TM^n} \le \rho_\nu \big\}$$

has a tensor product structure

$$K^{\mathrm{tr}}_{\infty,\nu} = [-\rho_\nu, \rho_\nu]^{n \dim M} \subset \mathbb{R}^{n \dim M}.$$

Hence Problem (18) is a minimization problem for a quadratic functional on a compact hypercube. Since the quadratic part $\frac{1}{2}g_{q_\nu}(\text{Hess}\,\hat{j}_\nu(0)\cdot,\cdot)$ of m_ν stems from the discretization of a finite element problem, it is a sparse matrix.

For strictly convex quadratic minimization problems on sets with tensor product structure, Kornhuber introduced the monotone multigrid method (MMG) [17]. We will use it to solve the quadratic problems (18). As it turns out MMG performs well even for nonconvex energies, if the admissible set is compact.

The monotone multigrid method is a generalization of the linear multigrid method. Like the latter it assumes a nested hierarchy of grids $\mathcal{G}_0, \dots, \mathcal{G}_L$ and grid-dependent prolongation operators $P_j : \mathbb{R}^{n_{j-1}} \to \mathbb{R}^{n_j}$, $j = 1, \dots, L$, where n_j is the number of vertices on the j-th level. In the easiest case the operators P_j are the ones known from linear multigrid theory. The linear smoothers are replaced by a projected Gauß–Seidel method. On the coarser grid levels, coarse grid approximations to the admissible defect set are constructed such that an admissible coarse grid correction does not lead to an inadmissible fine grid iterate. The following is a simple pseudocode implementation of a single multigrid correction step for the functional

$$J_{\text{alg}}(x) = \frac{1}{2}x^T Ax - bx \qquad \text{on } K = \prod_{i=1}^{n_L}[\mu_i^-, \mu_i^+],$$

where $A \in \mathbb{R}^{n_L \times n_L}$ is symmetric, $b \in \mathbb{R}^{n_L}$, $\mu_i^-, \mu_i^+ \in \mathbb{R}$ for all $1 \le i \le n_L$, and $K \neq \emptyset$. The e_i occurring in the code are the canonical basis vectors of $\mathbb{R}^{n_j}$.

```
// MMG iteration on level j
mmg_step(A_j, r_j, λ_j^-, λ_j^+, j)

    // nonlinear Gauss-Seidel smoother
    define J_j(w) = 1/2 w^T A_j w - r_j w
    for each i = 1,...,n_j do

        set (v_j)_i = argmin_{v ∈ [(λ_j^-)_i, (λ_j^+)_i]} J_j( Σ_{k=1}^{i-1} (v_j)_k e_k + v e_i )

    if (j>0)
        // construct coarse grid defect problem
        set A_{j-1} = P_j^T A_j P_j
        set r_{j-1} = P_j^T (r_j - A_j v_j)

        // compute monotone coarse defect obstacles
        compute λ_{j-1}^- such that v_j + P_j λ_{j-1}^- ≥ λ_j^-
        compute λ_{j-1}^+ such that v_j + P_j λ_{j-1}^+ ≤ λ_j^+

        // recurse
        set v_j = v_j + P_j mmg_step(A_{j-1}, r_{j-1}, λ_{j-1}^-, λ_{j-1}^+, j-1)

    return v_j

end
```

As the pseudocode implementation computes corrections it has to be called as

```
    set x^{k+1} = x^k + mmg_step(A, b - Ax^k, μ^- - x^k, μ^+ - x^k, L).
```

For details we refer the reader to the literature (e.g., [15, 17]).

In practice, the most important situation is a strictly convex quadratic functional m_ν. In this case, monotone multigrid converges globally with asymptotic multigrid speed.

Theorem 4.1 (Kornhuber [17]). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite, $b \in \mathbb{R}^n$, and let $K = \prod_{i=1}^n [\mu_i^-, \mu_i^+]$, with $\mu_i^-, \mu_i^+ \in \mathbb{R}$, $K \neq \emptyset$.*

1. *For any initial iterate $x^0 \in \mathbb{R}^n$, the MMG method converges to the unique minimizer of the functional $J_{alg}(x) = \frac{1}{2}x^T Ax - bx$ on K.*
2. *If strict complementarity holds (see, e.g., [11]), there exists a number $k_0 \in \mathbb{N}$ such that for all iterations $k > k_0$ the active set*

$$\mathcal{N}_k^\bullet = \{i \in 1,\dots,n \mid x_i^k = \mu_i^- \text{ or } x_i^k = \mu_i^+\}$$

remains invariant and MMG degenerates to a standard multigrid method.

If the functional J_{alg} is not strictly convex the method remains well defined. Scalar minimization in directions in which the functional is concave still has minimizers, because the admissible set K is compact. If there are two minimizers for a search direction e_i either one can be chosen by the algorithm. The proof of global convergence under mild additional assumptions will appear in a separate article. In Chapter 6 we will demonstrate the fast convergence numerically.

5 Numerical Aspects of Geodesic Finite Elements

Geodesic finite elements are more difficult to handle than regular finite elements, because they are only defined implicitly through minimization problems. Therefore, in this section we revisit geodesic finite element functions from an algorithmic point of view.

So far we have only considered PDE problems with a minimization formulation. We now restrict our attention even further to problems with functionals of the form

$$\mathcal{J}(v) = \int_\Omega W(\nabla v(x), v(x), x)\, dx, \qquad v \in H^1(\Omega, M),$$

where W is a scalar energy density assumed to be as smooth as necessary. For a given grid $\mathcal{G}$ with n vertices the corresponding algebraic energy is

$$J(\bar v) = \int_\Omega W\big(\nabla(\mathcal{E}^{-1}(\bar v))(x), (\mathcal{E}^{-1}(\bar v))(x), x\big)\, dx, \qquad \bar v \in M^n. \tag{19}$$

To compute this energy numerically for a given function $v_h = \mathcal{E}^{-1}(\bar v) \in V_h^M$ (needed in (16) to assess the quality of trust-region steps) we have to evaluate geodesic finite element function values $v_h(x)$ and derivatives $\nabla v_h(x)$ at (quadrature) points x in Ω. For the minimization by Riemannian trust-region methods we further need to evaluate first and second derivatives of J with respect to the coefficients $\bar v$. By the chain rule, this in turn requires derivatives of $v_h(x)$ and $\nabla v_h(x)$ with respect to the finite element coefficients. In this work we compute analytical expressions for the first derivatives. Doing the same for second derivatives is possible but the expressions get fairly large. In our numerical example we have approximated the Hessian of J by finite differences.

Remark 5.1. So far geodesic finite element functions have been defined in terms of coordinates on the standard simplex Δ. However, in the finite element literature the reference simplex $\widetilde{\Delta}$ is much more common. For example, quadrature rules in existing codes are usually given with respect to $\widetilde{\Delta}$. Since the formulas we present in this chapter are meant to be implemented in actual computer codes, we formulate them in coordinates on the reference simplex. Therefore the mapping $\mathcal{B} : \widetilde{\Delta} \to \Delta$ given in (2) appears in the equations.

5.1 Evaluation of Geodesic Finite Element Functions

Let $v_h : \Omega \to M$ be a geodesic finite element function and $x \in \Omega$. We want to compute $v_h(x) \in M$. If T is an element of $\mathcal{G}$ with $x \in T$, then $v_h(x)$ can be computed by geodesic simplicial interpolation between the coefficients $v_{T,1}, \dots, v_{T,d+1} \in M$ corresponding to the corners of T. More formally, let $\widetilde{\mathcal{F}}_T : T \to \widetilde{\Delta}$ be an affine mapping from T onto the reference simplex. Then we have $v_h(x) = \tilde{v}_T(\xi)$ with $\xi = \widetilde{\mathcal{F}}_T(x)$ and $\tilde{v}_T$ a function defined on the reference simplex. By construction of v_h, the value $\tilde{v}_T(\xi)$ is given by

$$\tilde{v}_T(\xi) = \operatorname*{arg\,min}_{q \in M} \tilde{f}_\xi(q), \qquad \tilde{f}_\xi(\cdot) = \sum_{i=1}^{d+1} \mathcal{B}(\xi)_i \operatorname{dist}(v_{T,i}, \cdot)^2. \tag{20}$$

In general the minimization problem (20) can only be solved numerically. Since its objective function $\tilde{f}_\xi$ is defined on a Riemannian manifold M we use a Riemannian trust-region method as presented in Section 4.2. Under the assumptions of Theorem 2.1, $\tilde{f}_\xi$ is C^∞ (Lemma 2.4) and strictly convex on an open geodesic ball containing the $v_{T,i}$ [16, Thm. 1.2].

With k the trust-region iteration number let $q_k \in M$ be the current iterate. We use the exponential map $\exp_{q_k} : T_{q_k}M \to M$ to define lifted functionals

$$\begin{aligned} \hat{f}_k \ &: \ T_{q_k}M \to \mathbb{R} \\ \hat{f}_k(s) &= \sum_{i=1}^{d+1} \mathcal{B}(\xi)_i \operatorname{dist}(v_{T,i}, \exp_{q_k} s)^2, \end{aligned}$$

and corresponding quadratic models

$$m_k(s) = \hat{f}_k(0) + g_{q_k}(\nabla \hat{f}_k(0), s) + \frac{1}{2} g_{q_k}(\operatorname{Hess} \hat{f}_k(0)s, s).$$

Using $\nabla \exp 0 = \mathrm{Id}$ we see that the gradient of $\hat{f}_k$ at $0 \in T_{q_k}M$ is

$$\nabla \hat{f}_k(0) = \sum_{i=1}^{d+1} \mathcal{B}(\xi)_i \frac{\partial}{\partial q} \operatorname{dist}(v_{T,i}, q)^2,$$

and that the Hessian is

$$\operatorname{Hess} \hat{f}_k(0) = \sum_{i=1}^{d+1} \mathcal{B}(\xi)_i \frac{\partial^2}{\partial q^2} \operatorname{dist}(v_{T,i}, q)^2.$$

The derivatives of $\operatorname{dist}(\cdot,\cdot)^2$ embody the geometry of the manifold M. In the appendix we have computed the relevant formulas for M being a unit sphere.

At each step k of the Riemannian trust-region method a constrained quadratic problem of the form

$$s_k = \operatorname*{arg\,min}_{s \in T_{q_k} M} m_k(s) \qquad \|s\| \leq \rho^k \tag{21}$$

has to be solved. The number of variables of this problem is $\dim M$ if the algorithm uses local coordinates on M, or a larger number if M is parametrized by an embedding into a Euclidean space. In many practical cases the number of variables is small, but there is no guarantee for this. In any case, the number is independent of the number of vertices in the finite element grid. The problems (21) can be solved, for example, by a preconditioned truncated conjugate gradient method as described in [11].

5.2 Gradients of Geodesic Finite Element Functions

To evaluate the energy J in (19) we also need to be able to compute the derivatives $\nabla v_h(x)$ at quadrature points $x \in \Omega$. Let again $v_h : \Omega \to M$ be a geodesic finite element function and $x \in \Omega$. The derivative of v_h at x (if it exists) is a linear map

$$\nabla v_h(x) : T_x \Omega \to T_{v_h(x)} M.$$

Again it is sufficient to consider functions on the reference simplex only. Indeed, if $\mathcal{G}$ is a grid, T an element of $\mathcal{G}$ such that x is in the interior of T and $\widetilde{\mathcal{F}}_T$ an affine mapping from T onto the reference simplex $\widetilde{\Delta}$, then

$$\nabla v_h(x) = \frac{\partial}{\partial x} \tilde{v}_T(\widetilde{\mathcal{F}}_T(x)) = \frac{\partial}{\partial \xi} \tilde{v}_T(\xi) \cdot \frac{\partial}{\partial x} \widetilde{\mathcal{F}}_T(x),$$

with $\tilde{v}_T : \widetilde{\Delta} \to M$ as defined in (20). The problem hence reduces to computing the gradient of a function $\tilde{v}_T : \widetilde{\Delta} \to M$ defined by geodesic interpolation.

In the proof of Lemma 2.2 the implicit function theorem was used to show under what circumstances the derivative $\partial \tilde{v}_T / \partial \xi$ exists. Here we use it again for the actual computation. Let $v_{T,1}, \dots, v_{T,d+1} \in M$ be the coefficients of v_h at the corners of T. The definition of geodesic simplicial interpolation on $\widetilde{\Delta}$ reads

$$\tilde{v}_T(\xi) = \Upsilon(v_T, \mathcal{B}(\xi)) = \operatorname*{arg\,min}_{q \in M} f_{v_T, \mathcal{B}(\xi)}(q),$$

with $f_{v_T, \mathcal{B}(\xi)}$ as defined in (5). By Lemma 2.4 the functional on the right is smooth, and the minimizer can hence also be characterized by

$$F(v_T, \mathcal{B}(\xi), \Upsilon(v_T, \mathcal{B}(\xi))) = 0, \tag{22}$$

where

$$F : M^{d+1} \times \Delta \times M \to TM$$

$$F(v, w, q) = \frac{\partial}{\partial q} f_{v,w}(q) = \sum_{i=1}^{d+1} w_i \frac{\partial}{\partial q} \operatorname{dist}(v_i, q)^2. \tag{23}$$

Taking the total derivative of (22) with respect to ξ and using $\mathcal{B}(\xi) = B\xi + c$ we get

$$\frac{\mathrm{d}}{\mathrm{d}\xi} F(v_T, \mathcal{B}(\xi), \Upsilon(v_T, \mathcal{B}(\xi)))$$
$$= \frac{\partial F(v_T, w, q)}{\partial w} \cdot B + \frac{\partial F(v_T, w, q)}{\partial q} \cdot \frac{\partial \Upsilon(v_T, \mathcal{B}(\xi))}{\partial \xi} = 0.$$

By Lemma 2.1 the matrix

$$\frac{\partial F}{\partial q} = \operatorname{Hess} f_{v_T, w} \tag{24}$$

is invertible, and hence $\partial \tilde{v}_T(\xi)/\partial\xi = \partial\Upsilon(v_T, \mathcal{B}(\xi))/\partial\xi$ can be computed as the solution of the linear system of equations

$$\frac{\partial F(v_T, w, q)}{\partial q} \cdot \frac{\partial \Upsilon(v_T, \mathcal{B}(\xi))}{\partial \xi} = -\frac{\partial F(v_T, w, q)}{\partial w} \cdot B. \tag{25}$$

Using the definition (23) we see that in coordinates $\partial F/\partial w$ is a $(\dim M)\times(d+1)$-matrix, where the i-th column is

$$\left(\frac{\partial F}{\partial w}^T\right)_i = \frac{\partial}{\partial q} \operatorname{dist}(v_{T,i}, q)^2.$$

Hence evaluating the derivative of a geodesic finite element function amounts to an evaluation of its value (to know where to evaluate the derivatives of F) and the solution of the symmetric linear system (25).

5.3 Derivatives of Finite Element Function Values with Respect to Coefficients

In order to find minima of the algebraic energy functional J by a Riemannian trust-region method we need its derivatives $\partial J/\partial \bar{v}_i$, $1 \le i \le n$. By the chain rule, the expression for these derivatives includes derivatives $\partial v_h(x)/\partial \bar{v}_i$, where $v_h \in V_h^M$ and x is a (quadrature) point in Ω. By the same reasoning as in Section 5.1 it follows that the $\partial v_h(x)/\partial \bar{v}_i$ can be obtained by computing the corresponding derivatives of geodesic simplicial interpolation functions on the reference simplex.

Let therefore $\tilde{v}_T : \widetilde{\Delta} \to M$ be a function given by geodesic simplicial interpolation, $v_i \in M$, $i = 1, \dots, d+1$ its corner coefficients, and let $\xi \in \widetilde{\Delta}$ be arbitrary but fixed coordinates. We want to compute the derivatives

$$\frac{\partial}{\partial v_i} \tilde{v}_T(\xi) = \frac{\partial}{\partial v_i} \Upsilon(v_1, \dots, v_{d+1}; \mathcal{B}(\xi)) \;:\; T_{v_i} M \to T_{\Upsilon(v, \mathcal{B}(\xi))} M$$

for all $i = 1, \dots, d+1$. By definition of Υ we have

$$F(v_1, \dots, v_{d+1}, \mathcal{B}(\xi), \Upsilon(v_1, \dots, v_{d+1}; \mathcal{B}(\xi))) = 0.$$

Taking the total derivative of this with respect to v_i gives

$$\frac{\mathrm{d}F}{\mathrm{d}v_i} = \frac{\partial F}{\partial v_i} + \frac{\partial F}{\partial q} \cdot \frac{\partial \Upsilon(v, \mathcal{B}(\xi))}{\partial v_i} = 0,$$

with

$$\frac{\partial F}{\partial v_i} = \mathcal{B}(\xi)_i \frac{\partial}{\partial v_i}\frac{\partial}{\partial q}\operatorname{dist}(v_i, q)^2$$

and $\partial F/\partial q$ as in (24). By Lemma 2.1 the matrix $\partial F/\partial q$ is invertible. Hence the derivative of $\Upsilon(v_1, \dots, v_{d+1}; \mathcal{B}(\xi))$ with respect to one of its coefficients v_i can be computed as a minimization problem to obtain the value $\Upsilon(v, \mathcal{B}(\xi))$ and the solution of the linear system of equations

$$\frac{\partial F}{\partial q} \cdot \frac{\partial}{\partial v_i}\Upsilon(v, \mathcal{B}(\xi)) = -\frac{\partial F}{\partial v_i}. \tag{26}$$

5.4 Derivatives of the Finite Element Gradient with Respect to Coefficients

Let $v_h : \Omega \to M$ be a geodesic finite element function and let $x \in \Omega$ be such that $\nabla v_h(x)$ exists. We assume that the conditions of Theorem 3.2 hold and hence for a fixed x we can view $\nabla v_h(x)$ as a function of the coefficient vector $\bar{v} = \mathcal{E}(v_h) \in M^n$. For each $i = 1, \dots, n$ we want to compute the derivative $\frac{\partial}{\partial \bar{v}_i}\nabla v_h(x)$. Note that since $\nabla v_h(x)$ is a linear map, its derivative is a trilinear form—written in coordinates it will have three indices.

Let T be an element of $\mathcal{G}$ such that $x \in T$, and $v_{T,i}$, $1 \le i \le d+1$ the coefficients at the corners of T. As in Section 5.2 we use the decomposition

$$\nabla v_h(x) = \frac{\partial}{\partial \xi}\tilde{v}_T(\xi) \cdot \frac{\partial}{\partial x}\widetilde{\mathcal{F}}_T(x),$$

where $\widetilde{\mathcal{F}}_T$ is an affine mapping from T onto the reference simplex $\widetilde{\Delta}$, and $\tilde{v}_T = \Upsilon(v_T, \mathcal{B}(\cdot))$ is defined on $\widetilde{\Delta}$. We can then write

$$\frac{\partial}{\partial \bar{v}_i}\nabla v_h(x) = \frac{\partial}{\partial \bar{v}_i}\Big(\frac{\partial}{\partial \xi}\tilde{v}_T(\xi) \cdot \frac{\partial}{\partial x}\widetilde{\mathcal{F}}_T(x)\Big) = \bigg(\frac{\partial}{\partial v_{T,*}}\frac{\partial}{\partial \xi}\tilde{v}_T(\xi)\bigg) \cdot \frac{\partial}{\partial x}\widetilde{\mathcal{F}}_T(x),$$

where $v_{T,*}$ is the coefficient of $\tilde{v}_T$ corresponding to the global coefficient $\bar{v}_i$. It is hence again sufficient to consider only functions on the reference simplex.

To compute the derivative of $\nabla \tilde{v}_T$ with respect to $v_{T,*}$ we take the total derivative of expression (25) with respect to $v_{T,*}$ and obtain

$$\begin{aligned} \frac{\partial F}{\partial q} \cdot \frac{\partial^2 \Upsilon}{\partial v_{T,*}\,\partial \xi} = -\frac{\partial^2 F}{\partial v_{T,*}\,\partial q} \cdot \frac{\partial \Upsilon}{\partial \xi} - \frac{\partial \Upsilon}{\partial v_{T,*}} \cdot \frac{\partial^2 F}{\partial q^2} \cdot \frac{\partial \Upsilon}{\partial \xi} \\ - \frac{\partial^2 F}{\partial v_{T,*}\,\partial w} \cdot B - \frac{\partial \Upsilon}{\partial v_{T,*}} \cdot \frac{\partial^2 F}{\partial q\,\partial w} \cdot B. \end{aligned} \tag{27}$$

In coordinates (and using the Einstein convention), this is

$$\begin{aligned} \frac{\partial F_l}{\partial q_j}\frac{\partial^2 \Upsilon_l}{\partial (v_{T,*})_i\,\partial \xi_k} = -\frac{\partial^2 F_l}{\partial (v_{T,*})_i\,\partial q_j}\frac{\partial \Upsilon_l}{\partial \xi_k} - \frac{\partial \Upsilon_l}{\partial (v_{T,*})_i}\Big[\frac{\partial^2 F}{\partial q^2}\Big]_{ljm}\frac{\partial \Upsilon_m}{\partial \xi_k} \\ - \frac{\partial^2 F_j}{\partial (v_{T,*})_i\,\partial w_l}B_{lk} - \frac{\partial \Upsilon_m}{\partial (v_{T,*})_i}\frac{\partial^2 F_j}{\partial q_m\,\partial w_l}B_{lk}. \end{aligned}$$

This time we need to compute the value of $\tilde{v}_T(\xi)$, solve a linear system for $\partial\Upsilon/\partial\xi$, and then solve the linear system (27) to obtain the desired value of $\partial^2\Upsilon/\partial V_{T,*}\partial\xi$. Various new derivatives of F appear. These are

$$\begin{aligned}
\frac{\partial^2 F_j}{\partial(v_{T,*})_i\,\partial w_k} &= \delta_{*k}\frac{\partial}{\partial(v_{T,*})_i}\frac{\partial}{\partial q_j}\operatorname{dist}(v_{T,*},q)^2 \\
\frac{\partial^2 F_k}{\partial(v_{T,*})_i\,\partial q_j} &= w_*\frac{\partial}{\partial(v_{T,*})_i}\Big[\frac{\partial^2}{\partial q^2}\operatorname{dist}(v_{T,*},q)^2\Big]_{jk} \\
\Big[\frac{\partial^2 F}{\partial q^2}\Big]_{ijk} &= \sum_{l=1}^{d+1} w_l\Big[\frac{\partial^3}{\partial q^3}\operatorname{dist}(v_{T,l},q)^2\Big]_{ijk} \\
\frac{\partial^2 F_j}{\partial q_i\,\partial w_k} &= \Big[\frac{\partial^2}{\partial q^2}\operatorname{dist}(v_{T,k},q)^2\Big]_{ij}.
\end{aligned}$$

Note that these are all third-order objects.

5.5 Summary: Required Information about M

In the preceding sections we have computed values of geodesic finite element functions, and the various derivatives that were necessary to implement geodesic finite element methods. For computations of values a Riemannian trust-region method was used, and derivatives were obtained by different applications of the implicit function theorem. We have seen that all information about the manifold M enters in form of different derivatives of the squared distance function $\operatorname{dist}(\cdot,\cdot)^2 : M\times M \to \mathbb{R}$. In summary, the derivatives used were

$$\frac{\partial}{\partial q}\operatorname{dist}(p,q)^2, \qquad \frac{\partial^2}{\partial p\partial q}\operatorname{dist}(p,q)^2, \qquad \frac{\partial^2}{\partial q^2}\operatorname{dist}(p,q)^2,$$
$$\frac{\partial^3}{\partial p\partial q^2}\operatorname{dist}(p,q)^2, \qquad \frac{\partial^3}{\partial q^3}\operatorname{dist}(p,q)^2.$$

Additionally, the exponential map on M was needed for the trust-region method. If the Hessian of the energy functional $\mathcal{J}$ were also to be computed analytically, then various forth-order derivatives would have to be added to the list.

For various important spaces such as SO(3) and the unit spheres S^m, $m\in\mathbb{N}$, the derivatives can be computed analytically. Formulas for the latter can be found in the appendix. If no closed-form expressions are available the derivatives can in principle be approximated numerically.

6 Example: Harmonic Maps in Liquid Crystals

We close this article by giving a numerical example. Our aim is two-fold. First, we want to show that the Riemannian trust-region method of Section 4.2 together with the monotone multigrid solver is able to efficiently solve geodesic finite element problems. Secondly, we want to numerically estimate the convergence order of the discretization by geodesic finite elements.

Let N and M be two Riemannian manifolds. A function $v : N \to M$ is

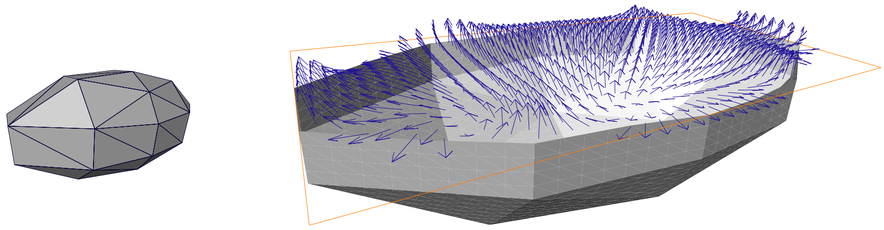

Figure 2: Left: Coarse grid. Right: Cut through the solution field on a grid obtained by four steps of uniform refinement.

called a harmonic map if it is a stationary point of the energy functional

$$\begin{aligned}\mathcal{J} \;&:\; H^1(N,M) \to \mathbb{R} \\ \mathcal{J}(v) &= \int_N |\nabla v|^2\,dx,\end{aligned} \tag{28}$$

where $\nabla : TN \to TM$ is the differential of v and the norm $|\cdot|$ is the one corresponding to the induced metric on the bundle $TN \otimes v^{-1}TM$. We consider the special case $N = \Omega \subset \mathbb{R}^d$ open and bounded and $M = S^2$ the unit sphere in $\mathbb{R}^3$, and try to solve the associated Dirichlet problem

$$\text{minimize} \quad \mathcal{J}(v) = \int_\Omega |\nabla v|^2\,dx \qquad \text{in } H^1(\Omega, S^2)$$

subject to

$$v = v_D \qquad \text{on } \partial\Omega,$$

with v_D a given set of boundary conditions. This is a standard model for equilibrium states of nematic liquid crystals, known as the one-constant approximation [13].

As coordinates on the unit sphere S^2 we use the canonical embedding into $\mathbb{R}^3$ (see the appendix for details). With the metric on S^2 induced by the embedding we obtain the coordinate representation

$$|\nabla v|^2 = \sum_{i=1}^{d}\sum_{\alpha=1}^{3}\Big(\frac{\partial v^\alpha}{\partial x^i}\Big)^2,$$

that is, ∇v is a $3\times d$-matrix and $|\cdot|$ the Frobenius norm.

As the domain Ω we choose an approximation of the ellipsoid

$$\Omega = \big\{x \in \mathbb{R}^3 \mid \frac{1}{4}x_1^2 + x_2^2 + x_3^2 \le 1\big\},$$

which we discretize by an unstructured simplicial grid $\mathcal{G}_0$ consisting of 40 vertices and 95 tetrahedra (Figure 2, left). The choice of an unstructured grid is

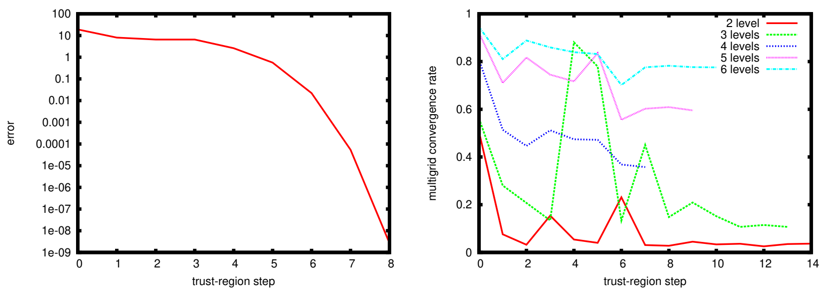

Figure 3: Left: Error after each trust-region step on a grid obtained by four steps of uniform refinement. Right: Average convergence rate of the monotone multigrid solver for each grid size and trust-region step.

deliberate, since we want to stress that geodesic finite elements do not rely on geometric structure in the grid. To create an example field of Dirichlet data v_D we first define a rotation axis $a(x) = \frac{(x_1,x_2,1)}{|(x_1,x_2,1)|}$ and then set

$$v_D(x) = R(a(x), -3\pi|x|)e_3,$$

where $R(a, \alpha)$ is the matrix corresponding to a rotation of an angle α around the axis a, and e_3 is the third canonical basis vector. The gradient of the energy functional (28) in the geodesic finite element space $V_h^{S^2}$ was computed analytically, whereas its Hessian was approximated by a finite difference method. The discretization and solution algorithms were implemented in C++ using the DUNE libraries [4].

We begin by measuring the efficiency of the trust-region solver. For this we use a grid $\mathcal{G}_4$ of $n_4 = 69\,304$ vertices, obtained from $\mathcal{G}_0$ by four steps of uniform refinement. We compute a reference solution $v^* \in (S^2)^{n_4}$ by solving the problem up to machine precision (Figure 2, right). Then, starting from

$$v^0 \in (S^2)^{n_4}, \qquad v_i^0 = \begin{cases} v_D(x_i) & \text{if } x_i \in \partial\Omega, \\ (1,0,0)^T & \text{else,} \end{cases}$$

we iterate again until the maximum norm of the correction steps drops below 10^{-9}. After each iteration ν, the error of the current iterate $v^\nu \in (S^2)^{n_4}$ with respect to the reference solution is computed as

$$e^\nu = \|\mathcal{E}^{-1}(v^\nu) - \mathcal{E}^{-1}(v^*)\|_1,$$

where $\|\cdot\|_1$ is the standard H^1-norm on the space $H^1(\Omega, \mathbb{R}^3)$. Figure 3, left, shows the error e^ν per iteration step ν. One can clearly see the local superlinear behavior predicted by the trust-region convergence theory.

Our next interest is the behavior of the solver as the grid is refined. We therefore repeat the previous experiment on a family of grids $\mathcal{G}_0, \dots, \mathcal{G}_5$ obtained from $\mathcal{G}_0$ by up to five steps of uniform refinement. On each grid $\mathcal{G}_j$ we compute a reference solution v_j^* and measure the solver convergence as before. Table 1 shows the number of trust-region iterations needed on each grid. We see that the number appears to be bounded from above independently from the grid

grid	$\mathcal{G}_0$	$\mathcal{G}_1$	$\mathcal{G}_2$	$\mathcal{G}_3$	$\mathcal{G}_4$	$\mathcal{G}_5$
vertices	40	208	1 311	9 245	69 305	536 433
overall it.	24	15	14	8	10	11
unsuccessful it.	20	1	1	0	1	1

Table 1: Number of iterations of the trust-region solver per grid size.

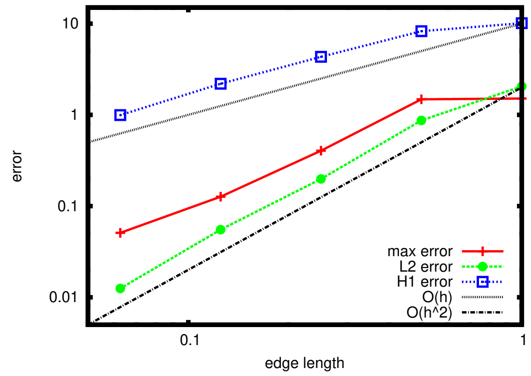

Figure 4: Discretization error as a function of normalized grid edge length.

size. This appears plausible considering the investigations in [27] for the linear case. It makes the method attractive, as using a finer grid to obtain a higher approximation quality does not automatically result in larger iteration numbers. The reason for the large number of unsuccessful iterations on the coarsest grid is unclear.

Figure 3, right, shows the convergence rates of the monotone multigrid solver. For each solve of the quadratic inner problem (18) we computed an average convergence rate. These rates are plotted for each grid and each trust-region step ν, including the unsuccessful ones. It can be seen from the figure that the rates stay well below 1, ever for the larger grids. From general multigrid theory we expect that there is an upper bound strictly lower than 1 even for $h \to 0$. The data we have assembled suggests this but is not comprehensive enough for a definite conclusion.

After having investigated the solver we now turn to the discretization error of geodesic finite elements. From the study on the solver speed we take the reference solutions v_j^*, $j = 0, \dots, 5$ computed on the grids $\mathcal{G}_0, \dots, \mathcal{G}_5$, respectively. We pick v_5^* to be the reference solution for the discretization error and compute the errors

$$e_V^k = \|\mathcal{E}^{-1}(v_k^*) - \mathcal{E}^{-1}(v_5^*)\|_V, \qquad k = 0, \dots, 4,$$

where V is one of $L^2(\Omega, \mathbb{R}^3)$, $L^\infty(\Omega, \mathbb{R}^3)$, or $H^1(\Omega, \mathbb{R}^3)$. Note that since geodesic finite element functions are not piecewise polynomials in $\mathbb{R}^3$, the error norms can only be computed with an additional error due to numerical quadrature.

Figure 6 shows the errors as functions of the mesh size h. We see that the H^1-error and L^2-error are linear and quadratic functions of h, respectively.

Hence we can reproduce the optimal convergence behaviour well-known from the linear theory even in this nonlinear case.

A Appendix: The Unit Sphere S^m

The geometry of the manifold M appears in the implementation of a geodesic finite element method in form of various derivatives of the squared distance function $\operatorname{dist}(\cdot,\cdot)^2 : M \times M \to \mathbb{R}$. A list of these derivatives has been given in Section 5.5. In this appendix we give explicit formulas for the case that M is the unit sphere S^m. In our numerical examples only the two-dimensional sphere S^2 was used. However, the formulas can be stated for all $m \in \mathbb{N}$ without increase in complexity.

The unit sphere S^m is the smooth compact manifold

$$S^m = \{p \in \mathbb{R}^{m+1} \mid \|p\| = 1\},$$

and global coordinates on S^m are naturally given by this embedding into $\mathbb{R}^{m+1}$. The tangent space at a point $p \in S^m$ is

$$T_pS^m = \{v \in \mathbb{R}^{m+1} \mid \langle p, v\rangle_{\mathbb{R}^{m+1}} = 0\}.$$

Hence tangent vectors $v \in T_pS^m$ can be treated as vectors in $\mathbb{R}^{m+1}$. S^m becomes a Riemannian manifold by inheriting the metric of the surrounding space

$$\langle v, w\rangle_{T_pS^m} := \langle v, w\rangle_{\mathbb{R}^{m+1}} \qquad \text{for all } v, w \in T_pS^m.$$

For a point $p \in S^m$ and a tangent vector $v \in T_pS^m$, the exponential map $\exp_p : T_pS^m \to S^m$ is then given by [2, Ex. 5.4.1]

$$\exp_p v = \cos|v| \cdot p + \frac{\sin|v|}{|v|} \cdot v.$$

The geodesics of S^m are the segments of great circles. Any two points $p, q \in S^m$ can be connected by such segments; hence S^m is geodesically complete. If $p \neq -q$ there is a unique shortest geodesic that connects p and q. For all pairs of points $p = -q$ there are infinitely many minimizing geodesics, each of length π. Hence the injectivity radius of S^m is $\operatorname{inj}(S^m) = \pi$. Any geodesic ball on S^m with radius $\rho < \pi/2$ is convex. The Riemannian distance between two points p and q is the length of the shortest arc of a great circle connecting p to q. It is given by

$$\operatorname{dist}(p, q) = \arccos\langle p, q\rangle. \tag{29}$$

To compute derivatives of $\operatorname{dist}(\cdot,\cdot)^2$ on S^m we use the following result [2, Prop. 5.3.2]. Part 1 covers the first derivatives, and Part 2 allows to cover the higher ones.

Lemma A.1. *Let M be a smooth Riemannian manifold isometrically embedded in a Euclidean space $\mathbb{R}^{m+1}$. For each $p \in M$ let $P_p : T_p\mathbb{R}^{m+1} \to T_pM$ be the orthogonal projection onto the tangent space at p.*

1. *Let $f : M \to \mathbb{R}$ be continuously differentiable and $\tilde{f}$ a smooth extension of f to a neighborhood of M in $\mathbb{R}^{m+1}$. Then*

$$\nabla f = P_p \widetilde{\nabla} \tilde{f}, \tag{30}$$

where ∇ is the gradient operator on M, and $\widetilde{\nabla}$ is the gradient in $\mathbb{R}^{m+1}$.

2. Let F, G be (tangential) vector fields on M, and $\tilde{F}$, $\tilde{G}$ smooth extensions of F and G to a neighborhood of M in $\mathbb{R}^{m+1}$. Then

$$\nabla_F G = P_p \widetilde{\nabla}_{\tilde{F}} \tilde{G},$$

where $\nabla_F G$ is the covariant derivative on M and $\widetilde{\nabla}_{\tilde{F}} \tilde{G}$ is the covariant derivative in $\mathbb{R}^{m+1}$.

As a suitable extension of the squared distance function we choose

$$\widetilde{\operatorname{dist}}(p,q)^2 = \operatorname{dist}\left(\frac{p}{|p|}, \frac{q}{|q|}\right)^2 = \arccos^2 \left\langle \frac{p}{|p|}, \frac{q}{|q|} \right\rangle.$$

This is well-defined and smooth on a neighborhood of S^m in $\mathbb{R}^{m+1}$. For any $p \in S^m$, the projection $P_p : T_p\mathbb{R}^{m+1} \to T_p S^m$ is given in coordinates by

$$(P_p)_{ij} = \delta_{ij} - p_i p_j.$$

For ease of notation we define $\alpha : [-1,1] \to \mathbb{R}$, $\alpha(x) := \arccos^2(x)$.

We now compute the derivatives of $\operatorname{dist}(\cdot,\cdot)^2$ listed in Section 5.5, beginning with the gradient

$$\nabla \operatorname{dist}(p,\cdot)^2 = \frac{\partial}{\partial q} \operatorname{dist}(p,q)^2 \in T_q S^m,$$

for arbitrary but fixed $p \in S^m$. Note that

$$\frac{\partial}{\partial q} \left\langle \frac{p}{|p|}, \frac{q}{|q|} \right\rangle = P_q p \qquad \text{if } |p| = |q| = 1.$$

With (29), (30), and $|p| = |q| = 1$ we get

$$\frac{\partial}{\partial q_i} \operatorname{dist}(p,q)^2 = \left(\frac{\partial}{\partial q} \operatorname{dist}(p,q)^2\right)_i = \alpha'(x)\Big|_{x=\langle p,q\rangle} (P_q p)_i.$$

The derivatives of $\alpha(x)$ will be discussed below. To simplify the notation further we will write α' to mean the value of the derivative of $\alpha(x)$ at $x = \langle p,q\rangle$, and similarly for higher derivatives.

For the second derivatives of $\operatorname{dist}(\cdot,\cdot)^2$ we note that

$$\frac{\partial}{\partial p_j}(P_q p)_i = (P_q)_{ij} = \delta_{ij} - q_i q_j \qquad \text{and} \qquad \frac{\partial}{\partial q_j}(P_q p)_i = -\delta_{ij}\langle p,q\rangle - q_i p_j.$$

This allows to compute

$$\left[\frac{\partial^2}{\partial q^2} \operatorname{dist}(p,q)^2\right]_{ij} = \alpha''(P_q p)_i (P_q p)_j - \alpha'(P_q)_{ij}\langle p,q\rangle.$$

Similarly we get the mixed derivative

$$\frac{\partial}{\partial p_i}\frac{\partial}{\partial q_j} \operatorname{dist}(p,q)^2 = \alpha''(P_p q)_i (P_q p)_j + \alpha'(P_p)_{ik}(P_q)_{kj},$$

where summation over pairs of equal indices is implied. Finally, the derivatives of third order are

$$\begin{aligned}\frac{\partial}{\partial p_i}\Big[\frac{\partial^2}{\partial q^2}\operatorname{dist}(p,q)^2\Big]_{jk} &= \alpha'''(P_p q)_i (P_q p)_j (P_q p)_k \\ &\quad + \alpha''(P_p)_{il}(P_q)_{jl}(P_q p)_k + \alpha''(P_p)_{il}(P_q p)_j (P_q)_{kl} \\ &\quad - \alpha''(P_p q)_i (P_q)_{jk}\langle p,q\rangle - \alpha'(P_p q)_i (P_q)_{jk}\end{aligned}$$

and

$$\begin{aligned}\Big[\frac{\partial^3}{\partial q^3}\operatorname{dist}(p,q)^2\Big]_{ijk} &= \alpha'''(P_q p)_i (P_q p)_j (P_q p)_k - \alpha''(P_q)_{ij}\langle p,q\rangle (P_q p)_k \\ &\quad + \alpha''(P_q p)_i q_j (P_q p)_k - \alpha''(P_q p)_j (P_q)_{ik}\langle p,q\rangle \\ &\quad + \alpha''(P_q p)_i (P_q p)_j q_k - \alpha''(P_q p)_i (P_q)_{jk}\langle p,q\rangle \\ &\quad + \alpha'(P_q)_{ij} q_k \langle p,q\rangle + \alpha'(P_q)_{ik} q_j \langle p,q\rangle - \alpha'(P_q p)_i (P_q)_{jk}.\end{aligned}$$

What is left are the derivatives of $\alpha(x) = \arccos^2(x)$. For each of them, a closed form expression is available. They are

$$\begin{aligned}\alpha'(x) &= -\frac{2\arccos(x)}{\sqrt{1-x^2}}, \\ \alpha''(x) &= \frac{2}{1-x^2} - \frac{2x\arccos(x)}{(1-x^2)^{3/2}}, \\ \alpha'''(x) &= \frac{6x}{(1-x^2)^2} - \frac{(4x^2-2)\arccos(x)}{(1-x^2)^{5/2}}.\end{aligned}$$

These expressions get numerically unstable around $x = 1$. There, the series expansions

$$\begin{aligned}\alpha'(x) &= -2 + \frac{2(x-1)}{3} + O((x-1)^2), \\ \alpha''(x) &= \frac{2}{3} - \frac{8}{15}(x-1) + O((x-1)^2), \\ \alpha'''(x) &= -\frac{8}{15} + \frac{24}{35}(x-1) + O((x-1)^2),\end{aligned}$$

have to be used instead. The series diverge for $x \to -1$, which is not surprising as $x = \langle p,q\rangle = -1$ corresponds to the case that q is in the cut locus of p.

Remark A.1. Treating tangent vectors as vectors in $\mathbb{R}^{m+1}$ with respect to the canonical basis there is simple and elegant, but it also has several disadvantages. First, the matrix representation of Hess $f_{v,w}$ (24) will have a nontrivial kernel, and the systems (25), (26), and (27) have to solved with a rank-aware algorithm. Also, the convergence result for the monotone multigrid method only holds if the maximum norm (17) on each tangent space of S^m is given with respect to a basis of the tangent space. For these reasons, we construct a basis for each tangent space of S^m, and use it for the representation of Hess $f_{v,w}$ and the trust-region corrections. Let $x \in \mathbb{R}^{m+1}$ be a point on S^m with $x_{m+1} \le 0$. Stereographic projection of x from the north pole onto the plane $x_{m+1} = 0$ is given by

$$\sigma(x) = (y_1,\dots,y_m) = \Big(\frac{x_1}{1-x_{m+1}},\dots,\frac{x_m}{1-x_{m+1}}\Big). \tag{31}$$

The derivative of the inverse of σ is given by

$$(\nabla(\sigma^{-1}))_{ij} = \frac{\partial(\sigma^{-1})_i}{\partial y_j} = \begin{cases} \frac{2\delta_{ij}(1+|y|^2)-4y_iy_j}{(1+|y|^2)^2} & \text{if } i \le m, \\ \frac{4y_j}{(1+|y|^2)^2} & \text{else.} \end{cases} \tag{32}$$

Let e_i, $i = 1, \dots, m$ be the canonical basis vectors of $\mathbb{R}^m$. Since the stereographic projection σ is conformal, the vectors $\hat{e}_i = \nabla(\sigma^{-1})e_i$, $i = 1, \dots, m$ form an orthogonal basis of T_xS^m. The basis remains orthogonal even if the factor $(1 + |y|^2)^{-2}$ is neglected. To avoid degeneracies near the north pole all points $x \in S^m$ with $x_{m+1} > 0$ have to be treated by stereographic projection from the south pole. This amounts to replacing the denominator in (31) by $1 + x_{m+1}$ and the second case of (32) by $-\frac{4y_j}{(1+|y|^2)^2}$.

References

[1] P.-A. Absil, C. G. Baker, and K. A. Gallivan. Trust-region methods on Riemannian manifolds. *Found. Comput. Math.*, 7(3):303–330, 2007.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.

[3] S. Bartels and A. Prohl. Constraint preserving implicit finite element discretization of harmonic map flow into spheres. *Math. Comp.*, 76(260): 1847–1859, 2007.

[4] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic interface for parallel and adaptive scientific computing. Part II: Implementation and tests in DUNE. *Computing*, 82(2–3):121–138, 2008.

[5] M. Berger. *A Panoramic View of Differential Geometry*. Springer, 2003.

[6] F. Bethuel. The approximation problem for Sobolev maps between two manifolds. *Acta Math.*, 167:153–206, 1991.

[7] O. Bottema. On the medians of a triangle in hyperbolic geometry. *Canad. J. Math.*, 10:502–506, 1958.

[8] D. Braess. *Finite Elemente*. Springer Verlag, 3rd edition, 2002.

[9] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature.* Springer Verlag, 1999.

[10] S. R. Buss and J. P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, 20:95–126, 2001.

[11] A. Conn, N. Gould, and P. Toint. *Trust-Region Methods*. SIAM, 2000.

[12] M. Crisfield and G. Jelenić. Objectivity of strain measures in the geometrically exact three-dimensional beam theory and its finite-element implementation. *Proc. R. Soc. Lond. A*, 455:1125–1147, 1999.

[13] P. de Gennes and J. Prost. *The Physics of Liquid Crystals*. Clarendon Press, 1993.

[14] K. Deckelnick, G. Dziuk, and C. M. Elliott. Computation of geometric partial differential equations and mean curvature flow. *Acta Numerica*, (14):139–232, 2005.

[15] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comp. Math.*, 27(1):1–44, 2009.

[16] H. Karcher. Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.*, 30:509–541, 1977.

[17] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. B.G. Teubner, 1997.

[18] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer Verlag, 1997.

[19] C. Löh and R. Sauer. Degree theorems and Lipschitz simplicial volume for non-positively curved manifolds of finite volume. Technical report, Universität Münster, 2007.

[20] N. Mermin. The topological theory of defects in ordered media. *Rev. Mod. Phys.*, 51(3):591–648, 1979.

[21] M. Moakher. Means and averaging in the group of rotations. *SIAM J. Matrix Anal. Appl.*, 24(1):1–16, 2002.

[22] W. Müller. *Numerische Analyse und Parallele Simulation von nichtlinearen Cosserat-Modellen*. PhD thesis, Karlsruher Institut für Technologie, 2009.

[23] P. Neff. *Geometrically exact Cosserat theory for bulk behaviour and thin structures. Modelling and mathematical analyis*. Habilitationsschrift, Technische Universität Darmstadt, 2003.

[24] O. Sander. Geodesic finite elements for Cosserat rods. *Int. J. Num. Meth. Eng.*, 82(13):1645–1670, 2010.

[25] J. Simo and L. Vu-Quoc. A three-dimensional finite-strain rod model. Part II: Computational aspects. *Comput. Methods Appl. Mech. Engrg.*, 58 (1):79–116, 1986.

[26] B. Tang, G. Sapiro, and V. Caselles. Diffusion of general data on non-flat manifolds via harmonic maps theory: The direction diffusion case. *Int. J. Comput. Vision*, 36(2):149–161, 2000.

[27] M. Weiser, A. Schiela, and P. Deuflhard. Asymptotic mesh independence of Newton's method revisited. *SIAM J. Numer. Anal.*, 42(5):1830–1845, 2005.