

## PROJECTION-BASED FINITE ELEMENTS FOR NONLINEAR FUNCTION SPACES\*

PHILIPP GROHS<sup>†</sup>, HANNE HARDERING<sup>‡</sup>, OLIVER SANDER<sup>‡</sup>, AND MARKUS  
SPRECHER<sup>§</sup>

**Abstract.** We introduce a novel type of approximation spaces for functions with values in a nonlinear manifold. The discrete functions are constructed by piecewise polynomial interpolation in a Euclidean embedding space, and then projected pointwise onto the manifold. We show optimal interpolation error bounds with respect to Lebesgue and Sobolev norms. Additionally, we show similar bounds for the test functions, i.e., variations of discrete functions. Combining these results with a nonlinear Céa lemma, we prove optimal  $L^2$  and  $H^1$  discretization error bounds for harmonic maps from a planar domain into a smooth manifold. All these error bounds are also verified numerically.

**Key words.** geometric finite elements, projection, nonlinear manifold, interpolation errors, discretization errors, harmonic maps

**AMS subject classifications.** 65N30, 65D05

**DOI.** 10.1137/18M1176798

We investigate the discrete approximation of functions from a Euclidean domain  $\Omega$  to a closed embedded submanifold  $M$  of  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ . Such functions are involved in a variety of partial differential equations (PDEs), from fields like liquid crystal physics [4] and micromagnetics [17]. In these applications, the manifold is  $M = S^2$ , the unit sphere in  $\mathbb{R}^3$ . In Cosserat-type material models [45, 38, 37] the manifold is  $M = \mathbb{R}^3 \times \text{SO}(3)$ , where  $\text{SO}(3)$  is the special orthogonal group. Further examples are the investigation of harmonic maps into manifolds [8], signal processing of manifold-valued signals [41], and the denoising of manifold-valued images [6].

Various approaches exist in the literature for numerically computing maps into nonlinear spaces. However, besides geodesic finite elements (see below), none of these methods uses explicitly defined conforming spaces of discrete functions. For harmonic maps into the unit sphere and related problems, [33, 4, 5, 31, 8, 7] constrained only the vertex values of linear finite element functions to be on the target manifold. Then, either finite differences or explicitly nonconforming finite element methods were used. A detailed convergence analysis is given in [7], which, however, only covers first-order functions. Alternative methods try to only discretize variations (which live in linear spaces) [47, 46]. However, such methods introduce spurious dependencies of the solution on aspects of the numerical solver [15]. Function spaces do not appear explicitly in these works.

A standard approach in Cosserat mechanics obtains space discretizations by interpolating on particular tangent spaces  $T_p M$  of  $M$  [37, 36]. Such methods need to include ad hoc switching between different tangent spaces to avoid instabilities when approaching the cut locus of  $p$ . While this approach could in principle be

\*Received by the editors March 22, 2018; accepted for publication (in revised form) December 18, 2018; published electronically February 19, 2019.

<http://www.siam.org/journals/sinum/57-1/M117679.html>

<sup>†</sup>Universität Wien, Fakultät für Mathematik, Oskar Morgenstern Platz 1, 1090 Wien, Austria (philipp.grohs@univie.ac.at).

<sup>‡</sup>Technische Universität Dresden, Institut für Numerische Mathematik, Zellescher Weg 12–14, 01069 Dresden, Germany (hanne.hardering@tu-dresden.de, oliver.sander@tu-dresden.de).

<sup>§</sup>ETH Zürich, Seminar für Angewandte Mathematik, Rämistrasse 101, 8092 Zürich, Switzerland (markus.sprecher@sam.math.ethz.ch).

used to define nonlinear finite element spaces, this aspect is never actually considered.

We are interested in functions of Sobolev smoothness. By this we mean functions from spaces

$$W^{k,p}(\Omega, M) := \left\{ v \in W^{k,p}(\Omega, \mathbb{R}^n) : v(x) \in M \text{ a.e.} \right\},$$

where we denote by  $W^{k,p}(\Omega, \mathbb{R}^n)$  the standard Sobolev space for  $k \in \mathbb{N}$  and  $p \in [1, \infty]$ . Throughout the paper,  $|\cdot|_{W^{k,p}}$  and  $\|\cdot\|_{W^{k,p}}$  will denote the corresponding Sobolev seminorm and full norm of  $\mathbb{R}^n$ -valued functions, respectively.

Spaces of approximating functions will be constructed by pointwise projection. Given a finite element grid of  $\Omega$ , and a set of values  $c_i \in M \subset \mathbb{R}^n$  at Lagrange points on  $\Omega$ , we construct nonlinear finite element functions by first interpolating in  $\mathbb{R}^n$  by piecewise polynomials in  $\mathbb{R}^n$ , and then projecting pointwise onto  $M$ . This results in a finite-dimensional set of functions  $V_h(\Omega, M)$  which, as it turns out, is a subspace of  $W^{1,p}$  for arbitrary  $p \geq 1$ . While the approach presented here is based on Lagrangian interpolation in  $\mathbb{R}^n$ , other linear finite element spaces can be used in principle (see [48] for an example).

The idea of averaging nonlinear data by Euclidean interpolation followed by a projection has been studied extensively in data analysis, robotics, and other fields. Gawlik and Leok [20] provide a list of references for interpolation on  $\text{SO}(3)$  alone. Using this approach to generalize finite element spaces has also appeared several times [48, 20, 44]. For functions taking values in the special orthogonal group  $\text{SO}(3)$ , [20] have studied  $C^1$ -norms of interpolation errors [20]. We will extend these results to general closed submanifolds of  $\mathbb{R}^n$  and to interpolation errors in Sobolev norms.

To this end, let  $Q_{\mathbb{R}^n}$  be the standard nodal interpolation operator for  $\mathbb{R}^n$ -valued Lagrangian finite elements, and set  $Q_M := \mathcal{P} \circ Q_{\mathbb{R}^n}$  the interpolation operator with a pointwise projection. For smooth manifolds, approximation qualities can be inferred from the linear ones of  $V_h(\Omega, \mathbb{R}^n)$ , as we can switch back and forth between discrete functions into  $M$  and into  $\mathbb{R}^n$  via the definition of  $Q_M$  and the identity

$$Q_{\mathbb{R}^n} \circ Q_M = Q_{\mathbb{R}^n}.$$

An alternative proof that uses the Lipschitz continuity of the closest-point projection has been given in [48].

A priori, test functions for manifold-valued settings are defined as variations of particular manifold-valued functions. We show that test functions for functions defined by polynomial interpolation and projection can also be constructed directly, using Euclidean interpolation followed by a projection. We show the same Sobolev interpolation error bounds for these discrete test functions as for the finite element functions themselves.

We then discuss finite element discretizations of PDEs with values in  $M$ . Prototypically, we focus on harmonic maps from a domain  $\Omega$  to  $M$ , which we regard as minimizers  $u$  of the Dirichlet energy in a suitable Sobolev space. Throughout the manuscript we assume that  $u$  is contained in a Sobolev space  $W^{k,p}$  such that the embedding theorem implies continuity of  $u$ . The corresponding discrete solution  $u_h$  is defined as a local minimizer of the same energy in  $V_h(\Omega, M)$ , which is well defined because  $V_h(\Omega, M)$  is suitably conforming.

To estimate  $\|v_h - u\|_{W^{1,2}}$  we combine a simple nonlinear Céa lemma with the interpolation results for  $V_h(\Omega, M)$ . To show optimal  $L^2$  bounds we use the abstract

theory of [27], showing that the four criteria stated there are fulfilled by projection-based finite elements. We will also provide inverse estimates. In classical finite element theory they are used in many proofs, e.g., in Nitsche's method of weighted norms for uniform convergence estimates [14]. In this work we will use them to justify a priori bounds on discrete minimizers of the harmonic map energy. Both interpolation and discretization error bounds are verified numerically in the two final sections.

There is one alternative construction for conforming finite element spaces for manifold-valued problems, known as geodesic finite elements [42, 43, 23, 24, 13, 40]. To evaluate the relative merits of the two methods we briefly revisit their theoretical relationship, and we repeat all numerical tests using geodesic finite elements. We observe that while geodesic finite elements yield lower errors, projection-based finite elements can be much faster if a cheap formula for the projection onto  $M$  is available. We measure a roughly 10-fold speed difference for  $M = S^2$ , whereas for  $M = \text{SO}(3)$ , where the projection needs to be computed iteratively, we observe hardly a difference.

**1. Projection-based finite element spaces.** Let  $\Omega$  be discretized by a finite union  $\mathcal{G}$  of affine-equivalent, regular, and quasi-uniform polyhedra  $T_h$ , such that the closures intersect in common faces. On  $\mathcal{G}$  we consider scalar-valued Lagrangian finite element spaces with the nodal basis  $(\phi_i)_{i \in I} : \Omega \rightarrow \mathbb{R}$  and associated Lagrange points  $(\xi_i)_{i \in I} \subset \Omega$ .

We will define the space  $V_h(\Omega, M)$  of projection-based finite elements as the image of an interpolation operator. First we consider the canonical interpolation operator  $Q_{\mathbb{R}^n}$  for continuous functions with values in  $\mathbb{R}^n$  into the space  $V_h(\Omega, \mathbb{R}^n)$  of  $\mathbb{R}^n$ -valued Lagrangian finite elements.

**DEFINITION 1.** *The interpolation operator  $Q_{\mathbb{R}^n} : C(\Omega, \mathbb{R}^n) \rightarrow V_h(\Omega, \mathbb{R}^n)$  corresponding to a set of basis functions  $(\phi_i)_{i \in I} : \Omega \rightarrow \mathbb{R}$  and nodes  $(\xi_i)_{i \in I} \subset \Omega$  is defined by*

$$Q_{\mathbb{R}^n} v := \sum_{i \in I} v(\xi_i) \phi_i \quad \forall v \in C(\Omega, \mathbb{R}^n).$$

For a manifold  $M$  embedded in  $\mathbb{R}^n$  and a function  $v \in C(\Omega, M)$  the values of  $Q_{\mathbb{R}^n} v$  will in general not be on  $M$  away from the  $\xi_i$ . To get  $M$ -valued functions we compose  $Q_{\mathbb{R}^n}$  pointwise with the closest-point projection

$$P : \mathbb{R}^n \rightarrow M, \quad P(q) := \arg \min_{r \in M} \|r - q\|_{\mathbb{R}^n},$$

where  $\|\cdot\|_{\mathbb{R}^n}$  denotes the Euclidean distance. While the closest-point projection is usually not well defined for all  $q \in \mathbb{R}^n$ , if  $M$  is regular enough it is well defined in a neighborhood  $U_\delta \subset \mathbb{R}^n$  of  $M$  [1].

This pointwise projection induces a superposition operator  $\mathcal{P}$  by

$$(\mathcal{P}v)(x) := P(v(x))$$

for all  $x \in \Omega$  and  $v : \Omega \rightarrow \mathbb{R}^n$ . We then define  $M$ -valued interpolation by composition of  $Q_{\mathbb{R}^n}$  and  $\mathcal{P}$ .

**DEFINITION 2.** *Set*

$$C(\Omega, M; \rho) := \left\{ v \in C(\Omega, M) : \text{diam}(v(T_h)) < \rho \quad \forall T_h \in \mathcal{G} \right\},$$

where the  $\text{diam}(B) := \sup_{p, q \in B} d(p, q)$  denotes the geodesic diameter of a subset  $B \subset M$ . Provided that  $\rho$  is small enough, define the interpolation operator

$$Q_M : C(\Omega, M; \rho) \rightarrow C(\Omega, M) \quad \text{by} \quad Q_M := \mathcal{P} \circ Q_{\mathbb{R}^n}.$$

The space of projection-based finite elements is defined as the range of this interpolation operator.

**DEFINITION 3.** Let  $\Omega \subset \mathbb{R}^s$ ,  $M \subset \mathbb{R}^n$  an embedded submanifold, and  $P: U_\delta \subset \mathbb{R}^n \rightarrow M$  the closest-point projection. For a given set of basis functions  $(\phi_i)_{i \in I}: \Omega \rightarrow \mathbb{R}$  we define

$$(1) \quad V_h(\Omega, M) := \left\{ v_h : \Omega \rightarrow M \text{ s.t. } \exists v \in C(\Omega, M) \text{ and } v_h = Q_M v \right\}.$$

As the operator  $Q_M$  only uses the values at the Lagrange nodes  $(\xi_i)_{i \in I}$ , we have the equivalent definition

$$V_h(\Omega, M) := \left\{ v_h : \Omega \rightarrow M, \exists (c_i)_{i \in I} \subset M \text{ s.t. } v_h = \mathcal{P} \left( \sum_{i \in I} c_i \phi_i \right) \right\}.$$

It has to be noted that while there exist nodal values  $(c_i)_{i \in I} \subset M$ ,  $c_i = v_h(\xi_i)$  for any function  $v_h \in V_h(\Omega, M)$ , for given values  $(c_i)_{i \in I} \subset M$  there exists an interpolating function  $v_h \in V_h(\Omega, M)$  only if the values  $c_i$  are close enough depending on  $M$  such that  $\sum_{i \in I} c_i \phi_i \subset U_\delta$ .

**1.1. Conformity.** The question of conformity of projection-based elements, i.e., whether  $V_h(\Omega, M) \subset W^{1,p}(\Omega, M)$  holds, can be reduced to the continuity of the superposition operator  $\mathcal{P}$  on  $W^{1,p}(\Omega, M)$  and of the operator  $Q_{\mathbb{R}^n}$ .

Denote by  $P'(x)[y] \in T_{P(x)}M$  the differential of the closest-point projection  $P$  at  $x \in U_\delta$  applied to  $y \in T_x \mathbb{R}^n \simeq \mathbb{R}^n$ . Let  $\hat{v}_h(x) = \sum_{i \in I} c_i \phi_i(x) \in \mathbb{R}^n$  for given coefficients  $c_i$ , and  $v_h = \mathcal{P} \hat{v}_h$ . By the chain rule we have

$$\frac{\partial}{\partial x_j} v_h(x) = P' \left( \sum_{i \in I} c_i \phi_i(x) \right) \left[ \sum_{i \in I} c_i \frac{\partial \phi_i}{\partial x_j}(x) \right]$$

for every  $x \in \Omega$  such that  $P$  is differentiable at  $\hat{v}_h(x)$  and  $\phi_i$  is differentiable at  $x \in \Omega$  for all  $i \in I$ .

If we assume that  $M$  is a smooth embedded submanifold, there exists a tubular neighborhood such that the closest-point projection is smooth [32, Prop. 6.1.8]. In particular, the pointwise norm of  $P'$  can be estimated in terms of the radius of curvature using explicit calculations in terms of the local parametrization of the manifold [1]. Thus, the  $W^{1,p}$ -conformity of  $V_h(\Omega, M)$  follows directly from the chain rule and smoothness of the Lagrange basis  $(\phi_i)_{i \in I} \subset W^{1,p}$ ,

$$|v_h|_{W^{1,p}} \leq \|P\|_{C^1} \sum_{i \in I} \|c_i\|_{\mathbb{R}^n} |\phi_i|_{W^{1,p}(\Omega)},$$

where  $\|P\|_{C^1}$  denotes the operator norm of the differential  $P'$ .

**1.2. Relationship to geodesic finite elements.** Projection-based finite elements are closely related to the geodesic finite elements proposed in [42, 43, 23] and analyzed in [24, 25]. Geodesic finite elements are constructed by replacing polynomial interpolation of values  $(c_i)_{i \in I}$

$$x \mapsto \sum_{i \in I} c_i \phi_i(x) = \arg \min_{q \in \mathbb{R}^n} \sum_{i \in I} \phi_i(x) \|c_i - q\|_{\mathbb{R}^n}^2$$

by the weighted Riemannian center of mass

$$(2) \quad x \mapsto \arg \min_{q \in M} \sum_{i \in I} \phi_i(x) d(c_i, q)^2,$$

where  $d(\cdot, \cdot)$  is the geodesic distance on  $M$ . Unlike the construction by pointwise projection, (2) is completely intrinsic and does not rely on an embedding space. Well-posedness of this definition under suitable conditions on the  $c_i$  is shown in [43, 25].

As observed independently by [48] and [20], we recover the projection-based interpolation if we replace the geodesic distance in (2) by the Euclidean distance of the embedding space

$$\begin{aligned} \arg \min_{q \in M} \sum_{i \in I} \phi_i(x) \|c_i - q\|_{\mathbb{R}^n}^2 &= \arg \min_{q \in M} \left( \|q\|_{\mathbb{R}^n}^2 - 2 \left\langle q, \sum_{i \in I} \phi_i(x) c_i \right\rangle \right) \\ &= \arg \min_{q \in M} \left\| q - \sum_{i \in I} \phi_i(x) c_i \right\|_{\mathbb{R}^n}^2 \\ &= P \left( \sum_{i \in I} \phi_i(x) c_i \right). \end{aligned}$$

This does not mean that projection-based finite elements are equal to geodesic finite elements for embedded manifolds. In general, even if the metric on  $M$  is obtained by an isometric embedding into Euclidean space, the distance  $d(\cdot, \cdot)$  is not the Euclidean distance in the surrounding space. Instead, projection-based interpolation can be interpreted as geodesic finite elements for a general metric space  $(M, d)$  with a non-intrinsic metric. As far as we know, no general existence theory and error estimates exist for this abstract setting.

**1.3. Preservation of isometries.** If an isometry  $T: M \rightarrow M$  commutes with the projection-based interpolation operator  $Q_M$ , then the finite element space  $V_h(\Omega, M)$  defined in (1) is equivariant under this isometry. In mechanics, this leads to the desirable property that discretizations of objective problems are again objective. Unfortunately, for projection-based finite elements this commutativity holds only under special circumstances.

**DEFINITION 4.** An isometry  $T: M \rightarrow M$  (w.r.t. the geodesic distance) is called extendable if there exists an isometry  $\tilde{T}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  with  $\tilde{T}(p) = T(p)$  for all  $p \in M$ .

Examples for extendable isometries  $T$  are orthogonal transformations for the sphere and multiplication with special orthogonal matrices for  $\text{SO}(n)$ .

**THEOREM 5.** Let  $M \subset \mathbb{R}^n$  be an embedded submanifold,  $P: \mathbb{R}^n \rightarrow M$  the closest-point projection,  $T: M \rightarrow M$  an extendable isometry, and  $(\phi_i)_{i \in I}$  a partition of unity. Then  $T$  commutes with  $Q_M$ .

*Proof.* Let  $\tilde{T}$  be an extension of  $T$  to  $\mathbb{R}^n$ . As an isometry maps closest distances to closest distances,  $\tilde{T}$  commutes with  $P$ . By the Mazur–Ulam theorem [34] there exists a linear map  $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$  with  $\tilde{T}(q) = A(q) + \tilde{T}(0)$  for all  $q \in \mathbb{R}^n$ , so that  $\tilde{T}$  obviously commutes with  $Q_{\mathbb{R}^n}$ .  $\square$

One can tell from this proof that only very few isometries  $T: M \rightarrow M$  are extendable. Indeed, in order to be extendable,  $T$  needs to be the restriction of a rigid body motion of  $\mathbb{R}^n$ . In contrast to this rather strong restriction, the geodesic interpolation rule (2) is equivariant under any isometry of  $M$  by construction.

**1.4. Discrete test functions and vector field interpolation.** The test function space for a function  $u : \Omega \rightarrow M$  consists of vector fields along  $u$  that correspond to intrinsic variations within the class of functions considered. For  $u \in H^1(\Omega, M)$ , we call the space of test functions  $H^1(\Omega, u^{-1}TM)$ . If we consider  $M \subset \mathbb{R}^n$  as an embedded submanifold, then  $H^1(\Omega, u^{-1}TM)$  can be canonically identified with a subset of  $H^1(\Omega, \mathbb{R}^n)$ .

We construct discrete test functions in the same manner, i.e.,  $\mathbf{v}_h$  is a discrete test function for  $u_h \in V_h(\Omega, M)$  if there exists a variation  $\gamma : \Omega \times (-\epsilon, \epsilon) \rightarrow M$  such that  $\gamma(\cdot, t) \in V_h(\Omega, M)$  for all  $t \in (-\epsilon, \epsilon)$ ,  $\gamma(\cdot, 0) = u_h$ , and  $\frac{d}{dt}\gamma(\cdot, 0) = \mathbf{v}_h$  [25, 44]. Writing this definition using the coefficients  $(c_i)_{i \in I} \subset M$  that constitute  $u_h$ , the set of all discrete test functions over the discrete function  $u_h$  can be defined as

$$W_h(\Omega, u_h^{-1}TM) := \left\{ \mathbf{v}_h \in L^2(\Omega, \mathbb{R}^n) : \exists (c_i)_{i \in I} \subset C((-\epsilon, \epsilon), M) \text{ s.t.} \right. \\ \left. c_i(0) = u_h(\xi_i) \forall i \in I \text{ and } \mathbf{v}_h = \frac{d}{dt} \Big|_{t=0} \mathcal{P} \left( \sum_{i \in I} c_i(t) \phi_i \right) \right\}.$$

Similar to the discrete functions themselves, discrete test functions can be constructed by polynomial interpolation followed by pointwise projection, as by the chain rule we have for any  $\mathbf{v}_h \in W_h$  and  $x \in \Omega$

$$\mathbf{v}_h(x) = P' \left( \sum_{i \in I} c_i(0) \phi_i(x) \right) \left( \sum_{i \in I} c'_i(0) \phi_i(x) \right),$$

and  $P'(y) : T_y \mathbb{R}^n \rightarrow T_{P(y)}M$ , the differential of the closest-point projection  $P$ , is again a projection.

**PROPOSITION 6.** *Let  $P : U_\delta \subset \mathbb{R}^n \rightarrow M$  be the closest-point projection onto a closed embedded  $C^2$ -submanifold  $M \subset \mathbb{R}^n$ . If  $y \in M$ , then  $P'(y) : T_y U_\delta \rightarrow T_y M$  is the orthogonal projection onto the tangent space  $T_y M$  with the canonical interpretation of  $T_y M$  as a subspace of  $\mathbb{R}^n$ . In general, the differential  $P' : TU_\delta \rightarrow TM$  is a projection onto the tangent bundle  $TM$ . For every  $y \in U_\delta$ ,  $P'(y) : T_y U_\delta \rightarrow T_{P(y)}M$  differs from the orthogonal projection onto  $T_{P(y)}M$  by the Lipschitz constant of  $P'$  and the distance of  $y$  to  $M$ .*

*Proof.* Let  $y \in M \subset U_\delta$ , and  $\xi \in T_y U_\delta = \mathbb{R}^n$ . We need to show that

$$\langle P'(y)(\xi) - \xi, \omega \rangle = 0$$

for all  $\omega \in T_y M \subset \mathbb{R}^n$ . To see this, we consider the curve  $c : (-\epsilon, \epsilon) \rightarrow M$  defined by  $c(t) = P(y + t\xi)$ , and a vector field  $w(t)$  along  $c$  with  $w(0) = \omega$ . As  $P$  is defined by minimization, the first variation yields at any  $t \in (-\epsilon, \epsilon)$

$$\langle P(y + t\xi) - (y + t\xi), w(t) \rangle = 0.$$

Differentiating this with respect to  $t$  yields

$$\begin{aligned} 0 &= \frac{d}{dt} \Big|_{t=0} \langle P(y + t\xi) - (y + t\xi), w(t) \rangle \\ &= \langle P'(y)(\xi) - \xi, \omega \rangle + \langle P(y) - y, w'(0) \rangle \\ &= \langle P'(y)(\xi) - \xi, \omega \rangle. \end{aligned}$$

Thus,  $P'(y)$  is indeed the orthogonal projection for  $y \in M$ . This immediately implies that  $P'$  restricted to  $TM$  is the identity, and hence  $P'$  is a projection even if  $y \notin M$ . If  $y \in U_\delta \setminus M$ ,  $P(y) - y$  does not vanish, and neither does the normal part of  $w'(0)$ . For  $\xi \in \mathbb{R}^n$  we can estimate

$$\|P'(y)(\xi) - P'(P(y))(\xi)\| \leq \text{Lip}(P') d(y, M) |\xi|,$$

which was asserted.  $\square$

Thus, the computation of the value at  $x \in \Omega$  of a test function along a discrete function  $u_h \in V_h(\Omega, M)$  corresponds to first interpolating given tangent vectors  $c'_i \in T_{u_h(\xi_i)}M \subset \mathbb{R}^n$  in  $\mathbb{R}^n$ , and then projecting the resulting piecewise polynomial function pointwise to  $T_{u_h(x)}M$ . Alternatively, by linearity we can first project the  $c'_i$  to  $T_{u_h(x)}M$ , and then interpolate the result in the vector space  $T_{u_h(x)}M$ .

In particular, we can define for  $u_h = \mathcal{P}(\sum_{i \in I} u_i \phi_i) \in V_h(\Omega, M)$  an interpolation operator  $Q_{u_h^{-1}TM} : C(\Omega, u_h^{-1}TM) \rightarrow W_h(\Omega, u_h^{-1}TM)$  by

$$Q_{u_h^{-1}TM} \mathbf{v} = \mathcal{P}_{u_h^{-1}TM} \circ Q_{\mathbb{R}^n}(\mathbf{v}).$$

Given some test function  $\mathbf{v}$  along a continuous function  $u$ , i.e.,  $(u, \mathbf{v}) \in C(\Omega, TM)$ , we can first interpolate  $u$  and then  $\mathbf{v}$ , as the interpolation of  $\mathbf{v}$  depends only on the values at the Lagrange nodes  $(\xi_i)_{i \in I}$ , where  $u$  and  $u_h$  agree.

**2. Interpolation error estimates.** In this chapter we will estimate the interpolation errors of  $Q_{\mathbb{R}^n}$  and  $Q_M$  in terms of the mesh width  $h$ . We also estimate the error of the test vector field interpolation operator  $Q_{u_h^{-1}TM}$ .

**2.1. Properties of Euclidean interpolation.** Proving interpolation error bounds for  $Q_M$  uses several standard results for interpolation in Euclidean spaces. We repeat some of them here for convenience.

Define the usual grid-dependent Sobolev norms

$$\|v\|_{W^{l,p}(\mathcal{G})} := \left( \sum_{T \in \mathcal{G}} \|v\|_{W^{l,p}(T)}^p \right)^{\frac{1}{p}}$$

for functions  $v \in C(\Omega)$  such that  $v|_T \in W^{l,p}(T)$  for all  $T \in \mathcal{G}$ . For the rest of this paper, this norm is meant whenever we speak of the  $W^{l,p}$ -norm of a discrete function, unless explicitly stated otherwise. As we assume shape regularity of the mesh, one can use the Sobolev embedding theorem and elementwise scaling to the reference element to prove that if  $lp > s$ ,  $Q_{\mathbb{R}^n}$  is continuous with respect to the grid-dependent  $W^{l,p}$ -norm, i.e., there exists  $C > 0$  such that we have

$$(3) \quad \|Q_{\mathbb{R}^n} v\|_{W^{l,p}(\mathcal{G})} \leq C \|v\|_{W^{l,p}(\Omega)} \quad \forall v \in W^{l,p}(\Omega).$$

Note that by the Sobolev embedding theorem,  $Q_{\mathbb{R}^n} v$  is well defined for all  $v \in W^{l,p}$  with  $lp > s$ . Under these assumptions, we have the following approximation error estimate for  $Q_{\mathbb{R}^n}$  [14, 12].

**THEOREM 7.** *Let  $\Omega \subset \mathbb{R}^s$  be a bounded Lipschitz domain,  $\mathcal{G}$  a shape-regular, affine-equivalent mesh on  $\Omega$ ,  $l \in \mathbb{N}$ ,  $p \in [1, \infty]$ , and  $(\phi_i)_{i \in I}$  Lagrangian nodal basis functions for polynomial order  $r \geq l - 1$ . Then on each element  $T$  for any  $m \in \mathbb{N}$*

with  $m \geq l$  and  $\min(m, r+1) > \frac{s}{p}$  we have

$$|v - Q_{\mathbb{R}^n} v|_{W^{l,p}(T, \mathbb{R}^n)} \leq C h_T^{\min(m, r+1)-l} |v|_{W^{\min(m, r+1), p}(T, \mathbb{R}^n)} \quad \forall v \in W^{m,p}(T, \mathbb{R}^n)$$

with the constant independent of  $v$  and  $h_T = \text{diam}(T)$ .

We will also need the following inverse inequalities.

**THEOREM 8.** *Consider a shape-regular, affine-equivalent, quasi uniform mesh  $\mathcal{G}$ , and two pairs  $(k, p)$  and  $(m, q)$  with  $1 \leq k \leq m < \infty$  and  $p, q \in [1, \infty]$ . Then on each mesh element  $T \in \mathcal{G}$  and for all discrete functions  $Q_{\mathbb{R}^n} v$  of polynomial order  $r$*

$$|Q_{\mathbb{R}^n} v|_{W^{m,q}(T)} \leq C h^{-(m-k)-s \max\{0, \frac{1}{p} - \frac{1}{q}\}} |Q_{\mathbb{R}^n} v|_{W^{k,p}(T)},$$

where the constant depends on the quasi uniformity and regularity parameters of the mesh, but not on  $h$ .

This result here is not as general as it could be. Inverse inequalities with weaker requirements on the mesh appear, e.g., in [16] (no quasi uniformity) and [21] (no shape regularity). We expect that these generalizations can help to extend the following results on  $M$ -valued interpolation as well.

**2.2.  $M$ -valued interpolation.** We now turn to error bounds for the  $M$ -valued interpolation operator  $Q_M := \mathcal{P} \circ Q_{\mathbb{R}^n}$ . Given  $v \in W^{m,p}(\Omega, M)$ , we estimate the error  $\|Q_M v - v\|_{W^{l,p}(\mathcal{G})}$ ,  $l \leq \min(m, r+1)$ , by observing that  $Q_{\mathbb{R}^n}(Q_M v) = Q_{\mathbb{R}^n} v$  and using the triangle inequality

$$\|Q_M v - v\|_{W^{l,p}} \leq \|Q_M v - Q_{\mathbb{R}^n}(Q_M v)\|_{W^{l,p}} + \|Q_{\mathbb{R}^n} v - v\|_{W^{l,p}}$$

(again in the grid-dependent norm). Denoting  $\hat{m} := \min(m, r+1)$ , both terms on the right can be bounded using Theorem 7, and we obtain

$$\|Q_M v - v\|_{W^{l,p}} \leq C h^{\hat{m}-l} (|Q_M v|_{W^{\hat{m},p}} + |v|_{W^{\hat{m},p}}).$$

It remains to show estimates of  $|Q_M v|_{W^{\hat{m},p}}$  in terms of Sobolev norms of  $v$ . Unlike in the Euclidean case the Sobolev seminorm  $|v|_{W^{\hat{m},p}}$  is not by itself sufficient to bound  $|Q_M v|_{W^{\hat{m},p}}$ , because lower-order derivatives appear by the chain rule. The proper quantity is the homogeneous norm  $|\cdot|_{W^{\hat{m},p}} + |\cdot|_{W^{1,\hat{m},p}}^{\hat{m}}$ , known, e.g., from [11]. It replaces the unwieldy smoothness descriptor used in corresponding results for geodesic finite elements [24].

**PROPOSITION 9.** *Let  $\hat{m} \geq 1$ ,  $p \in [1, \infty]$ , and  $M$  be such that the closest-point projection  $P$  is in  $C^{\hat{m}}$  in some  $\delta$ -neighborhood  $U_\delta \subset \mathbb{R}^n$  of  $M$ . Let  $Q_{\mathbb{R}^n} v$  be a discrete function from a Lipschitz domain  $T \subset \mathbb{R}^s$  into  $\mathbb{R}^n$  defined by interpolation of values  $(c_i)_{i \in I} \subset M$ . Suppose the  $(c_i)_{i \in I}$  are contained in a geodesic ball  $B_\rho$  of radius  $\rho$ , where  $\rho$  is small enough such that  $Q_{\mathbb{R}^n} v(T) \subset U_\delta$ . Then*

$$(4) \quad |Q_M v|_{W^{\hat{m},p}} \leq (1 + C \text{Lip}(P')\rho + C\|P\|_{C^{\hat{m}}}) |Q_{\mathbb{R}^n} v|_{W^{\hat{m},p}} + C\|P\|_{C^{\hat{m}}} \max\{|Q_{\mathbb{R}^n} v|_{W^{1,\hat{m},p}}^{\hat{m}}, |Q_{\mathbb{R}^n} v|_{W^{1,\hat{m},p}}\}.$$

*Proof.* Let  $\vec{a} \in \mathbb{N}^s$  be a multi-index with  $|\vec{a}|_1 = \hat{m}$ . By the chain rule, the derivative  $D^{\vec{a}} Q_M v = D^{\vec{a}}(\mathcal{P}(Q_{\mathbb{R}^n} v))$  can be written almost everywhere as a sum of terms of the form

$$P^{(k)}(Q_{\mathbb{R}^n} v(x)) [D^{\vec{a}_1} Q_{\mathbb{R}^n} v(x), \dots, D^{\vec{a}_k} Q_{\mathbb{R}^n} v(x)],$$



where  $1 \leq k \leq \hat{m}$ ,  $\vec{a}_1, \dots, \vec{a}_k \in \mathbb{N}^s \setminus \{(0, \dots, 0)\}$ , and  $\vec{a}_1 + \dots + \vec{a}_k = \vec{a}$ . An expansion of  $P$  around  $M$  yields

$$(5) \quad \|\mathcal{P}'(Q_{\mathbb{R}^n} v)\|_{L^\infty} \leq 1 + \text{Lip}(P')d(Q_{\mathbb{R}^n} v, M) \leq 1 + C \text{Lip}(P')\rho,$$

where  $d(Q_{\mathbb{R}^n} v, M) := \sup_{x \in T} \inf_{z \in M} \|Q_{\mathbb{R}^n} v(x) - z\|_{\mathbb{R}^n}$ , and  $\text{Lip}(P')$  denotes the Lipschitz constant of the map  $P' : \mathbb{R}^{2n} \rightarrow TM$ . Further, we have

$$\begin{aligned} \left\| P^{(k)}(Q_{\mathbb{R}^n} v(x)) [D^{\vec{a}_1} Q_{\mathbb{R}^n} v(x), \dots, D^{\vec{a}_k} Q_{\mathbb{R}^n} v(x)] \right\|_{L^p} \\ \leq \|\mathcal{P}^{(k)}(Q_{\mathbb{R}^n} v)\|_{L^\infty} \prod_{i=1}^k \|D^{\vec{a}_i} Q_{\mathbb{R}^n} v\|_{L^{\frac{\hat{m}p}{|\vec{a}_i|}}}. \end{aligned}$$

For  $k = 1$ , this yields

$$\|P'(Q_{\mathbb{R}^n} v(x)) [D^{\vec{a}} Q_{\mathbb{R}^n} v(x)]\|_{L^p} \leq (1 + C \text{Lip}(P')\rho) |Q_{\mathbb{R}^n} v|_{W^{\hat{m}, p}}.$$

For  $k \geq 2$ , we have by the Gagliardo–Nirenberg–Sobolev inequality (applied to first-order derivative terms of  $Q_{\mathbb{R}^n} v$ ),

$$\|D^{\vec{a}_i} Q_{\mathbb{R}^n} v\|_{L^{\frac{\hat{m}p}{|\vec{a}_i|}}} \leq C |Q_{\mathbb{R}^n} v|_{W^{\hat{m}, p}}^{\frac{|\vec{a}_i|-1}{\hat{m}-1}} |Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}^{1-\frac{|\vec{a}_i|-1}{\hat{m}-1}} + C |Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}.$$

Multiplying these terms and using Young's inequality yields

$$\prod_{i=1}^k \|D^{\vec{a}_i} Q_{\mathbb{R}^n} v\|_{L^{\frac{\hat{m}p}{|\vec{a}_i|}}} \leq C (|Q_{\mathbb{R}^n} v|_{W^{\hat{m}, p}} + \max \{|Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}^{\hat{m}}, |Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}\}).$$

Combining all of this leads to

$$\begin{aligned} \|D^{\vec{a}} Q_M v\|_{L^p} &\leq \sum \left\| P^{(k)}(Q_{\mathbb{R}^n} v(x)) [D^{\vec{a}_1} Q_{\mathbb{R}^n} v(x), \dots, D^{\vec{a}_k} Q_{\mathbb{R}^n} v(x)] \right\|_{L^p} \\ &\leq (1 + C \text{Lip}(P')\rho) \|D^{\vec{a}} Q_{\mathbb{R}^n} v\|_{L^p} \\ &\quad + C \|P\|_{C^{\hat{m}}} \left( |Q_{\mathbb{R}^n} v|_{W^{\hat{m}, p}} + \max \{|Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}^{\hat{m}}, |Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}\} \right). \quad \square \end{aligned}$$

If  $m \geq r + 1$ , then the highest-order derivatives  $D^{(\hat{m})} Q_{\mathbb{R}^n} v$  vanish. In that case, (4) reduces to

$$|Q_M v|_{W^{\hat{m}, p}} \leq C \|P\|_{C^{\hat{m}}} \max \{|Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}^{\hat{m}}, |Q_{\mathbb{R}^n} v|_{W^{1, \hat{m}p}}\}$$

for the mesh-dependent norm.

We can now state the main theorem.

**THEOREM 10.** *Consider the same setting as in Theorem 7. Let  $M \subset \mathbb{R}^n$  be an embedded submanifold, such that the closest-point projection  $P: U \subset \mathbb{R}^n \rightarrow M$  is in  $C^{\hat{m}}$ , where  $\hat{m} := \min(m, r + 1) > \frac{s}{p}$ . Then there exists  $\alpha > 0$ , depending on  $m, p$ , and  $s$  such that for all  $v \in W^{m, p}(\Omega, M)$  and  $0 \leq l \leq \hat{m}$  we can choose  $h$  small enough such that*

$$(6) \quad |Q_M v - v|_{W^{l, p}} \leq C h^{\hat{m}-l} \left[ (1 + C \text{Lip}(P')\rho + C \|P\|_{C^{\hat{m}}}) |v|_{W^{\hat{m}, p}} + C \|P\|_{C^{\hat{m}}} \max \{|v|_{W^{1, \hat{m}p}}^{\hat{m}}, |v|_{W^{1, \hat{m}p}}\} \right]$$

with the constants depending on the constants in Theorem 7 and Proposition 9.

*Proof.* As the map  $v \in W^{m,p}(\Omega, M)$  is continuous, we can choose the mesh width  $h$  small enough such that the assumptions of Proposition 9 are fulfilled with  $\rho := \max_{T_h} \max_{i,j} d(v(\xi_{i,T_h}), v(\xi_{j,T_h}))$ . For  $l = 0$ , we have, using (5) and Theorem 7,

$$\begin{aligned} \|Q_M v - v\|_{L^p} &\leq \int_0^1 \|\mathcal{P}'(v + t(Q_{\mathbb{R}^n} v - v))(Q_{\mathbb{R}^n} v - v)\|_{L^p} dt \\ &\leq (1 + C \operatorname{Lip}(P')\rho) \|Q_{\mathbb{R}^n} v - v\|_{L^p} \\ &\leq C (1 + \operatorname{Lip}(P')\rho) h^{\hat{m}} |v|_{W^{\hat{m},p}}. \end{aligned}$$

For  $l \geq 1$ , we use  $Q_{\mathbb{R}^n} \circ Q_M = Q_{\mathbb{R}^n}$ , Theorem 7, and Proposition 9 to estimate

$$\begin{aligned} \|Q_M v - v\|_{W^{l,p}} &\leq \|Q_M v - Q_{\mathbb{R}^n}(Q_M v)\|_{W^{l,p}} + \|Q_{\mathbb{R}^n} v - v\|_{W^{l,p}} \\ &\leq C h^{\hat{m}-l} (|v|_{W^{\hat{m},p}} + |Q_M v|_{W^{\hat{m},p}}) \\ &\leq C h^{\hat{m}-l} (|v|_{W^{\hat{m},p}} + (1 + C \operatorname{Lip}(P')\rho + C \|P\|_{C^{\hat{m}}}) |Q_{\mathbb{R}^n} v|_{W^{\hat{m},p}} \\ &\quad + C \|P\|_{C^{\hat{m}}} \max \{ |Q_{\mathbb{R}^n} v|_{W^{1,\hat{m}p}}^{\hat{m}}, |Q_{\mathbb{R}^n} v|_{W^{1,\hat{m}p}} \} ). \end{aligned}$$

By Theorem 7, we can estimate all arising seminorms of  $Q_{\mathbb{R}^n} v$  by corresponding seminorms of  $v$ .  $\square$

If the projection  $P$  is in  $W^{\hat{m}+1,\infty}$  in the neighborhood of  $M$ , we can estimate

$$\|P\|_{C^{\hat{m}}} \leq C(M) + CL(P)\rho,$$

where the constant  $L$  depends on the  $W^{k+1,\infty}$ -norm of  $P$ , and the constant  $C(M)$  depends on the extrinsic curvature of  $M$ . Instead of (6), we can then write

$$\begin{aligned} |Q_M v - v|_{W^{l,p}} &\leq C h^{\hat{m}-l} \left[ |v|_{W^{\hat{m},p}} + (C(M) \right. \\ &\quad \left. + CL(P)\rho) (|v|_{W^{\hat{m},p}} + \max \{ |v|_{W^{1,\hat{m}p}}^{\hat{m}}, |v|_{W^{1,\hat{m}p}} \}) \right]. \end{aligned}$$

If  $\{\xi_{i,T_h}\}_i$  denote the Lagrangian interpolation nodes in an element  $T_h$ , we have by the Sobolev embedding theorem for some  $\alpha > 0$

$$\rho = \max_{T_h} \max_{i,j} d(v(\xi_{i,T_h}), v(\xi_{j,T_h})) \leq C \|v\|_{C^{0,\alpha}} |\xi_{i,T_h} - \xi_{j,T_h}|^\alpha \leq C \|v\|_{W^{\hat{m},p}} h^\alpha.$$

Thus, the factor  $L(P)$  is multiplied with  $h^\alpha$ ,  $\alpha > 0$ , and becomes irrelevant for  $h \rightarrow 0$ .

We have seen that, due to the chain rule, estimates on  $Q_M$  obtained from the ones on  $Q_{\mathbb{R}^n}$  are always with respect to the homogeneous Sobolev seminorms of the type  $|\cdot|_{W^{m,q}} + |\cdot|_{W^{1,mq}}^m$ . As the term  $|\cdot|_{W^{1,mq}}^m$  does not scale correctly, we cannot expect general inverse estimates in the style of Theorem 8 for  $Q_M$ . An exception is the special case  $m = 1$ .

**THEOREM 11.** *Let the assumptions of Theorem 8 be fulfilled with  $k = m = 1$  and  $p > s$ . Then for all projected finite element functions  $Q_M v$  we have*

$$|Q_M v|_{W^{1,q}(T)} \leq C h^{-s \max\{0, \frac{1}{p} - \frac{1}{q}\}} |Q_M v|_{W^{1,p}(T)}.$$

*Proof.* The estimate follows directly from Proposition 9 and Theorem 8. We need the condition  $p > s$  in order to apply Theorem 7 to estimate

$$|Q_{\mathbb{R}^n} v|_{W^{1,p}} \leq |Q_M v|_{W^{1,p}} + |Q_{\mathbb{R}^n} Q_M v - Q_M v|_{W^{1,p}} \leq C |Q_M v|_{W^{1,p}}. \quad \square$$

**2.3.  $TM$ -valued interpolation.** In section 1.4, we have defined interpolation of a vector field  $\mathbf{v} \in C(\Omega, u_h^{-1}TM)$  along a discrete function  $u_h$  by  $(Q_{u_h^{-1}TM}\mathbf{v})(x) := P'(u_h(x))(Q_{\mathbb{R}^n}\mathbf{v}(x))$ . This definition is very similar to that of  $M$ -valued interpolation, with the difference that the pointwise projection  $P'(u_h(x))$  is even linear. This linearity makes proving optimal interpolation error bounds for vector fields along given discrete functions much easier than proving the error bounds for the discrete functions themselves.

**THEOREM 12.** *Let  $\Omega \subset \mathbb{R}^s$  be a bounded Lipschitz domain,  $mp > s$ ,  $u_h \in V_h(\Omega, M)$  such that  $\|u_h\|_{W^{m,p}}$  is bounded independently of  $h$ ,  $\mathbf{v} \in W^{m,p} \cap C(\Omega, u_h^{-1}TM)$ , and  $0 \leq l < \min(m, r+1)$ . Let the assumptions of Theorem 7 be satisfied. Assume further that  $P$  is in  $W^{l+1,\infty}$  on  $U_\delta$ . Then there exists a constant  $C > 0$  depending on  $P$  such that*

$$\begin{aligned} |Q_{u_h^{-1}TM}\mathbf{v} - \mathbf{v}|_{W^{l,p}} \\ \leq C h^{\min(m,r+1)-l} |\mathbf{v}|_{W^{\min(m,r+1),p}} \left[ 1 + \|u_h\|_{W^{m,p}} + \|u_h\|_{W^{1,mp}}^m \right]. \end{aligned}$$

*Proof.* As  $\mathbf{v}(x) = P'(u_h(x))(\mathbf{v}(x))$ , we can estimate for  $l = 0$ ,

$$\begin{aligned} \|Q_{u_h^{-1}TM}\mathbf{v} - \mathbf{v}\|_{L^p} &= \left( \int_{\Omega} |P'(u_h(x))(Q_{\mathbb{R}^n}\mathbf{v}(x) - \mathbf{v}(x))|^p dx \right)^{\frac{1}{p}} \\ &\leq \|P'(u_h)\|_{L^\infty} \|Q_{\mathbb{R}^n}\mathbf{v} - \mathbf{v}\|_{L^p}. \end{aligned}$$

For  $l > 0$ , we have using the chain rule

$$\begin{aligned} |Q_{u_h^{-1}TM}\mathbf{v} - \mathbf{v}|_{W^{l,p}} &\leq \|P'(u_h)\|_{L^\infty} \|Q_{\mathbb{R}^n}\mathbf{v} - \mathbf{v}\|_{W^{l,p}} \\ &\quad + C \sum_{k=1}^l \|\mathcal{P}^{(k+1)}(u_h)\|_{L^\infty} (\|u_h\|_{W^{k,\frac{p}{k}}} + \|u_h\|_{W^{1,s}}^k \\ &\quad + \|u_h\|_{W^{1,s}}) \|Q_{\mathbb{R}^n}\mathbf{v} - \mathbf{v}\|_{W^{l-k,\frac{ps}{s-kp}}}. \end{aligned}$$

Application of Theorem 7 concludes the proof.  $\square$

**3. Discretization error estimates for harmonic maps.** We use the interpolation results of the previous section to show optimal discretization error bounds for projection-based finite element approximations of harmonic maps. For an open domain  $\Omega \subset \mathbb{R}^s$  with piecewise  $C^1$ -boundary, a smooth Riemannian manifold  $(M, g)$ , and a smooth map  $v : \Omega \rightarrow M$ , we define the harmonic energy by

$$(7) \quad \mathcal{J}(v) := \frac{1}{2} \int_{\Omega} |dv|_g^2 dx.$$

For a review on harmonic maps between Riemannian manifolds we refer to [28]. In the following we will always assume that  $\mathcal{J}$  has a unique local minimizer within  $H_\varphi^1(\Omega, M)$ , where  $H_\varphi^1(\Omega, M)$  is the set of  $H^1(\Omega, M)$ -functions  $v$  in the same homotopy class as the given function  $\varphi$ , and with  $v = \varphi$  on  $\partial\Omega$ .

We construct discrete harmonic maps  $u_h$  by minimizing  $\mathcal{J}$  in the projection-based finite element space  $V_h(\Omega, M)$ . Generalizing the approach for the Euclidean case, we prove  $H^1$ -error bounds by combining a nonlinear Céa lemma with an interpolation error bound. We then use the non-Euclidean Aubin–Nitsche trick of [27] to obtain bounds on the  $L^2$ -error. We stress that all these results hold only under the assumption of the minimizer  $u$  being continuous.

**3.1. Ellipticity.** We start with a definition of ellipticity for manifold-valued functions. Unlike [24], we define ellipticity with respect to an extrinsic error measure. The definition is locally equivalent to the intrinsic definition of [24] but is easier to use in the case of embedded manifolds.

DEFINITION 13. Let  $M \subset \mathbb{R}^n$  be an embedded submanifold and  $\Omega \subset \mathbb{R}^s$  a domain. A functional  $\mathcal{J}: H \subset H^1(\Omega, M) \rightarrow \mathbb{R}$  is called  $H^1$ -elliptic around  $u \in H \cap C(\Omega, M)$  if there exist  $\lambda, \Lambda, \epsilon > 0$  such that for all  $v \in B_\epsilon^\infty(u)$ , with

$$B_\epsilon^\infty(u) := \{v \in H \cap C(\Omega, M) \mid d_{L^\infty}(u, v) \leq \epsilon\},$$

we have

$$\lambda|v - u|_{H^1}^2 \leq \mathcal{J}(v) - \mathcal{J}(u) \leq \Lambda|v - u|_{H^1}^2.$$

Having ellipticity it is straightforward to prove a nonlinear C  a lemma.

LEMMA 14. Let  $M \subset \mathbb{R}^n$  be an embedded submanifold of  $\mathbb{R}^n$ ,  $\Omega \subset \mathbb{R}^s$ , and  $\mathcal{J}: H \subset H^1(\Omega, M) \rightarrow \mathbb{R}$  a functional with a minimizer  $u \in H \cap C(\Omega, M)$  that is unique in a closed ball  $B_\epsilon^\infty(u)$ . Assume that  $\mathcal{J}$  is elliptic around  $u$ . Let  $V \cap B_\epsilon^\infty(u) \subset H$  be a nonempty subset and

$$v := \arg \min_{w \in V \cap B_\epsilon^\infty(u)} \mathcal{J}(w).$$

Then

$$|v - u|_{H^1} \leq \sqrt{\frac{\Lambda}{\lambda}} \inf_{w \in V \cap B_\epsilon^\infty(u)} |w - u|_{H^1}.$$

*Proof.* By the ellipticity we have for any  $w \in V \cap B_\epsilon(u)$

$$\lambda|v - u|_{H^1}^2 \leq \mathcal{J}(v) - \mathcal{J}(u) \leq \mathcal{J}(w) - \mathcal{J}(u) \leq \Lambda|w - u|_{H^1}^2.$$

Taking the square root yields the desired result.  $\square$

In the following  $\mathcal{J}$  will always denote the Dirichlet energy (7). If  $M$  is isometrically immersed in  $\mathbb{R}^n$ , it is well known [18, 28] that the Euler–Lagrange equation for critical points of  $\mathcal{J}$  is

$$(8) \quad \Delta u + \operatorname{tr} A(du, du) = 0,$$

where  $A$  denotes the second fundamental form of  $M$ . Thus,  $u \in H^2(\Omega, M)$  is a critical point of  $\mathcal{J}$  only if

$$(9) \quad \Delta u(x) \in T_{u(x)} M^\perp$$

almost everywhere. Written in local coordinates, (8) is a semilinear second-order elliptic system of partial differential equations. We show that it is also elliptic in the sense of Definition 13.

PROPOSITION 15. Let  $M$  be an embedded submanifold of  $\mathbb{R}^n$  such that the closest-point projection  $P: \mathbb{R}^n \supset U_\delta \rightarrow M$  is in  $C^2$ . Let further  $\Omega \subset \mathbb{R}^s$  and  $q > \max\{2, s\}$  have a combined Sobolev and Poincar   constant  $C_P$  for the estimate  $\|\cdot\|_{L^{\frac{2q}{q-2}}} \leq C_P \|\cdot\|_{H_0^1}$ , and let  $u \in W^{1,q}(\Omega, M)$  be a critical point of the harmonic energy  $\mathcal{J}$ . Let  $\kappa$  denote the largest principal curvature of  $M$ . Then if  $\|du\|_{L^q} < (\kappa C_P)^{-1}$ , the functional  $\mathcal{J}$  is elliptic around  $u$ .

*Proof.* For  $v \in H_\varphi^1(\Omega, M)$  we have

$$(10) \quad \mathcal{J}(v) - \mathcal{J}(u) = \frac{1}{2} |v - u|_{H^1}^2 - \langle v - u, \Delta u \rangle_{H^1, H^{-1}}.$$

Let  $\gamma(x, \cdot) : [0, 1] \rightarrow M$  be a smooth family of curves connecting  $u$  and  $v$  pointwise. Then  $\langle \dot{\gamma}(\cdot, 0), \Delta u \rangle = 0$  by (9), and

$$\begin{aligned} \langle v - u, \Delta u \rangle(x) &= \int_0^1 (1-t) \langle \ddot{\gamma}(x, t), \Delta u(x) \rangle dt \\ &= - \int_0^1 (1-t) \langle \ddot{\gamma}(x, t), \text{tr } A(du(x), du(x)) \rangle dt \end{aligned}$$

for almost every  $x$  in  $\Omega$ . Suppose that  $\gamma$  is a geodesic homotopy; then  $\ddot{\gamma} = A(\dot{\gamma}, \dot{\gamma})N$ , where  $N$  is the outer normal to  $M$ . Thus we obtain, using the Hölder inequality, the estimate

$$|\langle v - u, \Delta u \rangle_{H^1, H^{-1}}| \leq \frac{\kappa^2}{2} \|du\|_{L^q}^2 \|\dot{\gamma}\|_{L^{\frac{2q}{q-2}}}^2.$$

Set  $\alpha(x, t) := P(u(x) + t(v(x) - u(x)))$ , assuming that  $v \in B_\epsilon^\infty(u)$  with  $\epsilon$  small enough such that  $\alpha$  is well defined. Then  $\|\dot{\alpha}(x, t)\| \leq (1 + \text{Lip}(P')\epsilon)|v(x) - u(x)|$ . As  $\gamma$  is pointwise length minimizing, we can estimate

$$\|\dot{\gamma}\|_{L^{\frac{2q}{q-2}}} \leq \left\| \int_0^1 \|\dot{\alpha}(\cdot, t)\| dt \right\|_{L^{\frac{2q}{q-2}}} \leq (1 + \text{Lip}(P')\epsilon) \|v - u\|_{L^{\frac{2q}{q-2}}}.$$

Using the Sobolev and Poincaré inequalities, we obtain for  $\epsilon$  small enough, depending on the strict bound on  $\|du\|_{L^q}$  and  $\text{Lip}(P')$ ,

$$|\langle v - u, \Delta u \rangle_{H^1, H^{-1}}| \leq \frac{\kappa^2}{2} \|du\|_{L^q}^2 (1 + \text{Lip}(P')\epsilon)^2 C_P^2 |v - u|_{H^1}^2 < \frac{1}{2} |v - u|_{H^1}^2.$$

Plugging this back into (10) yields the assertion.  $\square$

**3.2. Discretization error estimates in  $H^1$ .** We will now combine the Céa lemma and the approximation properties of the space  $V_h(\Omega, M)$ .

**THEOREM 16.** *Let the assumptions of Proposition 15 be fulfilled, and let the local minimizer  $u$  of the harmonic energy be in  $H^m(\Omega, M)$  with  $m > \frac{s}{2}$ . Further assume that the assumptions of Theorem 10 are met for this  $m$  and  $p = 2$ . Additionally suppose that  $\|du\|_{L^q} < (\kappa C_P)^{-1}$ , where  $\kappa$  denotes the largest principal curvature of  $M$ , and  $C_P$  is the Sobolev and Poincaré constant of  $\Omega$  for the estimate  $\|\cdot\|_{L^{\frac{2q}{q-2}}} \leq C_P \|\cdot\|_{H_0^1}$ . Let  $\varphi$  be such that  $V_h \cap H_\varphi^1(\Omega, M)$  is not empty, and set*

$$V_{h;K} := \left\{ w \in V_h \cap H_\varphi^1(\Omega, M) : \|dw\|_{L^q} \leq K \right\},$$

with  $K$  large enough and  $h$  small enough such that  $Q_M u \in V_{h;K} \cap B_\epsilon^\infty(u)$ , where  $\epsilon$  is small enough such that  $\mathcal{J}$  is elliptic in an  $\epsilon$ -neighborhood of  $u$ . Set

$$(11) \quad u_h := \arg \min_{w \in V_{h;K} \cap B_\epsilon^\infty(u)} \mathcal{J}(w).$$

Then for  $h$  small enough and  $\hat{m} := \min(m, r+1)$  we have

$$(12) \quad |u_h - u|_{H^1} \leq C \sqrt{\frac{\Lambda}{\lambda}} h^{\hat{m}-1} \left[ |u|_{H^{\hat{m}}} + \max \left\{ |u|_{W^{1,2\hat{m}}}^{\hat{m}}, |u|_{W^{1,2\hat{m}}} \right\} \right].$$

If the grid is shape-regular, affine-equivalent, and quasi-uniform, the map  $u_h$  is indeed a local minimizer of  $\mathcal{J}$  in  $V_h \cap H_\varphi^1(\Omega, M)$ , if we additionally assume that  $q < \frac{2s}{s-2\min(m-1, r)}$  in case  $2\min(m-1, r) < s$ .

*Proof.* By restriction to  $V_{h;K} \cap B_\epsilon^\infty(u)$  and the choice of  $K$  and  $\epsilon$  we can apply Lemma 14 and Theorem 10 to obtain for  $h$  small enough depending on  $u$  and the arising constants that

$$|u_h - u|_{H^1} \leq C \sqrt{\frac{\Lambda}{\lambda}} h^{\hat{m}-1} \left[ |u|_{H^{\hat{m}}} + \max \{ |u|_{W^{1,2\hat{m}}}^{\hat{m}}, |u|_{W^{1,2\hat{m}}} \} \right],$$

where the constant  $C$  depends on  $P$ . To show that  $u_h$  is a local minimizer in  $V_h \cap H_\varphi^1$ , let  $k := \frac{2s}{s-2(\hat{m}-1)}$  if  $2(\hat{m}-1) < s$ , and arbitrarily large otherwise. By assumption  $\max\{s, 2\} < q < k$ . Thus, we can set  $\mu := (2^{-1} - q^{-1})(q^{-1} - k^{-1})^{-1}$ ,  $\delta := (1 + \mu)^{-1}(\hat{m} - 1 - s(2^{-1} - q^{-1})) > 0$ , and use  $L^p$ -interpolation, the  $H^1$  bound (12), the Sobolev embedding theorem, and Theorem 11 to estimate

$$\begin{aligned} |u_h - u|_{W^{1,q}} &\leq h^{-(\hat{m}-1)+\delta} |u - u_h|_{H^1} + h^{\mu^{-1}(\hat{m}-1-\delta)} |u - u_h|_{W^{1,k}} \\ &\leq Ch^\delta (\|u\|_{H^{\hat{m}}} + \|u\|_{H^{\hat{m}}}^{\hat{m}}) + h^{\mu^{-1}(\hat{m}-1-\delta)} (|u|_{W^{1,k}} + |u_h|_{W^{1,k}}) \\ &\leq Ch^\delta (\|u\|_{H^{\hat{m}}} + \|u\|_{H^{\hat{m}}}^{\hat{m}} + K). \end{aligned}$$

As  $\|u_h - u\|_{L^\infty} \leq C|u_h - u|_{W^{1,q}} \leq C(u)h^\delta$ , we can choose  $h$  small enough such that  $\|du_h\|_{L^q} < K$  and  $\|u_h - u\|_{L^\infty} < \epsilon$ , so that  $u_h$  does not lie on the boundary of  $V_{h;K} \cap B_\epsilon^\infty(u)$ , and is indeed a local minimizer in  $V_h \cap H_\varphi^1(\Omega, M)$ .  $\square$

**3.3. Discretization error estimates in  $L^2$ .** To obtain optimal  $L^2$ -discretization error estimates we apply [27, Thm. 2.13], which is a generalization of the Aubin–Nitsche lemma [14, Thm. 3.2.5]. With slightly adapted notation, it states the following.

**THEOREM 17.** *Let  $m > \frac{s}{2}$  and assume that  $u \in H_\varphi^m(\Omega, M) := H^m(\Omega, M) \cap H_\varphi^1(\Omega, M)$  is a minimizer of an elliptic, semilinear energy  $E : H_\varphi^m(\Omega, M) \rightarrow \mathbb{R}$  that has an  $H^2$ -regular dual problem, i.e., for all  $\mathbf{g} \in L^2(\Omega, u^{-1}TM)$  there exists a solution  $\mathbf{w} \in H^2(\Omega, u^{-1}TM)$  to*

$$\mathbf{w} \in H_0^1(\Omega, u^{-1}TM), \quad \delta^2 E(u)(\mathbf{w}, \mathbf{v}) = -(\mathbf{g}, \mathbf{v}) \quad \forall \mathbf{v} \in H_0^1(\Omega, u^{-1}TM)$$

with

$$\|\mathbf{w}\|_{W^{2,2}} \leq C \|\mathbf{g}\|_{L^2}.$$

For a given shape-regular, affine-equivalent, quasi-uniform grid  $\mathcal{G}$ , let  $S_h \subset H_\varphi^1(\Omega, M)$  be a discrete approximation space such that for all  $v \in H_\varphi^1 \cap C(\Omega, M; \rho)$  with  $v|_{T_h} \in H^m(T_h, M)$  for all  $T_h$ , there exists an approximating map  $v_I \in S_h$  with

$$(13) \quad |v_I|_{H^m} + |v_I|_{W^{1,2m}}^m \leq C (|v|_{H^m} + |v|_{W^{1,2m}}^m)$$

that fulfills the estimate

$$(14) \quad \|v - v_I\|_{L^2} + h |v - v_I|_{W^{1,2}} \leq C(v) h^l$$

for  $l \geq 2$ . Assume further that each discrete map  $v_h \in S_h \cap C(\Omega, M; \rho)$  fulfills inverse estimates of the form

$$(15) \quad |v_h|_{W^{1,p}} \leq C h^{-s \max\{0, \frac{1}{q} - \frac{1}{p}\}} |v_h|_{W^{1,q}}$$

for  $q > \max\{2, s\}$ . Finally, assume that for each vector field  $\mathbf{v} \in C(\Omega, v_h^{-1}TM)$  along a discrete function  $v_h \in S_h$  that is in  $W^{2,2}(T_h, v_h^{-1}TM)$  for each element  $T_h$ , there exists a variation  $\mathbf{v}_I$  of maps in  $S_h$  such that

$$(16) \quad \|\mathbf{v} - \mathbf{v}_I\|_{W^{1,2}} \leq C(v_h, \mathbf{v})h.$$

Then if the discrete minimizer

$$u_h := \arg \min_{v_h \in S_h} E(v_h)$$

fulfills the a priori error estimate

$$(17) \quad |u - u_h|_{W^{1,2}} \leq C(u)h^{l-1},$$

and on each element  $T_h$  the estimate

$$(18) \quad \|u_h\|_{W^{2,p_2}(T_h)}^{p_2} + \|u_h\|_{W^{1,2p_2}(T_h)}^{2p_2} \leq K_2^{p_2}$$

for a constant  $K_2$  and  $p_2 = \max\{2, \frac{s}{2}\}$  for  $s \neq 4$  (and  $p_2 > 2$  for  $s = 4$ ), we get

$$\|u - u_h\|_{L^2} \leq C(u)h^l.$$

The main difference to the Euclidean Aubin–Nitsche lemma is that in order to compare test vector fields along  $u$  and  $u_h$ , one needs to be transported into the space of the other along a suitable connecting curve. This transportation needs to preserve  $H^2$ -norms of vector fields. In [27] this preservation is proven for the case that the functions  $u$  and  $u_h$  have bounded grid-dependent  $H^2$ -norm. This leads to the additional assumption (18).

For technical reasons, we restrict ourselves to the practically relevant case  $s < 4$ . Other dimensions may be dealt with similarly as discussed in [27].

**THEOREM 18.** *Consider the setting of Theorem 16 with  $m \geq 2$ ,  $s < 4$ , and  $q \geq 4$ , including the assumptions for  $u_h$  to be a local minimizer in  $V_h \cap H_\varphi^1(\Omega, M)$ . Then*

$$\|u - u_h\|_{L^2} \leq C h^{\min(m,r+1)}.$$

*Proof.* We prove the assertion by verifying the assumptions of Theorem 17. First note that the harmonic map energy is indeed elliptic (Proposition 15), is semilinear, and has an  $H^2$ -regular dual problem [25].

The approximation error estimates for maps (13) and (14) are provided by Proposition 9 and Theorem 10, respectively. The inverse estimate (15) is given in Theorem 11. The interpolation error bound (16) for vector fields follows from Theorem 12. The  $H^1$  a priori bound (17) is proven in Theorem 16.

Thus, all that is left to show is (18), i.e., for a solution  $u_h$  of (11) we need to show that there exists a constant  $K_2$  such that the grid-dependent homogeneous  $W^{2,2} \cap W^{1,4}$ -norm of  $u_h$  is bounded, i.e.,

$$(19) \quad \int_\Omega |du_h|^4 dx + \sum_{T_h \in \mathcal{G}} \int_{T_h} |\nabla^2 u_h|^2 dx \leq K_2^2.$$

The boundedness of the first integral even without the elementwise partition follows because the assumptions  $q \geq 4$  and  $u_h \in V_{h,K} \cap B_\epsilon^\infty(u)$  imply  $\|du_h\|_{L^4} \leq C K$ .

Set  $\hat{u}_h = Q_{\mathbb{R}^n} u_h$ , and  $\hat{u}_I = Q_{\mathbb{R}^n} u$ . As the continuity of  $Q_M$  implies that  $\|\nabla^2 Q_M u\|_{L^2(\mathcal{G})}$  is bounded, by the triangle inequality it is enough to obtain an estimate on  $\|\nabla^2 u_h - \nabla^2 Q_M u\|_{L^2(\mathcal{G})}$ . By the chain rule we have

$$\begin{aligned} \|\nabla^2 u_h - \nabla^2 Q_M u\|_{L^2} &\leq \|\mathcal{P}'(\hat{u}_h)(\nabla^2 \hat{u}_h) - \mathcal{P}'(\hat{u}_I)(\nabla^2 \hat{u}_I)\|_{L^2} \\ &\quad + \|\mathcal{P}''(\hat{u}_h)(d\hat{u}_h, d\hat{u}_h) - \mathcal{P}''(\hat{u}_I)(d\hat{u}_I, d\hat{u}_I)\|_{L^2} \\ &\leq \|\mathcal{P}'\|_{L^\infty} \|\nabla^2 \hat{u}_h - \nabla^2 \hat{u}_I\|_{L^2} \\ &\quad + \text{Lip}(\mathcal{P}') \|\nabla^2 \hat{u}_I\|_{L^2} \|\hat{u}_h - \hat{u}_I\|_{L^\infty} \\ &\quad + \|\mathcal{P}''\|_{L^\infty} (\|d\hat{u}_h\|_{L^4}^2 + \|d\hat{u}_I\|_{L^4}^2). \end{aligned}$$

By (3), the  $W^{1,4}$ -bounds on  $u_h$  and  $Q_M u$  transfer to  $\hat{u}_h$  and  $\hat{u}_I$ , and  $\|\nabla^2 \hat{u}_I\|_{L^2}$  is bounded. Further, we choose an exponent  $a = a(s)$  such that  $W^{1,2} \hookrightarrow L^a$  which exists by the Sobolev embedding theorem. Then by Theorem 8 we have

$$\|\nabla^2 \hat{u}_h - \nabla^2 \hat{u}_I\|_{L^2} + \|\hat{u}_h - \hat{u}_I\|_{L^\infty} \leq C(h^{-1} + h^{-\frac{s}{a}}) \|\hat{u}_h - \hat{u}_I\|_{W^{1,2}}.$$

Thus, by (3) and Theorem 16, we have

$$\|\nabla^2 \hat{u}_h - \nabla^2 \hat{u}_I\|_{L^2} + \|\hat{u}_h - \hat{u}_I\|_{L^\infty} \leq C h^{\min\{m-1, r\} - \max\{1, \frac{s}{a}\}}.$$

By the Sobolev embedding theorem, for  $s \leq 4$  we can even choose the  $a$  such that  $a \geq s$ . Then  $\max\{1, \frac{s}{a}\} = 1$ , and from  $m \geq 2$  it follows that  $\min\{m-1, r\} \geq 1$ . Thus,  $\|\nabla^2 u_h\|_{L^2}$  can be bounded independently of  $h$ , and we obtain (19).  $\square$

**4. Numerical interpolation error tests.** We now show numerically that the optimal interpolation error orders predicted by Theorem 10 can really be observed in practice. We test this for maps into the unit sphere  $S^2$  and into  $\text{SO}(3)$ . All algorithms are implemented in C++ using the DUNE libraries [9].

**4.1. Maps into the unit sphere.** Our first example measures the  $L^2$  and  $H^1$  interpolation errors for maps into the unit sphere  $S^2 \subset \mathbb{R}^3$ . As the domain we use the square  $\Omega = (-5, 5)^2$ , and we measure the error of interpolating the inverse stereographic projection

$$(20) \quad p_{\text{st}} : \mathbb{R}^2 \rightarrow S^2, \quad p_{\text{st}}(x) := \left( \frac{2x_0}{|x|^2 + 1}, \frac{2x_1}{|x|^2 + 1}, \frac{|x|^2 - 1}{|x|^2 + 1} \right)^T,$$

restricted to  $\Omega$ . This function is in  $C^\infty$ , and we can therefore hope for optimal interpolation error orders.

We discretize the domain with the grid shown in Figure 1. Observe that it combines triangles and nonaffine quadrilateral elements. It is therefore slightly beyond the assumptions of Theorem 10. We create a sequence of grids by refining the initial grid uniformly up to six times. On each grid we compute  $Q_M p_{\text{st}}$  using projection-based finite elements of orders  $p = 1, 2, 3$ , and we measure the error  $\|Q_M p_{\text{st}} - p_{\text{st}}\|$  both in the  $L^2(\Omega, \mathbb{R}^3)$ -norm and the  $H^1(\Omega, \mathbb{R}^3)$ -seminorm. Sixth-order Gaussian quadrature rules are used for the integrals, but note that since projection-based finite element functions are not piecewise polynomials in  $\mathbb{R}^3$ , a small additional error due to numerical quadrature remains.

The results are plotted in Figure 2(a). As expected, the errors decay like  $h^{p+1}$  for the  $L^2$ -norm and like  $h^p$  for the  $H^1$ -seminorm. These are the optimal orders predicted by the theory in Chapter 2.



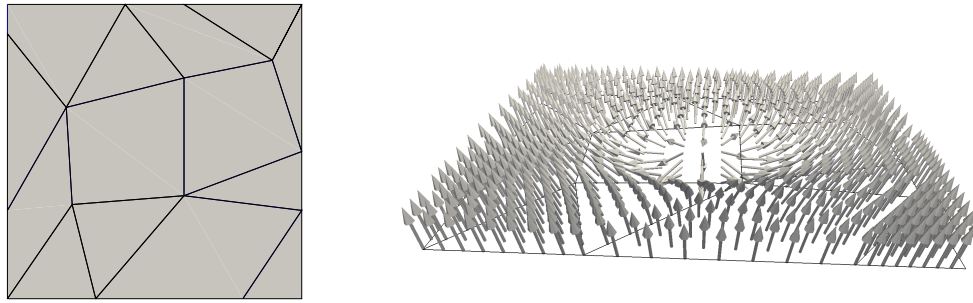


FIG. 1. Interpolating the inverse stereographic projection  $p_{st} : \mathbb{R}^2 \rightarrow S^2$ . Left: coarsest grid. Right: function values.

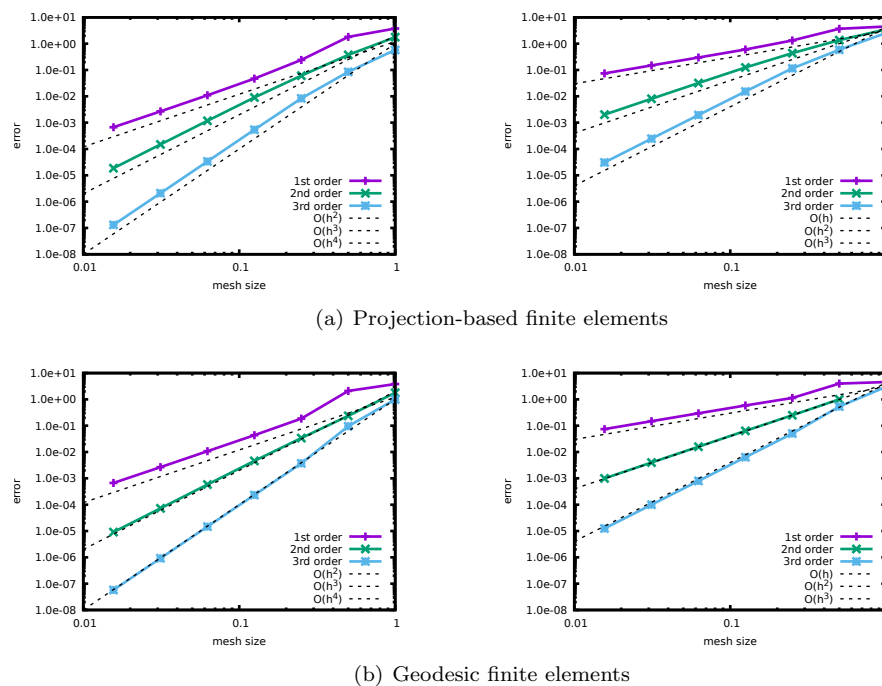


FIG. 2. Interpolation errors for the inverse stereographic projection. Left:  $L^2$ -norm; right:  $H^1$ -seminorm. The black dashed lines are at the same positions for both discretizations.

We now compare the projection-based discretization to a discretization using geodesic finite elements. The resulting errors per mesh size are shown in Figure 2(b). One can see that the same asymptotic orders are obtained, as predicted by the interpolation theory for geodesic finite elements [24, 26]. However, the constant is consistently better for geodesic finite elements for orders  $p = 2$  and  $p = 3$ . Supposedly, the reason for this is that the intrinsic construction of geodesic finite elements captures the geometry of  $S^2$  better.

**4.2. Maps into the special orthogonal group.** We repeat the experiment for a map into  $SO(3)$ . As shown in [39], the closest-point projection of a general matrix

onto  $O(3)$  is the orthogonal factor of the polar decomposition.<sup>1</sup> As such, it is defined for all  $A \in \mathbb{R}^{3 \times 3}$ , and it is unique if  $A$  is invertible. Furthermore, the polar factor will be in  $SO(3)$  if  $\det A > 0$ . In particular, this will be the case for all  $A$  close enough to  $SO(3)$ .

To numerically compute the polar factor of  $A$  we use the iteration defined by  $Q_0 := A$  and

$$(21) \quad Q_{k+1} := \frac{1}{2} (Q_k + Q_k^{-T}),$$

which is related to Newton's method for computing the square root of 1. Higham [29] showed that this iteration converges quadratically to the polar factor of  $A$ . To compute  $H^1$  norms of projection-based finite elements we also need the derivative of the polar factor with respect to  $A$ . Following [19], we use the iterative algorithm that results from differentiating (21).<sup>2</sup>

**5. Numerical discretization error tests for harmonic maps.** In this second section of numerical results we present measurements of the discretization errors of harmonic maps into  $S^2$  and  $SO(3)$ . These confirm the theoretical predictions of section 3. In addition, we again compare the results to geodesic finite elements.

For the domain of the example we use  $\Omega = (-5, 5)^2$  of the previous section, and we also reuse the grid from Figure 1. We will interpolate the function  $R : \Omega \rightarrow SO(3)$ ,

$$(22) \quad R(x) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\frac{\pi}{5}x_0) & -\sin(\frac{\pi}{5}x_0) \\ 0 & \sin(\frac{\pi}{5}x_0) & \cos(\frac{\pi}{5}x_0) \end{pmatrix} \begin{pmatrix} \cos(\frac{\pi}{5}x_1) & 0 & -\sin(\frac{\pi}{5}x_1) \\ 0 & 1 & 0 \\ \sin(\frac{\pi}{5}x_1) & 0 & \cos(\frac{\pi}{5}x_1) \end{pmatrix}.$$

Again we measure the errors in the  $L^2$ -norm and the  $H^1$ -seminorm on a set of seven grids obtained by uniform refinement. Interpolation errors for this scenario are plotted in Figure 3(a). As expected, we see the same optimal orders as for the case of mapping into the sphere. Figure 3(b) shows the corresponding errors obtained using geodesic finite elements. One observes the same effect as before: the convergence rates are the same, but the constant is lower. In fact, this effect is now more pronounced than before and seems to increase with the order  $p$ .

**5.1. Harmonic maps into the sphere  $S^2$ .** The example builds on top of the interpolation error measurements of section 4.1. Reusing the domain  $\Omega = (-5, 5)^2$  and grids from there, we compute harmonic maps from  $\Omega$  into  $S^2$  that take the values given by the inverse stereographic projection function  $p_{st}$  defined in (20) on the boundary of  $\Omega$  in spaces of projection-based finite elements of orders  $p = 1, 2, 3$  mapping into  $S^2$ . It is shown by [10] (see also [35]) that the inverse stereographic projection  $p_{st}$  itself minimizes the harmonic energy (7) in the first nontrivial homotopy group in  $H^1(\mathbb{R}^2, S^2)$ . This function is  $C^\infty$  and we can therefore hope for optimal discretization error orders.

Using the canonical embedding of  $S^2$  into  $\mathbb{R}^3$ , and the metric on  $S^2$  induced by the embedding, the integrand  $|dv|^2$  of (7) has the coordinate representation

$$|dv|^2 = \sum_{i=1}^s \sum_{a=1}^3 \left( \frac{\partial v^a}{\partial x^i} \right)^2,$$

<sup>1</sup>See section 5.3 for an alternative construction using quaternions.

<sup>2</sup>In a recent paper, [30] proposed a different method for computing the polar factor of a  $3 \times 3$  matrix, which they claim is faster than Newton iteration [30]. It is more complicated than the method used here, but for future research it appears to be an interesting alternative.

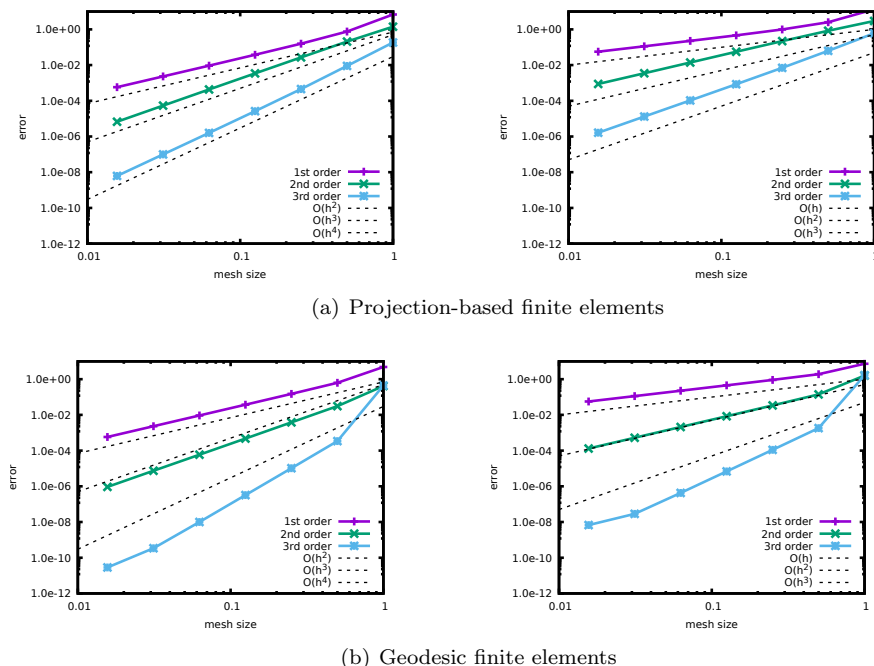


FIG. 3. Interpolation errors for a map into  $SO(3)$ . Left:  $L^2$ -norm; right:  $H^1$ -seminorm. The black dashed lines are at the same positions for both discretizations.

that is,  $dv$  is a  $3 \times s$ -matrix and  $|\cdot|$  the Frobenius norm. We compute minimizers of the discrete energy using the approach proposed in [42] for geodesic finite elements. Identifying discrete functions with sets of coefficients in  $S^2$ , we obtain an algebraic minimization problem posed on the nonlinear manifold  $(S^2)^N$ , where  $N$  is the number of Lagrange nodes on the grid. This minimization problem is solved using the Riemannian trust-region method introduced by [2] together with the inner monotone multigrid solver described in [42]. Gradient and Hessian of the energy functional are computed using the ADOL-C automatic differentiation software [49], and the formula derived in [3] to obtain the Riemannian Hessian from the Euclidean Hessian.

The Riemannian trust-region solver is set to iterate until the maximum norm of the correction drops below  $10^{-6}$ . We then compute solutions  $(p_{st})_p^k$ ,  $k = 0, \dots, 6$ ,  $p = 1, 2, 3$ , on the grids obtained by  $k$  steps of uniform refinement, and compute the errors

$$e_p^k = \|v_p^k - p_{st}\|, \quad k = 0, \dots, 6, \quad p = 1, 2, 3,$$

where  $\|\cdot\|$  is either the norm in  $L^2(\Omega, \mathbb{R}^3)$ , or the half norm in  $H^1(\Omega, \mathbb{R}^3)$ . Figure 4(a) shows the errors  $e_p^k$  as functions of the normalized mesh size  $h$ . We see that for  $p$ th-order finite elements the  $L^2$ -error decreases like  $h^{p+1}$ , and the  $H^1$ -error decreases like  $h^p$ . Hence we can reproduce the optimal convergence behavior predicted by Theorems 16 and 18.

We now compare the projection-based discretization to a discretization using geodesic finite elements. The resulting errors per mesh size are shown in Figure 4(b). One can see that the same asymptotic orders are obtained, as predicted by theory [24, 27]. As in the interpolation case, the constant is slightly better for geodesic finite

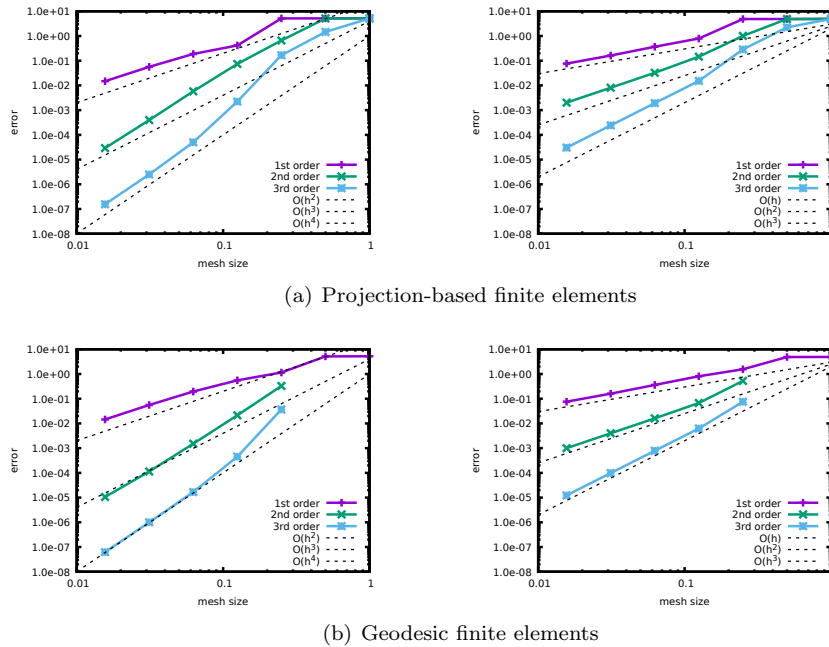


FIG. 4. Discretization errors for a harmonic map into  $S^2$  as a function of normalized grid edge length. Left:  $L^2$ -norm. Right:  $H^1$ -seminorm. The black dashed lines are at the same positions for both discretizations.

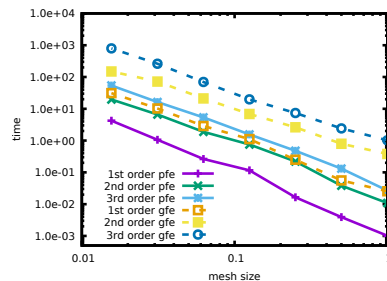


FIG. 5. Wall-time needed to compute the harmonic energy on seven different grids and approximation orders  $p = 1, 2, 3$ . Solid: projection-based finite elements. Dashed: geodesic finite elements.

elements. On the other hand, one can see that the graphs in Figure 4(b) do not contain values for the two coarsest grids and approximation orders 2 and 3. This is because the minimization problem that defines geodesic interpolation was actually ill-defined on at least one grid element in these cases. The problem does not happen for projection-based finite elements for this example.

The decisive argument for projection-based finite elements for this scenario, however, is run-time. Figure 5 plots the total time needed to compute the harmonic energy for the different finite element spaces and grid resolutions. Projection-based finite elements need only about 10% of the time of geodesic finite elements. This is of course because projection-based interpolation is given by a simple closed-form formula in the case of  $M = S^2$ , whereas for geodesic finite elements it involves numerically solving a small minimization problem (2). By means of automatic differentiation,

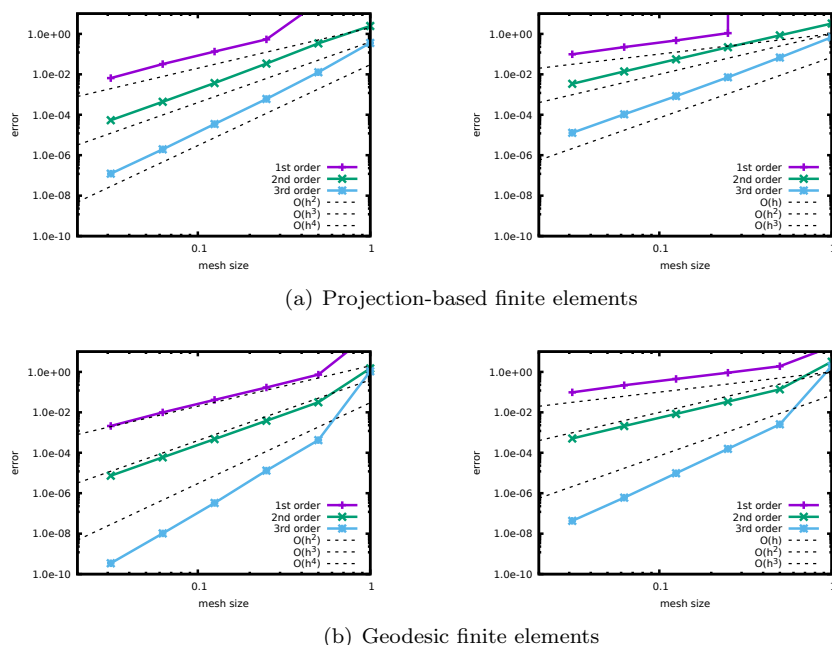


FIG. 6. Discretization errors for a harmonic map into  $SO(3)$  as a function of normalized grid edge length. Left:  $L^2$ -norm. Right:  $H^1$ -seminorm. The black dashed lines are at the same positions for both discretizations.

these differences appear in the computation of derivatives as well. In practical applications of sphere-valued problems, projection-based finite elements are therefore typically preferable to geodesic finite elements.

**5.2. Harmonic maps into  $SO(3)$ .** In the final example we compute minimizers of the harmonic energy in a space of functions mapping to  $SO(3)$ . As in section 4.2, we use the canonical embedding of  $SO(3)$  into  $\mathbb{R}^{3 \times 3}$  and the polar factor as the projection onto  $SO(3)$  (even though the implementation uses quaternions to actually store elements of  $SO(3)$ ). The iteration (21) used to compute the polar factor is a variant of a Newton method and therefore plays nicely with automatic differentiation systems like ADOL-C [22, Chap. 15].

We base our numerical test on the interpolation error test of section 4.2. On the domain  $\Omega$  given there, we look for minimizers of the harmonic energy in  $H^1(\Omega, SO(3))$ , subject to Dirichlet boundary and homotopy constraints given by the function  $R$  defined in (22). As the solution of this problem is not known in closed form, we compute discretization errors with respect to a numerical reference solution. For this we refine the grid uniformly six times and compute the solution there. We then trust this to be a good reference solution for grids with up to five steps of refinement.

Figure 6(a) shows the discretization error plots for the  $L^2$ - and  $H^1$ -errors, again for approximation spaces of orders up to 3. We observe the expected optimal discretization error rates in all cases.

Finally, we redo the experiment with geodesic finite elements. Figure 6(b) shows the discretization errors per mesh size for the same problem, but using a geodesic finite element discretization. The constant is again better than for projection-based finite elements, and the difference seems to increase with  $p$ .

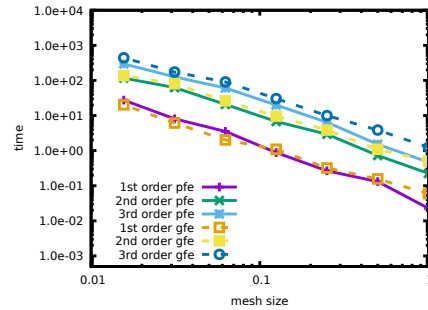


FIG. 7. Wall-time needed to compute the harmonic energy of the function  $R$  on seven different grids and approximation orders  $p = 1, 2, 3$ . Solid: projection-based finite elements. Dashed: geodesic finite elements.

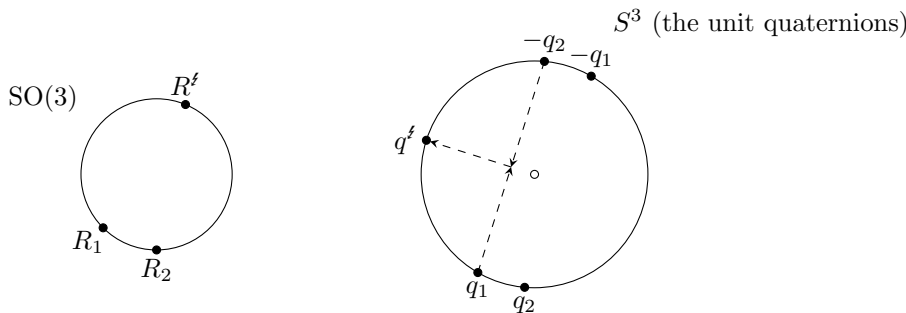


FIG. 8. How using quaternions to interpolate rotations can fail. The aim in this example is to average two rotations  $R_1$  and  $R_2$ . They are represented by the pairs of unit quaternions  $\pm q_1$  and  $\pm q_2$ , respectively. The point  $q^f$  results from projected interpolation between  $q_1$  and  $-q_2$ . The corresponding rotation is  $R^f$ , which is nowhere near the expected average.

When comparing the run-times again (Figure 7), the situation is vastly different. While projection-based finite elements were much faster than geodesic ones for the case of maps into the sphere, there is hardly a difference for  $M = \text{SO}(3)$ . This is because the projection from  $\mathbb{R}^{3 \times 3}$  to  $\text{SO}(3)$  is not given in a simple closed form but has to be computed iteratively (section 4.2). This puts the execution speed of projection-based finite elements on par with geodesic finite elements, if the projection is the polar decomposition in  $\mathbb{R}^{3 \times 3}$ .

**5.3. Interpolation in quaternion space.** It is tempting to replace the costly polar decomposition for the projection onto  $\text{SO}(3)$  by interpolation in four-dimensional quaternion space  $\mathbb{R}^4$  and projecting onto the unit quaternions  $S^3$ . Indeed, the unit quaternions are widely used for storing rotations, and the projection from  $\mathbb{R}^4$  onto the unit quaternions is simple and very cheap to compute.

However, in our experience this leads to severe numerical artifacts. The underlying cause is the fact that rotation matrices are represented by *pairs* of unit quaternions. When interpolating between  $m$  rotations, an implementation has to pick a representative from each of the  $m$  pairs, but this choice can influence the interpolation result. We have tried to illustrate this graphically for the case  $m = 2$  in Figure 8. There,  $R_1$  and  $R_2$  are two points on  $\text{SO}(3)$ , and the aim is to compute their average. Their quaternion representatives are  $\{q_1, -q_1\} \in S^3$  and  $\{q_2, -q_2\} \in S^3$ , respectively. If

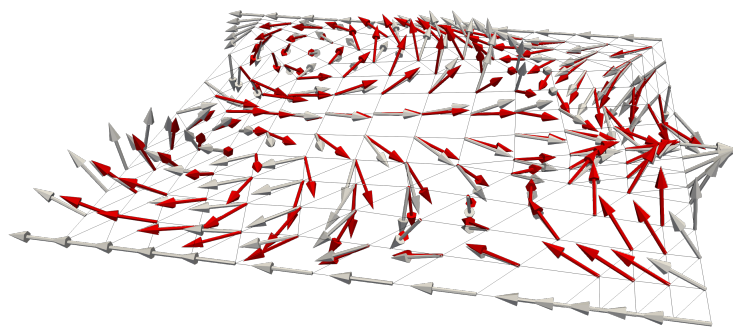


FIG. 9. Comparison of a discrete harmonic minimizers computed using polar decomposition (red) and quaternion interpolation (gray). The arrows are the first columns of the discrete rotation matrix fields.

$R_1$  and  $R_2$  are close to each other, then (say) the representatives  $q_1$  and  $q_2$  will be close to each other in  $S^3$ . Interpolating between  $q_1$  and  $q_2$  and projection onto  $S^3$  will produce a rotation “between”  $R_1$  and  $R_2$ , which is the expected outcome. The same happens when using the representatives  $-q_1$  and  $-q_2$  to interpolate.

However, an implementation may as well use the pairs  $q_1, -q_2$  or  $-q_1, q_2$  for interpolation. These, however, are far apart and will lead to interpolated rotations  $R^f$  that are far away from both  $R_1$  and  $R_2$ . For  $m = 2$ , simple strategies can be devised that always select suitable sets of representatives, but such strategies quickly become unfeasible for larger  $m$ .

These problems are not academic. We repeated the complete set of tests of section 5.2 and found a completely different numerical behavior. Not only did the trust-region solver need many more iterations, but also the computed solution had vastly higher energy levels than the one using the polar decomposition (ranging from 30 % up to a 10-fold increase). Exemplarily, Figure 9 shows the first directors (i.e., the first matrix columns) of the minimizer of the harmonic energy computed using the polar decomposition (red) and quaternion interpolation (gray). They are both similar, but certainly not identical. The energy at the minimizer is 19.161 for the quaternions and 10.059 for the polar factor.

For one-dimensional  $C^1$  finite elements, [20] compared the two interpolation approaches and reported smaller interpolation errors for the quaternion-based method. They did not mention any such stability problems, presumably because they consider one-dimensional domains only.

#### REFERENCES

- [1] T. J. ABATZOGLOU, *The minimum norm projection on  $C^2$ -manifolds in  $\mathbb{R}^n$* , Trans. Amer. Math. Soc., 243 (1978), pp. 115–122.
- [2] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [3] P.-A. ABSIL, R. MAHONY, AND J. TRUMPF, *An extrinsic look at the Riemannian Hessian*, in Geometric Science of Information, Lecture Notes in Comput. Sci. 8085, Springer, New York, 2013, pp. 361–368.
- [4] F. ALOUGES, *A new algorithm for computing liquid crystal stable configurations: The harmonic mapping case*, SIAM J. Numer. Anal., 34 (1997), pp. 1708–1726.
- [5] F. ALOUGES, E. KRITSIKIS, AND J.-C. TOUSSAINT, *A convergent finite element approximation for Landau–Lifschitz–Gilbert equation*, Phys. B, 407 (2012), pp. 1345–1349.

- [6] M. BACÁK, R. BERGMANN, G. STEIDL, AND A. WEINMANN, *A second order nonsmooth variational model for restoring manifold-valued images*, SIAM J. Sci. Comput., 38 (2016), pp. A567–A597.
- [7] S. BARTELS, *Numerical analysis of a finite element scheme for the approximation of harmonic maps into surfaces*, Math. Comp., 79 (2010), pp. 1263–1301.
- [8] S. BARTELS AND A. PROHL, *Constraint preserving implicit finite element discretization of harmonic map flow into spheres*, Math. Comp., 76 (2007), pp. 1847–1859.
- [9] P. BASTIAN, M. BLATT, A. DEDNER, C. ENGWER, R. KLÖFKORN, R. KORNUBER, M. OHLBERGER, AND O. SANDER, *A generic grid interface for parallel and adaptive scientific computing. Part II: Implementation and tests in DUNE*, Computing, 82 (2008), pp. 121–138.
- [10] A. BELAVIN AND A. POLYAKOV, *Metastable states of two-dimensional isotropic ferromagnets*, J. Experimental Theoret. Phys. Lett., 22 (1975), pp. 245–247.
- [11] G. BOURDAUD, *Superposition in homogeneous and vector valued Sobolev spaces*, Trans. Amer. Math. Soc., 362 (2010), pp. 6105–6130.
- [12] D. BRAESS, *Finite Elemente-Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*, Springer, Berlin, 2007.
- [13] S. R. BUSS AND J. P. FILLMORE, *Spherical averages and applications to spherical splines and interpolation*, ACM Trans. Graphics, 20 (2001), pp. 95–126.
- [14] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, Elsevier, Amsterdam, 1978.
- [15] 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. A, 455 (1999), pp. 1125–1147.
- [16] W. DAHMEN, B. FAERMANN, I. GRAHAM, W. HACKBUSCH, AND S. SAUTER, *Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method*, Math. Comp., 73 (2004), pp. 1107–1138.
- [17] A. DESIMONE, R. KOHN, S. MÜLLER, AND F. OTTO, *Recent analytical developments in micromagnetics*, in The Science of Hysteresis. Vol. 2 Physical Modeling, Micromagnetics, and Magnetization Dynamics, Elsevier, Amsterdam, 2006, pp. 269–381.
- [18] J. EELS AND L. LEMAIRE, *A report on harmonic maps*, Bull. Lond. Math. Soc., 10 (1978), pp. 1–68.
- [19] E. S. GAWLIK AND M. LEOK, *Iterative computation of the Fréchet derivative of the polar decomposition*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1354–1379.
- [20] E. S. GAWLIK AND M. LEOK, *Embedding-based interpolation on the special orthogonal group*, SIAM J. Sci. Comput., 40 (2018), pp. A721–A746, <https://doi.org/10.1137/17M1129416>.
- [21] I. G. GRAHAM, W. HACKBUSCH, AND S. A. SAUTER, *Finite elements on degenerate meshes: Inverse-type inequalities and applications*, IMA J. Numer. Anal., 25 (2005), pp. 379–407.
- [22] A. GRIEWANK AND A. WALTHER, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, 2nd ed., SIAM, Philadelphia, 2008.
- [23] P. GROHS, *Finite elements of arbitrary order and quasiinterpolation for Riemannian data*, IMA J. Numer. Anal., 33 (2013), pp. 849–874.
- [24] P. GROHS, H. HARDER, AND O. SANDER, *Optimal a priori discretization error bounds for geodesic finite elements*, Found. Comput. Math., 15 (2015), pp. 1357–1411, <https://doi.org/10.1007/s10208-014-9230-z>.
- [25] H. HARDER, *Intrinsic Discretization Error Bounds for Geodesic Finite Elements*, Ph.D. thesis, Freie Universität Berlin, 2015.
- [26] H. HARDER,  *$L^2$ -Discretization Error Bounds for Maps into Riemannian Manifolds*, preprint, arXiv:1612.06086, 2016.
- [27] H. HARDER,  *$L^2$ -discretization error bounds for maps into Riemannian manifolds*, Numer. Math., (2018), <https://doi.org/10.1007/s00211-017-0941-3>.
- [28] F. HÉLEIN AND J. C. WOOD, *Harmonic maps*, in Handbook of Global Analysis, Elsevier, Amsterdam, 2008, pp. 417–491.
- [29] N. J. HIGHAM, *Computing the polar decomposition with applications*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 1160–1174.
- [30] N. J. HIGHAM AND V. NOFERINI, *An algorithm to compute the polar decomposition of a  $3 \times 3$  matrix*, Numer. Algorithms, 73 (2016), pp. 349–369, <https://doi.org/10.1007/s11075-016-0098-7>.
- [31] E. KRITSIKIS, A. VAYSET, L. BUDA-PREJBEANU, F. ALOUGES, AND J.-C. TOUSSAINT, *Beyond first-order finite element schemes in micromagnetics*, J. Comput. Phys., 256 (2014), pp. 357–366.
- [32] J. M. LEE, *Introduction to Smooth Manifolds*, Springer, New York, 2003.



- [33] S.-Y. LIN AND M. LUSKIN, *Relaxation methods for liquid crystal problems*, SIAM J. Numer. Anal., 26 (1989), pp. 1310–1324.
- [34] S. MAZUR AND S. ULAM, *Sur les transformations isométriques d'espaces vectoriels normés*, C. R. Acad. Sci. Paris, 194 (1932), pp. 946–948.
- [35] C. MELCHER, *Chiral skyrmions in the plane*, Proc. R. Soc. A, 470 (2014), <https://doi.org/10.1098/rspa.2014.0394>.
- [36] W. MÜLLER, *Numerische Analyse und Parallele Simulation von nichtlinearen Cosserat-Modellen*, Ph.D. thesis, Karlsruher Institut für Technologie, 2009.
- [37] I. MÜNCH, *Ein geometrisch und materiell nichtlineares Cosserat-Modell—Theorie, Numerik und Anwendungsmöglichkeiten*, Ph.D. thesis, Universität Karlsruhe, 2007.
- [38] P. NEFF, *A geometrically exact Cosserat shell-model including size effects, avoiding degeneracy in the thin shell limit. Existence of minimizers for zero Cosserat couple modulus*, Math. Models Methods Appl. Sci., 17 (2007), pp. 363–392.
- [39] P. NEFF, J. LANKEIT, AND A. MADEO, *On Grioli's minimum property and its relation to Cauchy's polar decomposition*, Int. J. Engrg. Sci. (2014).
- [40] X. PENNEC, P. FILLARD, AND N. AYACHE, *A Riemannian framework for tensor computing*, Int. J. Comput. Vis., 66 (2006), pp. 41–66.
- [41] U. RAHMAN, I. DRORI, V. C. STODDEN, D. L. DONOHO, AND P. SCHRÖDER, *Multiscale representations for manifold-valued data*, Multiscale Model. Simul., 4 (2005), pp. 1201–1232.
- [42] O. SANDER, *Geodesic finite elements on simplicial grids*, Internat. J. Numer. Methods Engrg., 92 (2012), pp. 999–1025.
- [43] O. SANDER, *Geodesic finite elements of higher order*, IMA J. Numer. Anal., 36 (2016), pp. 238–266.
- [44] O. SANDER, *Test Function Spaces for Geometric Finite Elements*, preprint, arXiv:1607.07479, 2016.
- [45] O. SANDER, P. NEFF, AND M. BİRSAN, *Numerical treatment of a geometrically nonlinear planar Cosserat shell model*, Comput. Mech., 57 (2016), pp. 817–841.
- [46] J. SIMO, D. FOX, AND M. RIFAI, *On a stress resultant geometrically exact shell model. Part III: Computational aspects of the nonlinear theory*, Comput. Methods Appl. Mech. Engrg., 79 (1990), pp. 21–70.
- [47] J. SIMO AND L. VU-QUOC, *A three-dimensional finite-strain rod model. Part II: Computational aspects*, Comput. Methods Appl. Mech. Engrg., 58 (1986), pp. 79–116.
- [48] M. SPRECHER, *Numerical Methods for Optimization and Variational Problems with Manifold-Valued Data*, Ph.D. thesis, ETH Zürich, 2016.
- [49] A. WALTHER AND A. GRIEWANK, *Getting started with ADOL-C*, in Combinatorial Scientific Computing, U. Naumann and O. Schenk, eds., Chapman & Hall/CRC Press, Boca Raton, FL, 2012, pp. 181–202.