Optimal A Priori Discretization Error Bounds for Geodesic Finite Elements

Philipp Grohs, Hanne Hardering and
Oliver Sander

Optimal A Priori Discretization Error Bounds for Geodesic Finite Elements

Philipp Grohs, Hanne Hardering and Oliver Sander

May 14, 2013

Abstract

We prove optimal bounds for the discretization error of geodesic finite elements for variational partial differential equations for functions that map into a nonlinear space. For this we first generalize the well-known Céa lemma to nonlinear function spaces. In a second step we prove optimal interpolation error estimates for pointwise interpolation by geodesic finite elements of arbitrary order. These two results are both of independent interest. Together they yield optimal a priori error estimates for a large class of manifold-valued variational problems. We measure the discretization error both intrinsically using an H^1-type Finsler norm, and with the H^1-norm using embeddings of the codomain in a linear space. To measure the regularity of the solution we propose a nonstandard smoothness descriptor for manifold-valued functions, which bounds additional terms not captured by Sobolev norms. As an application we obtain optimal a priori error estimates for discretizations of smooth harmonic maps using geodesic finite elements, yielding the first high order scheme for this problem.

1 Introduction

This article investigates the numerical discretization of partial differential equations (PDEs) for functions whose codomain is a nonlinear Riemannian manifold M. Such problems arise, for example, in Cosserat-type material models [38, 47], liquid crystal physics [2, 28], and in image processing [52, 53]. Further we mention variational splines in manifolds [30], multi-body dynamics [31], and the investigation of harmonic maps into manifolds [14]. In signal processing of manifold-valued signals (see, e.g., [40]) any generalization of a linear variational method leads to a variational problem with values in a manifold.

The numerical approximation of solutions to such PDEs is difficult, because the relevant function spaces do not possess a linear structure. Therefore, standard discretization methods like finite elements cannot be used. Instead, various ad hoc methods have been proposed in the literature to discretize individual PDEs with particular codomains M. For example, to compute harmonic maps into the unit sphere S^2, Bartels and Prohl [6, 7] embedded S^2 into $\mathbb{R}^3$, and used first-order Lagrangian finite elements, constraining only the vertex values to be in S^2. In the literature on geometrically exact shells, the direction of the shell surface normal is frequently expressed as a set of angles, and the angles are discretized separately using finite elements [56]. For Cosserat continua (with values in $\mathbb{R}^3 \times SO(3)$), an alternative approach, used by Münch [36] and Müller [35], interpolates rotation vectors in $\mathfrak{so}(3)$ instead of in the group of rotations $SO(3)$. Finally, Simo et al. [48, 49] did not interpolate rotations at all. Rather, they kept the orientation at each quadrature point as a history variable, and updated it with linear interpolants of the corrections coming from a Newton method.

All these approaches have their shortcomings. Bartels and Prohl rely on an isometric embedding with corresponding projection, which is not always cheaply available in spaces other than S^2 (as an example we mention the space SPD(3) of symmetric positive definite 3×3 matrices, frequently used in applications [18]). Also, it is unclear whether their method can achieve higher than first-order convergence. The approach used by Münch and Müller fails when values close to the cut locus of the identity on $SO(3)$ are involved. Also, the dependence on a fixed tangent space of the codomain breaks objectivity. For the approach by Simo and coworkers [48], Crisfield and Jelenić [11] showed that it introduces a spurious dependence of the solution on the initial iterate and the parameters of the path-following mechanism.

With the notable exception of Bartels and Prohl, who proved weak convergence of their discrete solutions to weakly harmonic maps (see also Remark 7.3 below), no analytical investigations of any of the above discretization methods appear in the literature. Hence it is generally unknown whether these methods converge, and whether the nominal rate of of the approximation spaces is actually achieved. For the numerical approximation of explicitly given functions with values in a manifold, several theoretical results have been achieved in the recent years [13, 21, 22, 25, 40, 55, 57]. These methods are based on subdivision schemes, and it is unclear how they can be used for solving PDEs.

Recently, geodesic finite elements (GFE) have been introduced for partial differential equations with nonlinear codomains [42, 43, 44]. Based on the Karcher mean (or Riemannian center of mass), they form a natural generalization of Lagrangian finite elements of arbitrary order to the case where the codomain M is a nonlinear Riemannian manifold. Geodesic finite elements do not rely on an embedding of M into a linear space, and form a conforming discretization in the sense that geodesic finite element functions are H^1-functions [44, Thm. 5.1]. Also, they are equivariant under isometries of M. In mechanics, this leads to the desirable property that discretizations of objective problems are again objective.

In [42, 43, 44], numerical studies of the discretization error were performed. These studies involved geodesic finite elements of order up to three for functions mapping into the unit sphere S^2 and the special orthogonal group SO(3). In all cases optimal convergence orders in the L^2- and H^1-norms were observed. However, no analytical investigation of the discretization error was given at all. We make up for this with the present article, providing a complete, intrinsic convergence theory for geodesic finite elements for problems of variational type.

By "variational type" we mean the following setting. For a domain $\Omega \subset \mathbb{R}^d$ and M a Riemannian manifold, we look at minimization problems

$$u : \Omega \to M, \qquad u = \arg\min_{w \in H} \mathfrak{J}(w), \tag{1}$$

with $\mathfrak{J} : H \to \mathbb{R}$ a nonlinear functional. The domain H of $\mathfrak{J}$ is a set of functions $\Omega \to M$ of H^1 smoothness, which we discuss in detail in later sections. By construction, GFE functions are H^1 functions, and the set V^h of GFE functions for a given grid is a subset of H. We can therefore formulate a discrete problem by restricting $\mathfrak{J}$ to V^h. The discrete solution is

$$u^h = \arg\min_{w^h \in V^h} \mathfrak{J}(w^h), \tag{2}$$

i.e., we minimize the original energy functional over a finite-dimensional subset of the original set H.

As in the linear case, assessing the error of this numerical procedure is done in two steps. First, under an ellipticity assumption on the energy $\mathfrak{J}$, we show that u^h is a quasioptimal solution in the approximation space V^h, that is, the error between u^h and u is comparable to the approximation power of the space V^h (inspired by the linear theory we call such a result a *Céa lemma*). As it turns out, such a result can be proved easily in general metric spaces, using only certain convexity properties of the energy along geodesics, see Theorem 3.1. However, for the crucial H^1-type distance this convexity is difficult to verify in practice. We therefore also give a more elaborate result (Theorem 3.3), which allows to bound the H^1 distance using variations of the energy along geodesic homotopies. The results are independent of the construction of geodesic finite elements, and also cover other discretization methods.

Then, in a second step, the approximation power of the GFE spaces is assessed. In Theorem 5.4 we find that, provided that the solution u has a certain smoothness, the best approximation error of u in V^h decays like a power of the mesh size h.[1] We obtain the same orders as in the corresponding linear cases. All our arguments are completely intrinsic and the dependence of the approximation quality on the geometry of M turns out to only consist of iterated covariant derivatives of the logarithm mapping of M.

Combining these two results yields optimal convergence orders for the discretization error of geodesic finite element discretizations of general nonlinear elliptic variational problems (1) in Theorems 6.1 and 6.2. Compared to known results in the linear setting, the only important additional restriction of our results is that we require the solution u to have bounded first derivatives. As an application we give optimal a priori

---

[1]Similar results are shown in previous work by Grohs [24] for univariate nonlinear interpolation functions and the L^∞ norm, albeit with different methods.

error estimates for GFE discretizations of harmonic maps in Theorem 7.1 under certain assumptions of the sectional curvature of M.

We would like to emphasize that the two aforementioned results, viz. the nonlinear Céa lemmas and the approximation error estimate are highly interesting in their own right. For instance, the Céa lemmas apply to general approximation spaces other than GFE spaces, for example, the interpolation method used in [35, 36]. The approximation error estimates are also useful in the general context of approximating manifold-valued functions (see, e.g., [4, 40]).

A delicate issue is the proper choice of error measures in a nonlinear function space. In classical a priori bounds in linear spaces, a Sobolev-type half-norm $|u|$ of the solution u bounds the error $\|u - u^h\|$. Since there is no subtraction defined on the set H, we need to replace $\|u - u^h\|$ by a suitable distance metric in the function space H. We present two such metrics in Section 2.2, which reduce to $\|u - u^h\|_{H^1}$ if M is a linear space.

To generalize the term $|u|$, the covariant Sobolev half-norm is an obvious choice. However, in our expression for the interpolation error, terms appear that cannot be controlled by a Sobolev half-norm alone. In Section 2.4 we therefore introduce a slightly stronger concept, which we call the *smoothness descriptor*. We show that it provides information that is comparable to the actual Sobolev (half-)norms, but it does differ from them even in linear spaces. The question of whether our bounds also hold for covariant Sobolev norms is open.

We have structured the article as follows. In Chapter 2 we discuss the nonlinear spaces made up by functions $\Omega \to M$ of Sobolev smoothness. We propose two distance notions, and introduce the smoothness descriptor. In Chapter 3 we prove different forms of a nonlinear Céa lemma. Only then geodesic finite elements are introduced in Chapter 4. The second important part of the proof, the interpolation error bound, is shown in Chapter 5. This allows us to state a priori bounds for the discretization error for the discrete problem (2) in Chapter 6. Finally, in Chapter 7 we apply our results to harmonic maps and some of their generalizations. Under some regularity and curvature assumptions we obtain optimal error bounds for discrete harmonic maps of all approximation orders.

2 Nonlinear Function Spaces

Describing regularity of functions with a nonlinear codomain is a much less unified field than the corresponding linear theory. We introduce the notions that will be used in this article.

2.1 Sobolev Spaces

Let $\Omega \subset \mathbb{R}^d$ be open and bounded with Lipschitz boundary. On Ω we use canonical coordinates $x^1, \dots, x^d$. We use the notation $\partial^{\vec{k}}$ for the (weak) partial derivative of a d-variate function with respect to the multi-index $\vec{k} = (k_1, \dots, k_d) \in \mathbb{N}_0^d$, i.e.,

$$\partial^{\vec{k}} = \frac{\partial^{|\vec{k}|}}{(\partial x^d)^{k_d} \dots (\partial x^1)^{k_1}},$$

where we have written $|\vec{k}| := k_1 + \dots + k_d$. For a function $v : \Omega \to \mathbb{R}$ and an integrability parameter $p \in [1, \infty)$ we define the usual Sobolev half norms and norms

$$|v|_{W^{k,p}}^p := \int_\Omega \sum_{|\vec{k}|=k} |\partial^{\vec{k}} v(x)|^p \, dx, \qquad \|v\|_{W^{k,p}}^p := \sum_{j=0}^k |v|_{W^{j,p}}^p. \qquad (3)$$

We denote by $W^{k,p}(\Omega, \mathbb{R}^N)$ the set of functions $\Omega \to \mathbb{R}^N$ for which this quantity is finite componentwise. This set of functions forms a linear space. As an extension the space $W^{k,\infty}(\Omega, \mathbb{R}^N)$ is defined as the set of all functions $\Omega \to \mathbb{R}^N$ for which

$$\|v\|_{W^{k,\infty}} := \sum_{|\vec{k}| \leq k} \sup_{x \in \Omega} |\partial^{\vec{k}} v(x)| \, dx$$

is finite. For a simpler notation we will sometimes write $H^k(\Omega, \mathbb{R}^N)$ for $W^{k,2}(\Omega, \mathbb{R}^N)$.

Let now (M, g) be an n-dimensional Riemannian manifold with scalar product $\langle \cdot, \cdot \rangle_g$ and induced distance dist : $M \times M \to \mathbb{R}^+$. The following definition of a Sobolev space for functions with values in M is standard (see, e.g., [45]).

Definition 2.1. *Let $i : M \to \mathbb{R}^N$ be an isometric embedding (which always exists by [37]), $k \in \mathbb{N}_0$ and $p \in \mathbb{N} \cup \{\infty\}$. Define*

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

Again we will write $H^k(\Omega, M)$ for $W^{k,2}(\Omega, M)$. We shall also use the notation $C(\Omega, M)$ to denote continuous functions from Ω to M.

For nonlinear M these spaces obviously do not form vector spaces. However, under certain smoothness conditions the manifold structure of M is inherited. The following result is proved in [39].

Lemma 2.1. *If $k > d/p$, the spaces $W^{k,p}(\Omega, M)$ are Banach manifolds.*

Unfortunately, this lemma excludes the important case of $W^{1,2}(\Omega, M)$ with $d \geq 2$. However, even when $W^{k,p}(\Omega, M)$ is not a manifold we can still consider vector fields that are attached to a general continuous M-valued function.

Definition 2.2. *Let $u \in C(\Omega, M)$. We say that $W : \Omega \to TM$ is a vector field along u if $W(x) \in T_{u(x)}M$ for all $x \in \Omega$. The set of all vector fields along u is denoted by $u^{-1}TM$.*

For each continuous $u : \Omega \to M$, the set $u^{-1}TM$ forms a linear space which we now equip with two norms. The first is of L^p-type.

Definition 2.3. *Let $u \in C(\Omega, M)$. For a vector field $W \in u^{-1}TM$, and $p \in [1, \infty]$ we set*

$$|W|_{L^p}^p := \int_\Omega |W(x)|_{g(u(x))}^p \, dx,$$

with the obvious modifications for $p = \infty$.

The second one is a $W^{1,2}$-type norm, involving derivatives with respect to x. With $\frac{D}{dx^\alpha}$ we denote the covariant partial derivative along u with respect to x^α. In coordinates on Ω and M it reads

$$\frac{D}{dx^\alpha} W^l(x) := \frac{dW^l}{dx^\alpha}(x) + \Gamma_{ij}^l(u(x)) \frac{du^i}{dx^\alpha} W^j(x),$$

where we sum over repeated indices and denote with Γ_{ij}^l the Christoffel symbols associated to the metric of M.

Definition 2.4. *Let $u \in W^{1,\infty}(\Omega, M)$ and assume that the coordinate functions associated to the vector field $W \in u^{-1}TM$ are in $H^1(\Omega, \mathbb{R})$. We set*

$$|W|_{H^1}^2 := |W|_{L^2}^2 + \int_\Omega \left| \nabla_x W(x) \right|_g^2 \, dx := |W|_{L^2}^2 + \sum_{\alpha=1}^d \int_\Omega \left| \frac{D}{dx^\alpha} W(x) \right|_{g(u(x))}^2 \, dx. \tag{4}$$

For this norm we can show the following version of the Poincaré inequality.

Lemma 2.2 (Poincaré Inequality). *Let $u \in W^{1,\infty}(\Omega, M)$, and assume that $W \in u^{-1}TM$ with $W\big|_{\partial\Omega} = 0$. Then we have*

$$|W|_{L^2}^2 \leq C_1(\Omega) \sum_{\alpha=1}^d \int_\Omega \left| \frac{D}{dx^\alpha} W(x) \right|_{g(u(x))}^2 \, dx,$$

with $C_1(\Omega)$ the Poincaré constant of the domain Ω.

Proof. By the Poincaré inequality for $f : x \mapsto |W(x)|_{g(u(x))} \in \mathbb{R}$ we get

$$|W|_{L^2}^2 = \int_\Omega |W(x)|_{g(u(x))}^2 \, dx = \|f\|_{L^2}^2 \leq C_1 \sum_{\alpha=1}^d \left\| \frac{df}{dx^\alpha} \right\|_{L^2}^2.$$

Using the Cauchy inequality for g we may then calculate

$$\left| \frac{df}{dx^\alpha}(x) \right| = \frac{\langle W(x), \frac{D}{dx^\alpha} W \rangle_{g(u(x))}}{|W(x)|_{g(u(x))}} \leq \left| \frac{D}{dx^\alpha} W(x) \right|_{g(u(x))},$$

and the assertion follows. $\qquad\qquad\square$

We will frequently work with functions whose derivatives are bounded by a fixed constant $K > 0$. We therefore introduce the notation

$$H_K = H_K(\Omega, M) := \left\{ v \in W^{1,\infty}(\Omega, M) : \max_{\alpha=1\ldots,d} \left| \frac{d}{dx^\alpha} v(x) \right|_{g(v(x))} \leq K \right\}. \tag{5}$$

Note that functions in H_K are necessarily continuous.

2.2 Distance Measures in Nonlinear Function Spaces

To quantify the error between a function $u \in W^{k,p}(\Omega, M)$ and an approximation v of u in the same space, we need a distance measure on the nonlinear function space $W^{k,p}(\Omega, M)$. There are several ways to construct such a distance. The most natural one uses the embedding i used in Definition 2.1 to define the space $W^{k,p}(\Omega, M)$.

Definition 2.5. *For all $u, v \in W^{k,p}(\Omega, M)$ define*

$$\mathrm{dist}_{\mathrm{emb},W^{k,p}}(u, v) := \|i(u) - i(v)\|_{W^{k,p}}. \tag{6}$$

Since i is an isometry, the definition yields a distance metric. Also, it naturally degenerates to the standard Sobolev distance if M is a linear space.

This distance is convenient to evaluate, and defined even for functions u, v of little smoothness. However, aesthetically it is somewhat unpleasing, because it depends on the embedding i. A purely intrinsic distance can be defined using minimizing paths.

Definition 2.6. *Let H be a set of functions $\Omega \to M$, and $u, v \in H$. Suppose there is at least one continuously differentiable path γ from u to v in H. We denote by $\dot{\gamma} : \Omega \times [0,1] \to M$ the push-forward of $\frac{d}{dt}$ along γ, i.e., the vector field defined by*

$$\dot{\gamma}(x, t) := \frac{d}{dt} \gamma(x, t) \in T_{\gamma(x,t)} M.$$

For each $\gamma(t) \in H$ let there be a norm $|\cdot|_G$ on the space of vector fields along $\gamma(t)$, and define

$$\mathrm{dist}_G(u, v) := \inf_{\gamma \text{ path from } u \text{ to } v} \int_0^1 |\dot{\gamma}(t)|_G \, dt.$$

For each norm $|\cdot|_G$ we obtain a corresponding distance.

Definition 2.7. *For each $u, v \in C(\Omega, M)$ define*

$$\mathrm{dist}_{L^2}(u, v) := \inf_{\gamma \text{ path from } u \text{ to } v} \int_0^1 |\dot{\gamma}(t)|_{L^2} \, dt = \left(\int_\Omega \mathrm{dist}(u(x), v(x)) \, dx \right)^{1/2}$$

and

$$\mathrm{dist}_{L^\infty}(u, v) := \inf_{\gamma \text{ path from } u \text{ to } v} \int_0^1 |\dot{\gamma}(t)|_{L^\infty} \, dt = \sup_{x \in \Omega} \mathrm{dist}(u(x), v(x)).$$

Finally, for each $u, v \in W^{1,\infty}(\Omega, M)$ define

$$\mathrm{dist}_{W^{1,2}}(u, v) := \inf_{\gamma \text{ path from } u \text{ to } v} \int_0^1 |\dot{\gamma}(t)|_{H^1} \, dt. \tag{7}$$

The minimizing curves with respect to dist_{L^2} are called *geodesic homotopies*. They have the following useful property.

Remark 2.1. Let $\gamma : [0,1] \to C(\Omega, M)$ be a geodesic homotopy. Then for each $x \in \Omega$, the curve $\gamma(x, \cdot)$ is a geodesic on M.

Two functions that can be connected by a geodesic homotopy are called *geodesically homotopic*.

Defining distance using minimizing paths is a very elegant way of defining a distance, but it can be difficult to work with. Inside our proofs we will therefore frequently use a third error measure. It has a lot less mathematical structure than the two distance notions introduced above. However, we show below that it bounds both the embedded and the path-induced distance from above.

For the definition we need the exponential map $\exp(\cdot, \cdot)$ of M, as well as its inverse $\log(\cdot, \cdot)$. For both maps, the first argument denotes the base point $p \in M$. That is, $\exp(p, \cdot) : T_p M \to M$ and $\log(p, \cdot) : M \supset U \to T_p M$.

Definition 2.8. *Let $u, v \in W^{1,\infty}(\Omega, M)$. Define the quantity*

$$D_{1,2}(u,v)^2 := \int_\Omega |\log(u(x), v(x))|^2_{g(u(x))} \, dx + \sum_{\alpha=1}^d \int_\Omega \left| \frac{D}{dx^\alpha} \log(u(x), v(x)) \right|^2_{g(u(x))} dx. \tag{8}$$

In the linear case, this definition coincides with the usual H^1 error. It is, however, not a metric, since it is neither symmetric nor does it fulfill the triangle inequality.

The following lemma states that $D_{1,2}(u,v)$ provides an upper bound for $\|i(u) - i(v)\|_{H^1}$ for $u, v \in H_K$ as defined in (5). In the following we will write $A \lesssim B$ to say that a quantity A is bounded by a quantity B times a constant. If also the converse estimate holds we will sometimes write $A \sim B$.

Lemma 2.3. *For $u \in H_K$ and M isometrically embedded into Euclidean space we have the estimate*

$$\|i(u) - i(v)\|_{H^1} \lesssim D_{1,2}(u,v),$$

with the implicit constant only depending on K, the embedding i, and the geometry of M.

Proof. For simplicity we abuse notation and write $i(u) = u$, $i(v) = v$. Clearly we have

$$|u(x) - v(x)| \leq \text{dist}(u(x), v(x)) = |\log(u(x), v(x))|_{g(u(x))} \qquad \text{for almost all } x \in \Omega,$$

which takes care of the first term in the definition of $\|\cdot\|_{H^1}$. For the term associated with the derivative we put $v(x) = \exp(u(x), \log(u(x), v(x)))$ and compute, using the notation $\partial_1 \exp(p, w) = \frac{d}{dp} \exp(p, w)$, $\partial_2 \exp(p, w) = \frac{d}{dw} \exp(p, w)$ that

$$\frac{d}{dx^\alpha} v(x) = \partial_1 \exp\left(u(x), \log(u(x), v(x))\right) \frac{d}{dx^\alpha} u(x) + \partial_2 \exp\left(u(x), \log(u(x), v(x))\right) \frac{D}{dx^\alpha} \log(u(x), v(x)).$$

Then, since $\partial_1 \exp(p, 0) w = w$ for all $p \in M$, $w \in T_p M$, we have

$$\frac{d}{dx^\alpha} u(x) = \partial_1 \exp(u(x), 0) \frac{d}{dx^\alpha} u(x).$$

Hence, we can write the difference $\frac{d}{dx^\alpha} u(x) - \frac{d}{dx^\alpha} v(x)$ as a sum of the terms

$$\text{I} := \left[\partial_1 \exp(u(x), 0) - \partial_1 \exp\left(u(x), \log(u(x), v(x))\right) \right] \frac{d}{dx^\alpha} u(x)$$

and

$$\text{II} := \partial_2 \exp\left(u(x), \log(u(x), v(x))\right) \frac{D}{dx^\alpha} \log(u(x), v(x)).$$

The quantity II can be bounded in modulus by $\frac{D}{dx^\alpha} \log(u(x), v(x))$, up to a constant. By the Lipschitz continuity of $\partial_1 \exp$ in its second argument and the fact that $|\frac{d}{dx^\alpha} u(x)| \leq K$ by assumption, we can bound I up to a constant by $|\log(u(x), v(x))|$. This proves the statement. $\qquad \square$

6

Using a uniformity property of geodesic homotopies (which we prove in the following section), we can show that $D_{1,2}$ also bounds the distance $\text{dist}_{W^{1,2}}$ introduced in Definition 2.7.

Lemma 2.4. *For each $u, v \in H_K$ we have*

$$\text{dist}_{W^{1,2}}(u, v) \leq C_2 D_{1,2}(u, v),$$

where C_2 is the constant defined in (9).

Proof. Let Γ a geodesic homotopy (L^2-geodesic) from u to v. Then

$$\text{dist}_{W^{1,2}}(u, v) \leq \int_0^1 |\dot{\Gamma}(t)|_{H^1}\, dt \leq \sup_{t \in [0,1]} |\dot{\Gamma}(t)|_{H^1} \leq C_2 \inf_{t \in [0,1]} |\dot{\Gamma}(t)|_{H^1} \leq C_2 |\dot{\Gamma}(0)|_{H^1} = C_2 D_{1,2}(u, v).$$

$\square$

2.3 H^1-Uniformity of Geodesic Homotopies

The curves that induce the $\text{dist}_{W^{1,2}}$-distance are difficult to work with. The following result shows that geodesic homotopies are in some sense similar to these curves, provided the derivatives are bounded by a constant K. This will allow us to work with geodesic homotopies, and still obtain bounds in the $\text{dist}_{W^{1,2}}$-distance.

Lemma 2.5. *Assume that $u, v \in H_K(\Omega, M)$ and that Γ is a geodesic homotopy from u to v. Then*

$$\sup_{t \in [0,1]} \left|\dot{\Gamma}(\cdot, t)\right|_{H^1} \leq C_2 \inf_{t \in [0,1]} \left|\dot{\Gamma}(\cdot, t)\right|_{H^1}$$

with

$$C_2 = \sqrt{2} + 2^{d/2} C_3 \|\text{Rm}\|_g K \, \text{dist}_{L^\infty}(u, v), \tag{9}$$

where $\|\text{Rm}\|_g$ is the maximum norm of the Riemann curvature tensor Rm [12], and C_3 only depends on the geometry of M.

Proof. Since $t \mapsto \Gamma(x, t)$ is a geodesic we have that

$$\left|\dot{\Gamma}(x, t)\right|^2 = \text{dist}(u(x), v(x))^2, \qquad \text{for almost all } x \in \Omega,$$

independent of t. Hence

$$\left|\dot{\Gamma}(\cdot, t)\right|_{H^1}^2 = \text{dist}_{L^2}(u, v)^2 + U^2(t), \tag{10}$$

where we have defined

$$U^2(t) := \sum_{\alpha=1}^{d} \int_\Omega \left|\frac{D}{dx^\alpha} \dot{\Gamma}(x, t)\right|_{g(\Gamma(x,t))}^2\, dx.$$

We note that

$$\frac{D}{dx^\alpha} \frac{d}{dt} \Gamma(x, t) = \frac{D}{dt} \frac{d}{dx^\alpha} \Gamma(x, t),$$

as well as the fact that

$$J^\alpha(x, t) := \frac{d}{dx^\alpha} \Gamma(x, t)$$

satisfies the Jacobi differential equation

$$\frac{D^2}{dt^2} J^\alpha(x, t) = \text{Rm}\big(J^\alpha(x, t), \dot{\Gamma}(x, t)\big)\dot{\Gamma}(x, t).$$

Using this we can write for every $\alpha = 1, \ldots, d$

$$\frac{d}{dt} \left\langle \frac{D}{dx^\alpha} \dot{\Gamma}(x,t), \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right\rangle_{g(\Gamma(x,t))} = 2 \left\langle \frac{D}{dt} \frac{D}{dx^\alpha} \dot{\Gamma}(x,t), \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right\rangle_{g(\Gamma(x,t))}$$

$$= 2 \left\langle \frac{D^2}{dt^2} J^\alpha(x,t), \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right\rangle_{g(\Gamma(x,t))}$$

$$= 2 \left\langle \mathrm{Rm}\left(J^\alpha(x,t), \dot{\Gamma}(x,t) \right) \dot{\Gamma}(x,t), \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right\rangle_{g(\Gamma(x,t))}$$

$$\leq 2 \|\mathrm{Rm}\|_g \left| J^\alpha(x,t) \right|_{g(\Gamma(x,t))} \left| \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right|_{g(\Gamma(x,t))} \left| \dot{\Gamma}(x,t) \right|_{g(\Gamma(x,t))}^2.$$

For simplicity we shall omit the subscript $g(\Gamma(x,t))$ from now on.

Since we can write J^α as

$$J^\alpha(x,t) = \frac{d}{dx^\alpha} \exp\left(u(x), t \log(u(x), v(x)) \right),$$

we see that there exists a uniform constant C_3, only depending on the geometry of M such that

$$|J^\alpha(x,t)| \leq C_3 \max\left(\left| \frac{d}{dx^\alpha} u(x) \right|, \left| \frac{d}{dx^\alpha} v(x) \right| \right) \leq C_3 K.$$

We can use the previous considerations to bound the time derivative of U^2. Together with the Cauchy–Schwarz inequality we have

$$\left| \frac{d}{dt} U^2(t) \right| \leq 2C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v) \sum_{\alpha=1}^d \left(\int_\Omega \left| \frac{D}{dx^\alpha} \dot{\Gamma}(x,t) \right|^2 dx \right)^{1/2} \operatorname{dist}_{L^2}(u,v)$$

$$\leq 2^{(d+1)/2} C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v) \operatorname{dist}_{L^2}(u,v) U(t).$$

We divide by $U(t)$ to get

$$\left| \frac{dU(t)}{dt} \right| = \left| \frac{\frac{d}{dt} U^2(t)}{2U(t)} \right| \leq 2^{(d-1)/2} C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v) \operatorname{dist}_{L^2}(u,v).$$

The results above imply that

$$|U(t_2) - U(t_1)| \leq 2^{(d-1)/2} C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v) \operatorname{dist}_{L^2}(u,v) \tag{11}$$

for any $t_1, t_2 \in [0,1]$.

Now we can use (10) together with (11) to see that for $t_1, t_2 \in [0,1]$ we have

$$\frac{|\dot{\Gamma}(\cdot, t_1)|_{H^1}}{|\dot{\Gamma}(\cdot, t_2)|_{H^1}} \leq \sqrt{2} \, \frac{\operatorname{dist}_{L^2}(u,v) + U(t_1)}{\operatorname{dist}_{L^2}(u,v) + U(t_2)}$$

$$\leq \sqrt{2} \left[1 + \frac{|U(t_2) - U(t_1)|}{\operatorname{dist}_{L^2}(u,v) + U(t_2)} \right]$$

$$\leq \sqrt{2} \left[1 + \frac{2^{(d-1)/2} C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v) \operatorname{dist}_{L^2}(u,v)}{\operatorname{dist}_{L^2}(u,v) + U(t_2)} \right]$$

$$\leq \sqrt{2} + 2^{d/2} C_3 \|\mathrm{Rm}\|_g K \operatorname{dist}_{L^\infty}(u,v),$$

which finally proves the desired estimate. $\qquad \square$

2.4   The Smoothness Descriptor

We have given one definition of Sobolev regularity of functions $u : \Omega \to M$ in Section 2.1. A natural alternative is the covariant Sobolev norm

$$|u|_{H_{\mathrm{cov}}^k} := \sum_{\dim \vec{\beta} = k} \left(\int_\Omega \left| \mathcal{D}^{\vec{\beta}} u(x) \right|_{g(u(x))}^2 \right)^{1/2} dx, \qquad \|u\|_{H_{cov}^k} = \sum_{i=1}^k |u|_{H_{cov}^i}.$$

8

Here the symbol $\mathcal{D}^{\vec{\beta}} u$ means covariant partial differentiation along u with respect to the multi-index $\vec{\beta}$ in the sense that

$$\mathcal{D}^{\vec{\beta}} u := \frac{D}{dx^{\beta_k}} \cdots \frac{D}{dx^{\beta_2}} \frac{d}{dx^{\beta_1}} u, \qquad \vec{\beta} \in \{1, \ldots, d\}^k, \ k \in \mathbb{N}_0. \tag{12}$$

Additionally we define $\mathcal{D}^{\vec{\beta}} u := 1$ (a constant function $\Omega \to \mathbb{R}$) if $\dim \vec{\beta} = 0$. For a shorter notation we introduce the symbol

$$[d] := \{1, \ldots, d\}.$$

Note that (12) differs from the usual multi-index notation, which cannot be used because covariant partial derivatives do not commute.

Clearly, for linear M, these definitions coincide with the usual Sobolev half norms and norms (3). However, they cannot control all terms appearing in the nonlinear Bramble–Hilbert-Lemma in Section 5 (details are given in Remark 5.1). Therefore, we define the following alternative.

Definition 2.9 (Smoothness Descriptor). *For a function $u : \Omega \to M$, $k \geq 1$ and $p \in [1, \infty]$, define the homogeneous k-th order smoothness descriptor*

$$\dot{\Theta}_{p,k,\Omega}(u) := \sum_{\substack{\vec{\beta}_j \in [d]^{m_j}, \ j=1,\ldots,k \\ \sum_{j=1}^k m_j = k}} \left(\int_\Omega \prod_{j=1}^k \left| \mathcal{D}^{\vec{\beta}_j} u(x) \right|_{g(u(x))}^p dx \right)^{1/p},$$

with the usual modifications for $p = \infty$. Further, we define the L^p part

$$\dot{\Theta}_{p,0,\Omega}(u) := \min_{q \in M} \left(\int_\Omega |\text{dist}(u(x), q)|^p \, dx \right)^{1/p},$$

and the corresponding inhomogeneous smoothness descriptor

$$\Theta_{p,k,\Omega}(u) := \sum_{i=0}^k \dot{\Theta}_{p,i,\Omega}(u).$$

We will be mostly dealing with the case $p = 2$, for which we will omit the parameter p in the notation, i.e.,

$$\dot{\Theta}_{k,\Omega} := \dot{\Theta}_{2,k,\Omega} \quad \text{and} \quad \Theta_{k,\Omega} := \Theta_{2,k,\Omega}.$$

Note that we use a superposed dot to denote homogeneous quantities.

Remark 2.2. A function u with $\Theta_{k,\Omega}(u) < \infty$ must be uniformly continuous if $k > d/2$. Furthermore, in that case we have

$$\text{diam}(u) := \sup_{x,y \in \Omega} \text{dist}(u(x), u(y)) \lesssim \Theta_{k,\Omega}(u).$$

Both these assertions are direct consequences of the Sobolev embedding theorem.

To better present the smoothness descriptors Θ we discuss their relationships to other measures of regularity. For simplicity we restrict our analysis to the case $p = 2$. First, it follows directly from the definition that the smoothness descriptor Θ is a stronger notion than the covariant Sobolev norm.

Lemma 2.6. $\|u\|_{H^k_{cov}(\Omega,M)} \leq \Theta_{k,\Omega}(u).$

Proof. The proof follows immediately by noting that all terms which occur in the definition of $\|u\|_{H^k_{\text{cov}}(\Omega,M)}$ also occur in the definition of $\Theta_{k,\Omega}(u)$. $\qquad \square$

In the other direction we show that the Sobolev norm with respect to an embedding also bounds Θ from above, if k is sufficiently large.

Lemma 2.7. *Let i be an isometric embedding of M into a Euclidean space. For $k > \frac{d}{2}$ we have $\Theta_{k,\Omega}(u) \lesssim \|i \circ u\|_{H^k}^k.$*

Note that the smoothness descriptor is bounded by *the k-th power* of the corresponding norm.

Proof. Identify u with $i \circ u$ for simplicity. We need to estimate terms of the form

$$\left(\int_\Omega \prod_{j=1}^k \left| \mathcal{D}^{\vec{\beta}_j} u(x) \right|^2_{g(u(x))} dx \right)^{\frac{1}{2}} \tag{13}$$

with $\vec{\beta}_j \in [d]^{m_j}$, $j = 1, \ldots, k$, and $\sum_{j=1}^k m_j \leq k$. It will be no loss of generality to assume the most difficult case $\sum_{j=1}^k m_j = k$. First we deduce from the definition of the covariant derivative that any term of the form (13) can be estimated by a finite linear combination of terms of the form

$$\left(\int_\Omega \prod_{j=1}^k \left| \partial^{\vec{k}_j} u(x) \right|^2_{g(u(x))} dx \right)^{\frac{1}{2}}, \tag{14}$$

with $\sum_{j=1}^k |\vec{k}_j| = k$.

Now, for any values p_j, $j \in 1, \ldots, k$ with $\sum_{j=1}^k \frac{1}{p_j} \leq \frac{1}{2}$, by Hölder's inequality we can bound (14) by

$$\left(\int_\Omega \prod_{j=1}^k \left| \partial^{\vec{k}_j} u(x) \right|^2_{g(u(x))} dx \right)^{\frac{1}{2}} \leq \prod_{j=1}^k \left(\int_\Omega \left| \partial^{\vec{k}_j} u(x) \right|^{p_j}_{g(u(x))} dx \right)^{\frac{1}{p_j}} \leq \prod_{j=1}^k \|u\|_{W^{|\vec{k}_j|, p_j}}.$$

We make the specific choice

$$\frac{1}{p_j} = \frac{1}{2} - \frac{k - |\vec{k}_j|}{d} + \frac{(k-1)(k-d/2)}{kd}.$$

With this choice and $|\vec{k}| = k$ we have that

$$\sum_{j=1}^k \frac{1}{p_j} = \frac{k}{2} - \frac{k(k-1)}{d} + \frac{(k-d/2)(k-1)}{d} = \frac{1}{2}.$$

We shall now use the Sobolev embedding theorem which states that

$$\|u\|_{W^{l,p}} \lesssim \|u\|_{W^{k,2}},$$

whenever

$$\frac{1}{2} - \frac{k-l}{d} < \frac{1}{p}.$$

Setting $l = |\vec{k}_j|$ and $p = p_j$ for each $j = 1, \ldots, k$ we arrive at the desired statement. $\square$

A result similar to Lemma 2.7 can also be established for $p \neq 2$. In summary, our smoothness descriptor is an appropriate covariant way to measure smoothness of an M-valued function.

We finally show that the smoothness descriptor has a particular homogeneity property, also enjoyed by conventional Sobolev seminorms in linear spaces.

Definition 2.10. *Let T_1, T_2 be two domains in $\mathbb{R}^d$, and $\mathcal{F} : T_1 \to T_2$ a C^∞-diffeomorphism. For $l \in \mathbb{N}_0$ we say that $\mathcal{F}$ scales with h of order l if we have*

$$\sup_{x \in T_2} \left| \partial^{\vec{k}} \mathcal{F}^{-1}(x) \right| \lesssim h^{|\vec{k}|} \qquad \text{for all } \vec{k} \in \mathbb{N}_0^d, \ |\vec{k}| = 0, \ldots, l, \tag{15a}$$

$$|\det(\nabla \mathcal{F}(x))| \sim h^{-d} \qquad \text{for all } x \in T_1 \text{ (where } \nabla \mathcal{F} \text{ is the Jacobian of } \mathcal{F}), \tag{15b}$$

$$\sup_{x \in T_1} \left| \frac{d}{dx^\alpha} \mathcal{F}(x) \right| \lesssim h^{-1} \qquad \text{for all } \alpha = 1, \ldots, d. \tag{15c}$$

Such an $\mathcal{F}$ will be used to move finite element functions to the reference element and back, without losing approximation orders, see Section 5 below.

10

Lemma 2.8. *Let T_1, T_2 be two domains in $\mathbb{R}^d$, and $\mathcal{F} : T_1 \to T_2$ a map that scales with h of order l. Then for any $u : T_1 \to M$, $k \leq l$ and $p \in [1, \infty]$ we have*

$$\dot{\Theta}_{p,k,T_2}(u \circ \mathcal{F}^{-1}) \lesssim h^{-d/p} h^k \Theta_{p,k,T_1}(u).$$

Note that we bound the homogeneous smoothness descriptor by the inhomogeneous one.

Proof. It follows directly from the chain rule and the product rule that for any $m \in \mathbb{N}_0$ and $\vec{\beta} \in [d]^m$ the expression $\mathcal{D}^{\vec{\beta}}\left(u \circ \mathcal{F}^{-1}\right)$ can be written as a linear combination of terms of the form

$$\mathcal{D}^{\vec{\tau}} u(\mathcal{F}^{-1}(x)) \prod_{i=1}^{l} \partial^{\vec{k}_i} \left(\mathcal{F}^{-1}\right)^{j_i}(x)$$

with $\vec{\tau} \in [d]^n$,

$$\sum_{i=1}^{l} |\vec{k}_i| = m, \quad n \leq m, \ l \leq m,$$

and $\left(\mathcal{F}^{-1}\right)^{j_i}$ denoting the j_i-th coordinate of $\mathcal{F}^{-1}$. Using the scaling assumption (15a) we can therefore estimate the quantity

$$\left|\mathcal{D}^{\vec{\beta}}\left(u \circ \mathcal{F}^{-1}\right)\right|_{g(u(\mathcal{F}^{-1}(x)))}$$

by terms of the form

$$\left|\mathcal{D}^{\vec{\tau}} u(\mathcal{F}^{-1}(x)) \prod_{i=1}^{l} \partial^{\vec{k}_i} \left(\mathcal{F}^{-1}\right)^{j_i}(x)\right|_{g(u(\mathcal{F}^{-1}(x)))} \lesssim h^m \left|\mathcal{D}^{\vec{\tau}} u(\mathcal{F}^{-1}(x))\right|_{g(u(\mathcal{F}^{-1}(x)))}.$$

Therefore, every integrand

$$\prod_{j=1}^{k} \left|\mathcal{D}^{\vec{\beta}_j}\left(u \circ \mathcal{F}^{-1}\right)(x)\right|^p_{g(u(\mathcal{F}^{-1}(x)))}$$

in the definition of the homogeneous smoothness descriptor $\dot{\Theta}_{p,k,T_2}(u \circ \mathcal{F}^{-1})$ can be estimated pointwise by terms of the form

$$h^{pk} \prod_{j=1}^{k} \left|\mathcal{D}^{\vec{\tau}_j} u\left(\mathcal{F}^{-1}(x)\right)\right|^p_{g(u(\mathcal{F}^{-1}(x)))} \tag{16}$$

with

$$\vec{\tau}_j \in [d]^{n_j}, \quad n_j \in \mathbb{N}_0, \ \sum_{j=1}^{k} n_j \leq k.$$

Now, integrating (16) over T_2, and using the substitution $y = \mathcal{F}^{-1}(x)$, introduces an additional factor $h^{-d/p}$. Together with the scaling assumption (15b) we obtain the desired estimate. $\qquad\square$

Remark 2.3. The attentive reader will have noticed that only properties (15a) and (15b) have been used for the proof of Lemma 2.8. We will require the third assumption (15c) later, when we use scaling to derive local element-wise interpolation error estimates in Theorem 5.3 below.

3   Ellipticity and Céa's Lemma

Recall that we are trying to approximate the solution u of the variational problem

$$u = \arg\min_{w \in H} \mathfrak{J}(w) \tag{17}$$

by a minimizer v on a set $V \subset H$

$$v = \arg\min_{w \in V} \mathfrak{J}(w), \tag{18}$$

where H is a suitable set of functions, possibly fulfilling Dirichlet conditions. The classical linear Céa lemma assumes that H is a Hilbert space, and gives an estimate for the error between v and u in terms of the optimal approximation error $\inf_{w \in V} \|u - w\|_H$ of the approximation space V [8]. In this section we show analogous results when H consists of manifold-valued functions.

We proceed in two steps. Céa-type lemmas can be formulated and proved elegantly in general metric spaces. We show this in Section 3.1, and also give a reformulation for the case that H has a smooth structure with a Finsler norm. These results require certain convexity or ellipticity properties of the energy along distance-realizing curves. They are of independent interest, but they also illustrate some of the ideas of the subsequent section. There we allow variations over certain nonminimizing curves. The resulting Céa lemma is the basis of the discretization error bounds for geodesic finite elements in Chapter 6.

3.1 Céa's Lemma Based on Variations Along Curves

We start in an abstract setting. Suppose H is a metric space with distance function $\mathrm{dist}(\cdot, \cdot)$, and u and v are solutions of the minimization problems (17) and (18), respectively. We will refer to v as a *quasioptimal solution* if

$$\mathrm{dist}(u, v) \leq C \inf_{w \in V} \mathrm{dist}(u, w)$$

for a constant $C > 0$.

The main assumption leading to quasioptimality is a notion of strong convexity along curves. The following definition is taken from [3].

Definition 3.1. *A functional $\mathfrak{J} : H \to \mathbb{R}$ is called λ-convex along the curve $\gamma : [0,1] \to H$ if there is a $\lambda > 0$ such that*

$$\mathfrak{J}(\gamma(t)) \leq (1-t)\,\mathfrak{J}(\gamma(0)) + t\,\mathfrak{J}(\gamma(1)) - \frac{1}{2}\lambda t(1-t)\,\mathrm{dist}(\gamma(0), \gamma(1))^2$$

for all $t \in [0,1]$.

With this assumption, a metric version of the Céa Lemma follows almost immediately.

Theorem 3.1. *Assume that H is a metric space, and let $\mathfrak{J} : H \to \mathbb{R}$. Suppose that $u \in H$ is a minimizer of $\mathfrak{J}$ and let V be a subset of H for which the minimization problem*

$$v := \arg\min_{w \in V} \mathfrak{J}(w)$$

has a solution. Assume that there exists a curve γ with $\gamma(0) = u$ and $\gamma(1) = v$, along which the energy $\mathfrak{J}$ is λ-convex. Further, assume that $\mathfrak{J}$ is quadratically bounded around u in the sense that there is a constant $\Lambda > 0$ such that

$$\mathfrak{J}(w) - \mathfrak{J}(u) \leq \frac{\Lambda}{2}\,\mathrm{dist}(u, w)^2 \tag{19}$$

for all $w \in V$. Then

$$\mathrm{dist}(u, v) \leq \sqrt{2}\,\sqrt{\frac{\Lambda}{\lambda}}\,\inf_{w \in V} \mathrm{dist}(u, w).$$

Proof. Inserting $t = \frac{1}{2}$ into the definition of λ-convexity yields that

$$\mathrm{dist}(u, v)^2 \leq \frac{4}{\lambda}\left(\mathfrak{J}(v) - \mathfrak{J}(u)\right).$$

Since v is a minimizer on V, we can write

$$\mathrm{dist}(u, v)^2 \leq \frac{4}{\lambda}\inf_{w \in V}\left(\mathfrak{J}(w) - \mathfrak{J}(u)\right).$$

By (19), the right-hand-side can be bounded as desired. $\qquad\square$

A slightly more involved argument allows to get rid of the factor $\sqrt{2}$.

We now consider the case that H has a differentiable structure, which implies that we can have curves $\gamma : [0,1] \to H$ with well-defined tangent vectors $\dot{\gamma}$. We also assume that there is a norm $|\cdot|$ defined on these tangent vectors. The following alternative condition on $\mathfrak{J}$ is frequently convenient.

Definition 3.2. *We say that $\mathfrak{J}$ is elliptic along a differentiable curve $\gamma : [0,1] \to H$ if it is twice continuously differentiable along γ, and if there exist positive constants λ, Λ such that*

$$\lambda|\dot{\gamma}(t)|^2 \le \frac{d^2}{dt^2}\mathfrak{J}(\gamma(t)) \le \Lambda|\dot{\gamma}(t)|^2 \tag{20}$$

for all $t \in [0,1]$.

This concept is related to convexity in the following way. Assume that for each pair $w_1, w_2 \in H$ there is a differentiable path from w_1 to w_2, parametrized by arc length, that realizes the distance $\mathrm{dist}(w_1, w_2)$. We call such paths *(constant-speed) geodesics*.

Lemma 3.1. *Let $\mathfrak{J}$ be elliptic in the sense of Definition 3.2 along a given constant speed geodesic $\gamma : [0,1] \to H$. Then $\mathfrak{J}$ is λ-convex along that curve. If additionally $\gamma(0)$ is a minimizer of $\mathfrak{J}$, then γ is quadratically bounded in the sense of (19) along γ, with constant Λ.*

In particular, we see that the requirements of Theorem 3.1 are strictly weaker, because they are implied by ellipticity, but require no smoothness.

Proof. Set $f := \mathfrak{J}(\gamma) : [0,1] \to \mathbb{R}$. By the ellipticity assumption f is twice continuously differentiable. We first show that the lower bound on f'' implies that $\mathfrak{J}$ is λ-convex along γ. Pick $0 \le t \le 1$ and apply Taylor's formula to f at t. This gives

$$f(0) \ge f(t) + f'(t)(0-t) + \frac{1}{2}\lambda|\dot{\gamma}(s_1)|^2(0-t)^2$$

and

$$f(1) \ge f(t) + f'(t)(1-t) + \frac{1}{2}\lambda|\dot{\gamma}(s_2)|^2(1-t)^2,$$

where $0 \le s_1 \le t$ and $t \le s_2 \le 1$. Since γ is a constant speed geodesic we have $|\dot{\gamma}(s_1)|^2 = |\dot{\gamma}(s_2)|^2 = \mathrm{dist}(\gamma(0), \gamma(1))^2$. Multiply the first inequality by t, the second one by $1-t$, and add them to obtain the assertion.

Next we show that $f''(t) \le \Lambda|\dot{\gamma}(t)|^2$ implies $\mathfrak{J}(\gamma(1)) - \mathfrak{J}(\gamma(0)) \le \frac{\Lambda}{2}\mathrm{dist}(\gamma(0), \gamma(1))^2$. Using that $f'(0) = 0$ by assumption we can directly compute

$$\mathfrak{J}(\gamma(1)) - \mathfrak{J}(\gamma(0)) = \int_0^1 f'(t)\,dt - \int_0^1 f'(0)\,dt = \int_0^1 \int_0^t f''(s)\,ds\,dt$$

$$= \int_0^1 (1-t)f''(t)\,dt \le \int_0^1 (1-t)\Lambda|\dot{\gamma}(t)|^2\,dt = \frac{\Lambda}{2}\mathrm{dist}(\gamma(0), \gamma(1))^2.$$

$\square$

As an immediate consequence of Theorem 3.1 together with Lemma 3.1 we get the following result.

Theorem 3.2. *Let H be a Banach manifold with norm $|\cdot|$, and $\mathfrak{J} : H \to \mathbb{R}$ a functional. Assume that $u \in H$ is a minimizer of $\mathfrak{J}$, and that $\mathfrak{J}$ is elliptic along constant speed geodesics, with constants λ, Λ. For a $V \subset H$ set*

$$v := \underset{w \in V}{\arg\min}\,\mathfrak{J}(w),$$

assuming that this is well-defined. Then we have that

$$\mathrm{dist}(u,v) \le \sqrt{2}\,\sqrt{\frac{\Lambda}{\lambda}}\,\inf_{w \in V}\mathrm{dist}(u,w).$$

13

Using this theorem for the space $W^{1,2}(\Omega, M)$ together with the norm $|\cdot|_{H^1}$ defined in Definition 2.4 yields the following corollary, provided that $W^{1,2}(\Omega, M)$ is a Banach manifold. Unfortunately the latter only holds for $d = 1$.

Corollary 3.1. *Assume that $W^{1,2}(\Omega, M)$ is a Banach manifold, and let $\mathfrak{J} : W^{1,2}(\Omega, M) \to \mathbb{R}$. Assume that $u \in W^{1,2}(\Omega, M)$ is a minimizer of $\mathfrak{J}$, and that $\mathfrak{J}$ is elliptic along constant speed geodesics in $W^{1,2}(\Omega, M)$, starting in u. Let $V \subset W^{1,2}(\Omega, M)$ and*

$$v = \arg\min_{w \in V} \mathfrak{J}(w) \qquad \text{(the "discrete" solution)}.$$

Then we have that

$$\operatorname{dist}_{W^{1,2}}(u, v) \leq \sqrt{2} \, \sqrt{\frac{\Lambda}{\lambda}} \inf_{w \in V} \operatorname{dist}_{W^{1,2}}(u, w).$$

This corollary is the natural extension of the standard Céa lemma to nonlinear function spaces.

3.2 Céa's Lemma Using Geodesic Homotopies

When trying to apply the results of the previous section we encounter two problems. First, for the energies $\mathfrak{J}$ and domains Ω of our interest we consider variational problem formulations in $W^{1,2}(\Omega, M)$ and in general this space does not possess the structure of a Banach manifold [16, 26, 27]. Hence the results based on Banach manifolds cannot be used. Secondly, even if the space $W^{1,2}(\Omega, M)$ turns out to be a Banach manifold, it is difficult to work with constant speed geodesics in these spaces. In particular it is not easy to verify ellipticity properties along these curves for important energies, such as the harmonic energy.

To overcome these issues we generalize the approach somewhat. Instead of geodesics in $W^{1,2}$ we now consider geodesic homotopies. However, we still obtain bounds in terms of a $W^{1,2}$-like measure, namely the quantity $D_{1,2}$ introduced in (8). While this quantity is of little interest in itself, the result will allow to bound the discretization error of geodesic finite elements in terms of the embedded distance (6) and the geodesic distance (7). The proof is based on the H^1-uniformity of geodesic homotopies shown in Section 2.3. The price we pay is that we additionally have to assume the existence of a constant $K > 0$ such that $u \in H_K$ (the first weak derivatives are bounded by K), and $V \subset H_K$.

Theorem 3.3. *Let $H \subset W^{1,2}(\Omega, M)$, $\mathfrak{J} : H \to \mathbb{R}$ and $K > 0$. Assume that $u \in H \cap H_K(\Omega, M)$ is a stationary point of $\mathfrak{J}$ w.r.t. variations along geodesic homotopies in $H \cap H_K(\Omega, M)$ starting in u.*

For a second constant $L > 0$ denote

$$H_L^u := \big\{ v \,:\, \operatorname{dist}_{L^\infty}(u, v) \leq L \big\},$$

and assume that $\mathfrak{J} : H_K \cap H_L^u \cap H \to \mathbb{R}$ is elliptic along geodesic homotopies that start in u. Let $V \subset H_K \cap H_L^u \cap H$ and

$$v := \arg\min_{w \in V} \mathfrak{J}(w).$$

Then we have that

$$D_{1,2}(u, v) \leq C_2^2 \sqrt{\frac{\Lambda}{\lambda}} \inf_{w \in V} D_{1,2}(u, w),$$

with C_2 the uniformity constant (9), only depending on d, the product KL and the curvature of M.

Proof. For $w \in V$ define

$$f_w(t) := \mathfrak{J}(\Gamma(t))$$

with $\Gamma : [0, 1] \to H_K \cap H_L^u \cap H$ a geodesic homotopy from u to w. We have

$$\mathfrak{J}(w) - \mathfrak{J}(u) = \int_0^1 f_w'(t)\, dt - \int_0^1 f_w'(0)\, dt = \int_0^1 \int_0^t f_w''(s)\, ds\, dt = \int_0^1 (1 - t) f_w''(t)\, dt.$$

By the ellipticity assumption (20) we have

$$\lambda \int_0^1 (1 - t) |\dot{\Gamma}(t)|_{H^1}^2\, dt \leq \mathfrak{J}(w) - \mathfrak{J}(u) \leq \Lambda \int_0^1 (1 - t) |\dot{\Gamma}(t)|_{H^1}^2\, dt.$$

Now we use Lemma 2.5 which shows that

$$\frac{\lambda}{C_2^2}|\dot{\Gamma}(0)|_{H^1}^2 \leq \mathfrak{J}(w) - \mathfrak{J}(u) \leq \Lambda C_2^2 |\dot{\Gamma}(0)|_{H^1}^2,$$

where the constant C_2 depends only on d, K, L and the curvature of M. Noting further that

$$D_{1,2}(u,w)^2 = |\dot{\Gamma}(0)|_{H^1}^2$$

immediately yields that

$$D_{1,2}(u,v)^2 \leq \frac{C_2^2}{\lambda}\big[\mathfrak{J}(v) - \mathfrak{J}(u)\big] \leq \frac{C_2^2}{\lambda}\big[\mathfrak{J}(w) - \mathfrak{J}(u)\big]$$

for all $w \in V$. Furthermore we have

$$\mathfrak{J}(w) - \mathfrak{J}(u) \leq C_2^2 \Lambda D_{1,2}^2(u,w).$$

Together, we obtain that

$$D_{1,2}(u,v) \leq C_2^2 \sqrt{\frac{\Lambda}{\lambda}} \inf_{w \in V} D_{1,2}(u,w).$$

$\square$

Replacing $|\dot{\Gamma}(0)|_{H^1}^2$ by $|\dot{\Gamma}(1)|_{H^1}^2$ in the arguments above shows the following corollary.

Corollary 3.2. *With the notation of the previous theorem, we also have the estimate*

$$D_{1,2}(v,u) \leq C_2^2 \sqrt{\frac{\Lambda}{\lambda}} \inf_{w \in V} D_{1,2}(w,u).$$

The restriction that $u \in H_K$ does not appear in the linear theory. The question whether Theorem 3.3 can be shown without it is open.

Remark 3.1. Requiring that the approximation space V consists only of functions with derivatives bounded by K may lead to a restriction when considering families of approximation spaces V^h associated with a mesh width $h > 0$. If V^h are chosen as GFE spaces as introduced in Section 4 we show in Theorem 6.2 that the condition $V^h \subset H_K$ can be dispensed with, provided that u is sufficiently regular.

4 Geodesic Finite Elements

In Chapter 3 very little has been required from the approximation spaces V. For the theory based on distance-realizing curves in Section 3.1 only the existence of a minimizer of $\mathfrak{J}$ in V was asked. In Section 3.2 we additionally needed that the approximating functions that make up V have their derivatives bounded by a constant K.

In this section we present geodesic finite elements (GFE) as one particular example of a suitable space V. They have originally been introduced in [43, 44], but for completeness we give a brief review. The definition consists of two parts. First, nonlinear interpolation functions are constructed that interpolate values given on a reference element. Then, for a given grid, these interpolation functions are pieced together to form global finite element functions.

4.1 Geodesic Interpolation

Let T_{ref} be an open bounded subset of $\mathbb{R}^d$, which we will call reference element. Particular instances are the reference simplex $\big\{x \in \mathbb{R}^d \mid x^\alpha \geq 0, \ \alpha = 1, \ldots, d, \ \sum_{\alpha=1}^d x^\alpha \leq 1\big\}$, and the reference cube $[0,1]^d$. On T_{ref} we assume the existence of a set of Lagrangian interpolation polynomials, i.e., a set of Lagrange nodes $a_i \in T_{\text{ref}}$, $i = 1, \ldots, m$, and corresponding polynomial functions $\lambda_i : T_{\text{ref}} \to \mathbb{R}$ of order p such that

$$\lambda_i(a_j) = \delta_{ij} \qquad \text{for all } 1 \leq i, j \leq m,$$

and

$$\sum_{i=1}^{m} \lambda_i \equiv 1. \tag{21}$$

We now generalize Lagrange interpolation to values in a manifold. Let $v_i \in M$, $i = 1, \ldots, m$ be given values at the Lagrange nodes $a_i \in T_{\text{ref}}$. We want to construct a function $\Upsilon_v : T_{\text{ref}} \to M$ such that $\Upsilon_v(a_i) = v_i$ for all $i = 1, \ldots, m$. The following definition was given and motivated in [44].

Definition 4.1. *Let $\{\lambda_i, i = 1, \ldots, m\}$ be a set of p-th order scalar Lagrangian shape functions, and let $v_i \in M$, $i = 1, \ldots, m$ be values at the corresponding Lagrange nodes. We call*

$$\Upsilon \ : \ M^m \times T_{ref} \to M$$

$$\Upsilon(v_1, \ldots, v_m; x) = \operatorname*{arg\,min}_{q \in M} \sum_{i=1}^{m} \lambda_i(x) \operatorname{dist}(v_i, q)^2 \tag{22}$$

p-th order geodesic interpolation on M.

For fixed coefficients $v_1, \ldots, v_m$ we set $\Upsilon_v(\cdot) := \Upsilon(v_1, \ldots, v_m; \cdot)$ and obtain the desired function.

It is easy to verify that this definition reduces to p-th order Lagrangian interpolation if M is a linear space and $\operatorname{dist}(\cdot, \cdot)$ the standard distance. For the nonlinear case and $p = 1$, well-posedness of the definition under certain restrictions on the v_i is a classic result by Karcher [34]. For $p \geq 2$, where the λ_i can become negative, well-posedness has been proved in [44]. The interpolation function Υ is infinitely differentiable both as a function of the v_i and of the local coordinates x. This and several other features is discussed in [43, 44].

Since the values of Υ_v are defined as solutions of a minimization problem, we can also characterize them by the corresponding first-order optimality condition (see, for instance, [34]). We will make use of this representation in the interpolation error bound in Chapter 5.

Lemma 4.1. *For any $q \in M$ denote by $\log(q, \cdot) : M \supset U \to T_q M$ the inverse of the exponential map of M at q. Then $q^* := \Upsilon(v_1, \ldots, v_m; x)$ is (locally uniquely) characterized by the first-order condition*

$$\sum_{i=1}^{m} \lambda_i(x) \log(q^*, v_i) = 0 \ \in \ T_{q^*} M. \tag{23}$$

Interpolation error bounds for geodesic finite elements are based on the fact that the shape functions λ_i are exact on polynomials of degree no greater than p, meaning that

$$\sum_{i=1}^{m} \lambda_i(x) q(a_i) = q(x) \tag{24}$$

for all polynomials $q : T_{\text{ref}} \to \mathbb{R}$ of degree less than or equal to p. Using this we can prove the following technical property.

Lemma 4.2. *For all multi-indices $\vec{l}$ with $|\vec{l}| \leq p$ and all functions $f : T_{ref} \to \mathbb{R}$ we have*

$$\sum_{i=1}^{m} \lambda_i(x) (a_i - x)^{\vec{l}} f(x) = 0.$$

Proof. We start by fixing some arbitrary $x_* \in T_{\text{ref}}$. Then we can write

$$\sum_{i=1}^{m} \lambda_i(x) (a_i - x_*)^{\vec{l}} f(x_*) = \sum_{i=1}^{m} \lambda_i(x) p_{x_*}(a_i)$$

where

$$p_{x_*}(y) := (y - x_*)^{\vec{l}} f(x_*)$$

is a polynomial of degree $|\vec{l}|$. By (24) we get

$$\sum_{i=1}^{m} \lambda_i(x) \left(a_i - x_*\right)^{\vec{l}} f(x_*) = p_{x_*}(x).$$

Since by definition $p_{x_*}(x_*) = 0$, this implies

$$\sum_{i=1}^{m} \lambda_i(x_*) \left(a_i - x_*\right)^{\vec{l}} f(x_*) = 0,$$

which, by the arbitrariness of x_*, implies the statement. $\qquad\square$

4.2 Geodesic Finite Element Functions

Let now Ω be a domain in $\mathbb{R}^d$. Suppose we have a conforming grid $\mathcal{G}$ for Ω with elements not necessarily restricted to simplices. Let $n_i \in \Omega$, $i = 1, \ldots, |n|$ be a set of Lagrange nodes such that for each element T of $\mathcal{G}$ there are m nodes $a_{T,i}$ contained in T, and such that the p-th order interpolation problem on T is well posed.

Definition 4.2 (Geodesic Finite Elements)**.** *Let $\mathcal{G}$ be a conforming grid on Ω, and let M be a 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 interpolation in the sense that*

$$v^h|_T(x) = \Upsilon\big(v_{T,1}, \ldots, v_{T,m}; \mathcal{F}_T(x)\big),$$

where the $\mathcal{F}_T : T \to T_{ref}$ are element mappings (typically affine or multilinear), and the $v_{T,i}$ are values in M corresponding to the Lagrange nodes $a_{T,i}$. The space of all such functions v^h will be denoted by $V_{p,\mathcal{G}}^M$.

This definition reduces to standard (vector-valued) Lagrangian finite elements if M is a linear space with the usual Euclidean distance.

The following property is crucial for our analysis, because we always assume that the approximation space V is a subset of the solution space. The proof is given in [43] and [44].

Theorem 4.1. *$V_{p,\mathcal{G}}^M(\Omega) \subset H^1(\Omega, M)$ for all $p \geq 1$.*

While this holds for all grids $\mathcal{G}$ and all polynomial orders p, we note that geodesic finite element spaces are generally not nested. This means that in general $V_{p,\mathcal{G}}^M \not\subset V_{p+1,\mathcal{G}}^M$ and $V_{p,\mathcal{G}}^M \not\subset V_{p,\mathcal{G}'}^M$ if $\mathcal{G}'$ is a uniform refinement of $\mathcal{G}$. See [44, Chap. 4] for a brief discussion.

Remark 4.1. In numerical algorithms one uses the algebraic representation of $V_{p,\mathcal{G}}^M$, that is, a function $v^h \in V_{p,\mathcal{G}}^M$ is identified with a set of nodal coefficients $\bar{v} \in M^{|n|}$. However, note that $V_{p,\mathcal{G}}^M$ is not globally homeomorphic to $M^{|n|}$; in fact, it is not even globally a manifold. This is so because for certain sets of coefficients there is more than one interpolation function (see [42] for a simple example). On the other hand, it is shown in [44] that for many $\bar{v} \in M^{|n|}$ there is only a single interpolating function v^h, and then there is a diffeomorphism mapping a neighborhood of $\bar{v}$ in $M^{|n|}$ to a neighborhood of v^h in $V_{p,\mathcal{G}}^M$. In this sense the space $V_{p,\mathcal{G}}^M$ contains many small "manifold patches". Its global structure, however, remains unclear.

To prove quasi-optimality in Theorem 3.3 we had to make the assumption that the discrete space V contains only functions with first derivatives bounded by a global constant K. While it is obvious that each GFE function has bounded first derivatives, a global bound for all functions of a space $V_{p,\mathcal{G}}^M$ exists only if M has finite diameter. This global bound depends on the grid size h. The specific nature of this dependence will allow us in Theorem 6.2 to circumvent the restriction $V \subset H_K$ and obtain discretization error bounds without constraints on the ansatz space. For later use there we therefore state the following simple result, which holds for all orders p, and for M with bounded or unbounded diameter.

Lemma 4.3. *Let $\mathcal{G}$ be such that $\mathcal{F}_T$ scales with h of order p for each element T of $\mathcal{G}$. Then for each function $v^h \in V_{p,\mathcal{G}}^M$ we have*

$$\Theta_{\infty,1,\Omega}(v^h) \leq Ch^{-1},$$

where the constant C depends on the values of v^h at the Lagrange nodes.

In order to assess the approximation properties of the spaces $V_{p,\mathcal{G}}^M$, we finally construct the pointwise interpolation operator mapping continuous functions with values in M to elements in $V_{p,\mathcal{G}}^M$. As in the classical linear case we first define the interpolant on a reference element. We start by fixing the reference element T_{ref} with Lagrangian interpolation nodes a_i and corresponding local basis functions λ_i, $i = 1, \ldots, m$. Given a function $u : T_{\mathrm{ref}} \to M$ its local Lagrangian interpolant on T_{ref} is defined by

$$\mathbb{I}_{T_{\mathrm{ref}}} u(x) := \Upsilon \left(u(a_1), \ldots, u(a_m); x \right).$$

Likewise, for a general element T with associated mapping $\mathcal{F}_T : T \to T_{\mathrm{ref}}$, the local Lagrangian interpolant is given by

$$\mathbb{I}_T u(x) = \mathbb{I}_{T_{\mathrm{ref}}} \left(u \circ \mathcal{F}_T^{-1} \right) (\mathcal{F}_T(x)), \quad x \in T.$$

With these notions at hand we can define the *geodesic Lagrange interpolant* of a continuous function $u : \Omega \to M$.

Definition 4.3. *For each continuous function $u : \Omega \to M$ define the geodesic Lagrange interpolant $\mathbb{I}_{\mathcal{G}} u \in V_{p,\mathcal{G}}^M$ by*

$$\mathbb{I}_{\mathcal{G}} u(x) = \mathbb{I}_T u(x), \quad x \in T, \ T \in \mathcal{G}.$$

Note that unlike in the linear case, this interpolant is not always unique.

5 Interpolation Error Estimates

The goal of this section is to derive estimates of optimal order for the interpolation error between a function $u : \Omega \to M$ and its interpolant $\mathbb{I}_{\mathcal{G}} u$. To motivate our proof, we briefly review how interpolation error estimates can be obtained in the linear case $M = \mathbb{R}$. There, we start with an error bound on the reference element.

Theorem 5.1. *Let $u : T_{ref} \to \mathbb{R}$ satisfy $u \in H^k(T_{ref})$ with $k > d/2$. Then we have*

$$\| u - \mathbb{I}_{T_{ref}} u \|_{H^1(T_{ref})} \lesssim | u |_{H^k(T_{ref})}.$$

In order to turn Theorem 5.1 into an estimate for a small element T, say, $T = h T_{\mathrm{ref}}$ with $\mathcal{F}_T(x) := h^{-1} x$ and a function $u : T \to \mathbb{R}$ we use the fact that the Sobolev seminorm satisfies the subhomogeneity property

$$| u \circ \mathcal{F}_T^{-1} |_{H^k(T_{\mathrm{ref}})} \lesssim h^k h^{-d/2} | u |_{H^k(T)}. \tag{25}$$

We obtain the factor h^k from the k-fold application of the chain rule to $u \circ \mathcal{F}_T^{-1}(x) = u(hx)$, and the factor $h^{-d/2}$ from the integral transformation formula.

Then, denoting $v := u \circ \mathcal{F}_T^{-1} : T_{\mathrm{ref}} \to \mathbb{R}$ and using $\frac{d}{dx^\alpha} \mathcal{F}_T = h^{-1}$ for all $\alpha = 1, \ldots, d$, we get

$$| u - \mathbb{I}_T u |_{H^1(T)}^2 = h^{-2} \sum_{\alpha=1}^d \int_T \left| \left(\frac{d}{dx^\alpha} v - \frac{d}{dx^\alpha} \mathbb{I}_{T_{\mathrm{ref}}} v \right) \right|^2 \circ \mathcal{F}_T(x) \, dx = h^{-2} h^d | v - \mathbb{I}_{T_{\mathrm{ref}}} v |_{H^1(T_{\mathrm{ref}})}^2.$$

We can now invoke Theorem 5.1 and get the estimate

$$| u - \mathbb{I}_T u |_{H^1(T)}^2 \lesssim h^{-2} h^d | u \circ \mathcal{F}_T^{-1} |_{H^k(T_{\mathrm{ref}})}^2$$

which, together with (25) yields the classical estimate

$$\| u - \mathbb{I}_T u \|_{H^1(T)} \lesssim h^{k-1} | u |_{H^k(T)}.$$

To obtain a similar result for nonlinear codomains M we first need a generalization of Theorem 5.1. We prove such a result in Section 5.1, where the norm on the left becomes the quantity $D_{1,2}$, and the norm on the right becomes the smoothness descriptor Θ. Then, in Section 5.2 we assemble these local estimates to establish optimal approximation rates for the geodesic finite element spaces $V_{p,\mathcal{G}}^M$. This works because the smoothness descriptor Θ also has the subhomogeneity property (25) (Lemma 2.8).

18

5.1 Nonlinear Elementwise Estimates

In this section we prove a nonlinear generalization of the linear element-wise approximation result of Theorem 5.1. We cope with the implicit definition (22) of the interpolant by a clever use of the equilibrium condition (23).

Let $\log(p, \cdot) : M \to T_p M$ be the inverse of the exponential map at p. Denote by ∇_1, ∇_2 the covariant derivative of a bivariate function with respect to the first and second argument, respectively. In particular, for $l \in \mathbb{N}$ we will require the derivatives

$$\nabla_2^l \log(p, q) : (T_q M)^l \to T_p M$$

and

$$\nabla_2^l \nabla_1 \log(p, q) : T_p M \otimes (T_q M)^l \to T_p M;$$

more precisely their norms

$$\|\nabla_2^l \log(p, q)\| = \sup_{v_1, \ldots, v_l \in T_q M} \frac{\left|\nabla_2^l \log(p, q) (v_1, \ldots, v_l)\right|_{g(p)}}{\prod_{i=1}^l |v_i|_{g(q)}}$$

and

$$\|\nabla_2^l \nabla_1 \log(p, q)\| = \sup_{\substack{v_1, \ldots, v_l \in T_q M \\ w \in T_p M}} \frac{\left|\nabla_2^l \nabla_1 \log(p, q) (w, v_1, \ldots, v_l)\right|_{g(p)}}{|w|_{g(p)} \prod_{i=1}^l |v_i|_{g(q)}}.$$

Now we can state and prove a nonlinear elementwise approximation result.

Theorem 5.2. *Let* $u \in W^{k,2}(T_{ref}, M)$, *and* $\mathbb{I}_{T_{ref}} u$ *its p-th order geodesic interpolation. For $k > d/2$ and $p \geq k - 1$, we have*

$$\int_{T_{ref}} \left|\log(\mathbb{I}_{T_{ref}} u(x), u(x))\right|^2_{g(\mathbb{I}_{T_{ref}} u(x))} dx \lesssim \mathcal{C}_{1,u}^2(T_{ref}) \dot{\Theta}_{k, T_{ref}}(u)^2, \tag{26}$$

and for any $\alpha = 1, \ldots, d$

$$\int_{T_{ref}} \left|\frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{ref}} u(x), u(x))\right|^2_{g(\mathbb{I}_{T_{ref}} u(x))} dx \lesssim \left(\mathcal{C}_{1,u}^2(T_{ref}) + \mathcal{C}_{2,u}^2(T_{ref})\right) \dot{\Theta}_{k, T_{ref}}(u)^2, \tag{27}$$

where

$$\mathcal{C}_{1,u}(T_{ref}) := \sup_{1 \leq l \leq k} \sup_{\substack{p \in \mathbb{I}_{T_{ref}} u(T_{ref}) \\ q \in u(T_{ref})}} \left\|\nabla_2^l \log(p, q)\right\|$$

and

$$\mathcal{C}_{2,u}(T_{ref}) := \sup_{1 \leq l \leq k} \sup_{\substack{p \in \mathbb{I}_{T_{ref}} u(T_{ref}) \\ q \in u(T_{ref})}} \left\|\nabla_2^l \nabla_1 \log(p, q)\right\|.$$

The implicit constants are independent of u and M and only depend on the basis functions λ_i.

Note that the left hand sides of (26) and (27) make up the quantity $D_{1,2}(\mathbb{I}_{T_{ref}} u, u)$.

Proof. We split the proof into eight steps.

Step 1 We first prove (26). Using the balance law (23) we obtain for any $x \in T_{ref}$ that

$$\sum_{i=1}^m \lambda_i(x) \log\left(\mathbb{I}_{T_{ref}} u(x), u(a_i)\right) = 0.$$

Adding a zero we rewrite this as

$$\log(\mathbb{I}_{T_{ref}} u(x), u(x)) = \log(\mathbb{I}_{T_{ref}} u(x), u(x)) - \sum_{i=1}^m \lambda_i(x) \log\left(\mathbb{I}_{T_{ref}} u(x), u(a_i)\right),$$

and call the right hand side $\varepsilon(x) \in T_{\mathbb{I}_{T_{ref}} u(x)} M$. To obtain (26), we need to control the L^2-norm of the function ε.

Step 2 Next we define the helper function

$$G(x, y) := \log\left(\mathbb{I}_\Delta u(x), u(y)\right),$$

and perform a Taylor expansion of G in its second argument around $y = x$ (note that for fixed x, the function G takes its values in a vector space).

In what follows we shall use the notation $\partial_y^{\vec{k}} G(x, y)$ for the partial derivatives of G with respect to its second argument and the multi-index $\vec{k}$. The Taylor expansion then reads

$$G(x, y) = \sum_{|\vec{l}| < k} \frac{(y - x)^{\vec{l}}}{\vec{l}!} \partial_y^{\vec{l}} G(x, x) + \sum_{|\vec{k}| = k} R_{\vec{k}}(x, y)(y - x)^{\vec{k}}, \tag{28}$$

where

$$R_{\vec{k}}(x, y) = \frac{|\vec{k}|}{\vec{k}!} \int_0^1 (1 - t)^{|\vec{k}| - 1} \partial_y^{\vec{k}} G(x, x + t(y - x)) \, dt.$$

We can express the terms $\log\left(\mathbb{I}_{T_{\mathrm{ref}}} u(x), u(a_i)\right)$ occurring in the definition of ε in the form (28) and get

$$\varepsilon(x) = G(x, x) - \sum_{i=1}^m \lambda_i(x) \left(\sum_{|\vec{l}| < k} \frac{(a_i - x)^{\vec{l}}}{\vec{l}!} \partial_y^{\vec{l}} G(x, x) + \sum_{|\vec{k}| = k} R_{\vec{k}}(x, a_i)(a_i - x)^{\vec{k}} \right).$$

Using that the weight functions λ_i form a partition of unity on T_{ref} (21) we get

$$-\varepsilon(x) = \sum_{0 < |\vec{l}| < k} \sum_{i=1}^m \lambda_i(x) \frac{(a_i - x)^{\vec{l}}}{\vec{l}!} \partial_y^{\vec{l}} G(x, x) + \sum_{i=1}^m \lambda_i(x) \sum_{|\vec{k}| = k} R_{\vec{k}}(x, a_i)(a_i - x)^{\vec{k}}, \tag{29}$$

where the zeroth order derivative cancels with $G(x, x)$.

Step 3 By the assumption $p \geq k - 1$ we can apply Lemma 4.2 with $f(x) = \partial_y^{\vec{l}} G(x, x)$ to each sum $\sum_{i=1}^m \lambda_i(x) \frac{(a_i - x)^{\vec{l}}}{\vec{l}!} \partial_y^{\vec{l}} G(x, x)$ in (29) and see that the first addend in (29) is zero. Hence we can write $\varepsilon(x)$ as the sum

$$\varepsilon(x) = \sum_{i=1}^m \varepsilon_i(x) \qquad \text{with} \qquad \varepsilon_i(x) := -\lambda_i(x) \sum_{|\vec{k}| = k} R_{\vec{k}}(x, a_i)(a_i - x)^{\vec{k}}.$$

We now treat each term ε_i separately. For simplicity, we may assume, after a suitable translation (depending on the index i) that $a_i = 0$ and thus we arrive at the pointwise estimate

$$|\varepsilon_i(x)|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} \lesssim \sum_{|\vec{k}| = k} \left| \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}| - 1} \partial_y^{\vec{k}} G(x, tx) x^{\vec{k}} \, dt \right|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))}$$

$$\leq \sum_{|\vec{k}| = k} \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}| - 1} \left| \partial_y^{\vec{k}} G(x, tx) \right|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} x^{\vec{k}} \, dt, \quad (30)$$

where we have used that the λ_i are bounded on T_{ref}.

Step 4 In order to untangle the derivatives of log and u in the expression $\partial_y^{\vec{k}} G(x, y)$, we use the chain rule which yields

$$\left| \partial_y^{\vec{k}} G(x, y) \right|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} \lesssim \sum_{\substack{1 \leq l \leq k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \nabla_2^l \log\left(\mathbb{I}_{T_{\mathrm{ref}}} u(x), u(y)\right) \left(\mathcal{D}^{\vec{\beta}_1} u(y), \dots, \mathcal{D}^{\vec{\beta}_l} u(y) \right). \tag{31}$$

We repeat that we use the notation

$$\nabla_2^l \log(p, q) : (T_q M)^l \to T_p M$$

to denote the l-th order covariant derivative of the function $q \mapsto \log(p, q)$, which is an l-multilinear form.

20

Remark 5.1. We record here that this is the point where the smoothness descriptor $\dot{\Theta}$ (defined in Section 2.4) becomes necessary. Indeed, (31) already indicates that control over products of covariant derivatives of lower order is required whenever $\nabla_2^l \log \neq 0$. Note also that in the linear case we have $\nabla_2^l \log = 0$ for all $l > 1$ and therefore the usual Sobolev seminorm $|\cdot|_{H^k}$ is sufficient to obtain the desired control over terms of the form (31). Keeping this in mind it is easy to see that in the linear case our proof yields exactly the expected bounds for the interpolation error in the Sobolev seminorm.

By (30) we get

$$
\int_{T_{\text{ref}}} |\varepsilon_i(x)|^2_{g(\mathbb{I}_{T_{\text{ref}}} u(x))} \, dx \lesssim \int_{T_{\text{ref}}} \left(\int_0^1 t^{|\vec{k}|-1} \left| \partial_y^{\vec{k}} G(x, tx) \right|_{g(\mathbb{I}_{T_{\text{ref}}} u(x))} x^{\vec{k}} \, dt \right)^2 dx
$$

$$
\lesssim \sum_{\substack{1 \le l \le k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \int_{T_{\text{ref}}}
$$

$$
\left(\int_0^1 t^{|\vec{k}|-1} \left| \nabla_2^l \log \left(\mathbb{I}_{T_{\text{ref}}} u(x), u(tx) \right) \left(\mathcal{D}^{\vec{\beta}_1} u(tx), \ldots, \mathcal{D}^{\vec{\beta}_k} u(tx) \right) \right|_{g(\mathbb{I}_{T_{\text{ref}}} u(x))} x^{\vec{k}} \, dt \right)^2 dx
$$

$$
\le \sum_{\substack{1 \le l \le k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \int_{T_{\text{ref}}} \left(\int_0^1 t^{|\vec{k}|-1} H^{\vec{\beta}_1, \ldots, \vec{\beta}_l}(x, tx) x^{\vec{\beta}} \, dt \right)^2 dx,
$$

where we have put

$$
H^{\vec{\beta}_1, \ldots, \vec{\beta}_l}(x, y) := \left\| \nabla_2^l \log \left(\mathbb{I}_{T_{\text{ref}}} u(x), u(y) \right) \right\| \prod_{j=1}^l \left| \mathcal{D}^{\vec{\beta}_j} u(y) \right|_{g(u(y))}.
$$

Step 5 In the appendix we have collected a few estimates for remainder terms of Taylor series. We can use Lemma A.1 for the functions $H^{\vec{\beta}_1, \ldots, \vec{\beta}_l}$, which gives us the bound

$$
\int_{T_{\text{ref}}} |\varepsilon_i(x)|^2_{g(\mathbb{I}_{T_{\text{ref}}} u(x))} \, dx \lesssim \sum_{\substack{1 \le l \le k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \int_{T_{\text{ref}}} \sup_{x \in T_{\text{ref}}} |H^{\vec{\beta}_1, \ldots, \vec{\beta}_l}(x, y)|^2_{g(u(y))} \, dy
$$

$$
\le \sup_{1 \le l \le k} \sup_{\substack{p \in \mathbb{I}_{T_{\text{ref}}} u(T_{\text{ref}}) \\ q \in u(T_{\text{ref}})}} \left\| \nabla_2^l \log(p, q) \right\|^2 \sum_{\substack{1 \le l \le k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \int_{T_{\text{ref}}} \prod_{j=1}^l \left| \mathcal{D}^{\vec{\beta}_j} u(x) \right|^2_{g(u(x))} \, dx
$$

$$
= \mathcal{C}^2_{1,u}(T_{\text{ref}}) \dot{\Theta}_{k, T_{\text{ref}}}(u)^2.
$$

This concludes the first part of the proof.

Step 6 We now turn to the estimate for the first derivatives. To that end we need to bound the L^2-norm of $\frac{D}{dx^\alpha} \varepsilon(x)$. Since the functions λ_i have uniformly bounded first derivatives on T_{ref}, by the product rule, we can further reduce the problem to bounding the L^2-norm of

$$
\frac{D}{dx^\alpha} \sum_{|\vec{k}|=k} \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}|-1} \partial_y^{\vec{k}} G(x, tx) x^{\vec{k}} \, dt
$$

for $\alpha = 1, \ldots, d$.

Using the chain rule we get

$$
\frac{D}{dx^\alpha} \sum_{|\vec{k}|=k} \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}|-1} \partial_y^{\vec{k}} G(x, tx) x^{\vec{k}} \, dt = \text{I}(x, x) + \text{II}(x),
$$

with

$$\mathrm{I}(z,x) := \sum_{|\vec{k}|=k} \frac{d}{dx^\alpha} \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}|-1} \partial_y^{\vec{k}} G(z,tx) x^{\vec{k}} \, dt, \qquad z \in T_{\mathrm{ref}}$$

and

$$\mathrm{II}(x) := \sum_{|\vec{k}|=k} \frac{|\vec{k}|}{\vec{k}!} \int_0^1 t^{|\vec{k}|-1} \partial_y^{\vec{k}} \frac{D}{dx^\alpha} G(x,z) \Big|_{z=tx} x^{\vec{k}} \, dt.$$

Step 7 To bound the L^2-norm of II we may again use Lemma A.1 and proceed exactly in the same fashion as for the proof of (26) in **Step 5** above. More precisely, by the chain rule we can bound

$$\left| \partial_y^{\vec{k}} \frac{D}{dx^\alpha} G(x,z) \Big|_{z=y} \right|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} \lesssim$$

$$\sum_{\substack{1 \le l \le k, \ \vec{\beta}_j \in [d]^{m_j} \\ \sum_{j=1}^l m_j = k}} \left| \nabla_2^l \nabla_1 \log \left(\mathbb{I}_{T_{\mathrm{ref}}} u(x), u(y) \right) \left(\frac{d}{dx^\alpha} \mathbb{I}_{T_{\mathrm{ref}}} u(x), \mathcal{D}^{\vec{\beta}_1} u(y), \dots, \mathcal{D}^{\vec{\beta}_l} u(y) \right) \right|_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))},$$

where we recall that ∇_1 denotes the covariant derivative of the vector field $p \mapsto \log(p,q)$. Using that $\mathbb{I}_{T_{\mathrm{ref}}} u$ has uniformly bounded derivatives, and arguing exactly as in **Step 4** and **Step 5**, we obtain

$$\int_{T_{\mathrm{ref}}} |\mathrm{II}(x)|^2_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} \, dx \lesssim \sup_{1 \le l \le k} \sup_{\substack{p \in \mathbb{I}_{T_{\mathrm{ref}}} u(T_{\mathrm{ref}}) \\ q \in u(T_{\mathrm{ref}})}} \left\| \nabla_2^l \nabla_1 \log(p,q) \right\|^2 \dot{\Theta}_{k,T_{\mathrm{ref}}}(u)^2 = \mathcal{C}_{2,u}^2(T_{\mathrm{ref}}) \dot{\Theta}_{k,T_{\mathrm{ref}}}(u)^2.$$

Step 8 The bound for $\mathrm{I}(z,x)$ is more subtle. At first sight it looks as if a bound for $\mathrm{I}(z,x)$ would require derivatives of order $k+1$ of u, which may not be available. However, by Lemma A.2 applied to the function $U(\cdot) := G(z,\cdot) : T_{\mathrm{ref}} \to T_{\mathbb{I}_{T_{\mathrm{ref}}} u(z)} M$ for every fixed $z \in T_{\mathrm{ref}}$, we can write

$$\mathrm{I}(z,x) = \sum_{|\vec{l}|=k-1} \frac{(-1)^{k-1}}{\vec{l}!} x^{\vec{l}} \partial_y^{\vec{l}+\vec{e}_\alpha} G(z,x),$$

which contains only derivatives of the desired order k. Here, for any $\alpha \in \{1, \dots, d\}$, we use the notation $\vec{e}_\alpha \in \mathbb{N}_0^d$ for the unit vector which is 1 in its α-th digit and 0 everywhere else.

Now we can proceed as above in **Step 3** (using Lemma A.1) to show that

$$\int_{T_{\mathrm{ref}}} |\mathrm{I}(x,x)|^2_{g(\mathbb{I}_{T_{\mathrm{ref}}} u(x))} \, dx \lesssim \sup_{1 \le l \le k} \sup_{\substack{p \in \mathbb{I}_{T_{\mathrm{ref}}} u(T_{\mathrm{ref}}) \\ q \in u(T_{\mathrm{ref}})}} \left\| \nabla_2^l \log(p,q) \right\|^2 \dot{\Theta}_{k,T_{\mathrm{ref}}}(u)^2$$

which proves (27).

$\square$

The previous theorem has bounded the interpolation error on the reference element. We now derive an estimate on a general element T.

Theorem 5.3. *With the same assumptions on p,k as in Theorem 5.2 and $T, \mathcal{F}_T : T \to T_{ref}$ as in Lemma 2.8, we have the element-wise estimate*

$$D_{1,2}(\mathbb{I}_T u, u) \lesssim h^{k-1} C_{M,T} \Theta_{k,T}(u),$$

with

$$C_{M,T}(u) := \mathcal{C}_{1,u}(T) + \mathcal{C}_{2,u}(T). \tag{32}$$

The implicit constant is independent of M and only depends on the basis functions λ_i.

Proof. We use the representation

$$\mathbb{I}_T u = \mathbb{I}_{T_{\text{ref}}} \left(u \circ \mathcal{F}_T^{-1} \right) \circ \mathcal{F}_T.$$

As a first step we prove the desired estimate for the L^2-part. Putting $v := u \circ \mathcal{F}_T^{-1} : T_{\text{ref}} \to M$, $y = \mathcal{F}_T(x)$, and using (15b) we get

$$\int_T |\log(\mathbb{I}_T u(x), u(x))|^2_{g(\mathbb{I}_T u(x))} \, dx = \int_{T_{\text{ref}}} |\log(\mathbb{I}_{T_{\text{ref}}} v(y), v(y))|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(y))} \left| \det \left(\nabla \mathcal{F}_T(\cdot) \right) \right|^{-1} \circ \mathcal{F}_T^{-1}(y) \, dy$$
$$\leq h^d \int_{T_{\text{ref}}} |\log(\mathbb{I}_{T_{\text{ref}}} v(y), v(y))|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(y))} \, dy.$$

By Theorem 5.2 we can further estimate

$$h^d \int_{T_{\text{ref}}} |\log(\mathbb{I}_{T_{\text{ref}}} v(y), v(y))|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(y))} \, dy \lesssim h^d \mathcal{C}_{1,u}(T)^2 \dot{\Theta}_{k,T_{\text{ref}}}(v)^2.$$

Finally we use Lemma 2.8 and the definition of v to arrive at

$$h^d \mathcal{C}_{1,u}(T)^2 \dot{\Theta}_{k,T_{\text{ref}}}(v)^2 \lesssim h^{2k} \mathcal{C}_{1,u}(T)^2 \Theta_{k,T}(u)^2.$$

Hence we have shown the L^2-part of the assertion.

We go on to estimate the quantity

$$\int_T \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_T u(x), u(x)) \right|^2_{g(\mathbb{I}_T u(x))} \, dx$$

for an $\alpha \in \{1, \ldots, d\}$. The chain rule yields that

$$\int_T \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_T u(x), u(x)) \right|^2_{g(\mathbb{I}_T u(x))} \, dx$$
$$\leq \int_T \left[\left| \frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{\text{ref}}} v(\cdot), v(\cdot)) \right|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(\cdot))} \circ \mathcal{F}_T(x) \right] \cdot \left| \frac{d}{dx^\alpha} \mathcal{F}_T(x) \right|^2 \, dx.$$

Now we use the scaling assumption (15c) for the term $\left| \frac{d}{dx^\alpha} \mathcal{F}_T(x) \right|$ to bound the previous quantity by

$$h^{-2} \int_T \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{\text{ref}}} v(\cdot), v(\cdot)) \right|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(\cdot))} \circ \mathcal{F}_T(x) \, dx.$$

We can now again use the substitution $y = \mathcal{F}_T(x)$ and, using (15b), get the bound

$$h^{-2} \int_T \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{\text{ref}}} v(\cdot), v(\cdot)) \right|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(\cdot))} \circ \mathcal{F}_T(x) \, dx \leq h^{d-2} \int_{T_{\text{ref}}} \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{\text{ref}}} v(y), v(y)) \right|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(y))} \, dy.$$

Now we can again invoke Theorem 5.2 to deduce the estimate

$$h^{d-2} \int_{T_{\text{ref}}} \left| \frac{D}{dx^\alpha} \log(\mathbb{I}_{T_{\text{ref}}} v(y), v(y)) \right|^2_{g(\mathbb{I}_{T_{\text{ref}}} v(y))} \, dy \lesssim h^{d-2} \mathcal{C}_{2,u}(T)^2 \dot{\Theta}_{k,T_{\text{ref}}}(v)^2.$$

Finally, applying Lemma 2.8 to $\dot{\Theta}_{k,T_{\text{ref}}}(v)$ yields the desired bound. $\qquad\square$

5.2 Global Interpolation Error Bounds

We now use Theorem 5.3 to obtain a global approximation result. The necessary grid regularity is formalized in the following definition.

23

Definition 5.1. *We say that a grid $\mathcal{G}$ is of width h if for each element T of $\mathcal{G}$ the map $\mathcal{F}_T$ from T to its reference element scales with h (of order p, where p is the order of the Lagrange shape functions used in the construction of the GFE spaces).*

A particular instance of such grids are shape regular triangulations with element diameters of the order of h. However, the definition also covers more general cases, such as grids where the $\mathcal{F}_T$ are polynomials.

Theorem 5.4. *Let Ω be a domain with a conforming grid $\mathcal{G}$ of width h, and let $\mathbb{I}_\mathcal{G}$ be the pointwise interpolation operator onto the space of p-th order geodesic finite elements on $\mathcal{G}$. If $k > d/2$ and $p \geq k-1$ we have the estimate*

$$D_{1,2}(\mathbb{I}_\mathcal{G} u, u) \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{k,\Omega}(u) \tag{33}$$

with

$$C_{M,\mathcal{G}}(u) := \sup_{T \in \mathcal{G}} C_{M,T}(u),$$

and $C_{M,T}(u)$ as defined in (32). The implicit constants are independent of M and only depend on the shape functions λ_i. For $h \to 0$ the constant $C_{M,\mathcal{G}}(u)$ approaches the limit

$$\lim_{h \to 0} C_{M,\mathcal{G}}(u) = \sup_{1 \leq l \leq k} \sup_{q \in u(\Omega)} \left\| \nabla_2^l \log(q,q) \right\| + \sup_{1 \leq l \leq k} \sup_{q \in u(\Omega)} \left\| \nabla_2^l \nabla_1 \log(q,q) \right\|. \tag{34}$$

Proof. The bound (33) follows from applying Theorem 5.3 element-wise and summing up. To show (34), note that u is uniformly continuous, because $k > d/2$. Therefore, the sets $u(T), \mathbb{I}_T u(T)$ converge to single points as h goes to zero. $\square$

The error estimates of Theorem 5.4 assess the error between a function and its Lagrange interpolant whenever the given function is of smoothness $k > d/2$. In particular, in three dimensions our results require that $u \in W^{k,2}(\Omega, M)$ with $k > 3/2$. The same requirement is needed for the linear theory, since as a minimal requirement to define the Lagrange interpolant an embedding into continuous functions is needed.

However, numerical experiments (not detailed here) indicate nevertheless optimal approximation properties of both linear and geodesic finite element spaces even for $k \leq d/2$. In the linear setting, this stronger result is proved using the Clément interpolation operator [10]. A generalization of this technique to nonlinear finite element spaces would be interesting.

Remark 5.2. For the linear case $M = \mathbb{R}$ we have $\log(q_1, q_2) = q_2 - q_1$,

$$\nabla_2^l \log(q_1, q_2) = \begin{cases} 1 & \text{if } l = 1 \\ 0 & \text{if } l > 1, \end{cases} \qquad \text{and} \qquad \nabla_2^l \nabla_1 \log(q_1, q_2) = 0 \qquad l \geq 1.$$

Therefore

$$C_{M,\mathcal{G}} = 1$$

for any grid $\mathcal{G}$ of size h.

We also remark that the same argument as the one in Theorem 5.4 also allows to obtain error estimates in terms of

$$D_{1,\infty}(v,w) := \sup_{x \in \Omega} \log(v(x), w(x)) + \sum_{\alpha=1}^d \sup_{x \in \Omega} \frac{D}{dx^\alpha} \log(v(x), w(x))$$

and $\Theta_{\infty,k,\Omega}$. Note that in this case no relation between d and k is required, since pointwise interpolation is defined in $W^{k,\infty}(\Omega, M)$ for all $k > 0$. The proof proceeds as the one for Theorem 5.4, except that the remainder terms occurring in the proof of Theorem 5.2 (e.g., in **Step 2**) have to be estimated in the ∞-norm (this is actually simpler than the L^2 norm bounds).

Theorem 5.5. *Let Ω be a domain with a conforming grid $\mathcal{G}$ of width h and let $p \geq k-1$. Then we have the estimates*

$$D_{1,\infty}(\mathbb{I}_\mathcal{G} u, u) \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{\infty,k,\Omega}(u)$$

and

$$\text{dist}_{L^\infty}(\mathbb{I}_\mathcal{G} u, u) \lesssim h^k C_{M,\mathcal{G}}(u) \Theta_{\infty,k,\Omega}(u).$$

The implicit constants are independent of M and only depend on the shape functions λ_i.

Additionally, we obtain the following stability of the pointwise interpolation operator.

Corollary 5.1. *There exists a constant C_4 which only depends on M and the shape functions λ_i (but not on h) such that*

$$\Theta_{\infty,1,\Omega}\left(\mathbb{I}_{\mathcal{G}}u\right) \leq C_4\Theta_{\infty,1,\Omega}\left(u\right).$$

Proof. We assume that our manifold M is smoothly embedded into $\mathbb{R}^N$. With the ansatz and notation of the proof of Lemma 2.3 we obtain

$$\frac{d}{dx^\alpha}u(x) = \partial_1\exp\left(\mathbb{I}_{\mathcal{G}}u(x), \log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right)\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)$$
$$+ \partial_2\exp\left(\mathbb{I}_{\mathcal{G}}u(x), \log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right)\frac{D}{dx^\alpha}\log(\mathbb{I}_{\mathcal{G}}u(x), u(x))$$

for all $\alpha \in \{1,\ldots,d\}$, which implies that

$$\left|\partial_1\exp\left(\mathbb{I}_{\mathcal{G}}u(x), \log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right)\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right| \lesssim \left|\frac{d}{dx^\alpha}u(x)\right| + \left|\frac{D}{dx^\alpha}\log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right| \tag{35}$$

for all $\alpha \in \{1,\ldots,d\}$. Now we use the fact that

$$\partial_1\exp\left(\mathbb{I}_{\mathcal{G}}u(x), 0\right)\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x) = \frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x),$$

together with the Lipschitz continuity of $\partial_1\exp(p,w)$ in w to get that, up to a constant C independent of h,

$$\left|\partial_1\exp\left(\mathbb{I}_{\mathcal{G}}u(x), \log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right)\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x) - \frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right| \leq C\left|\log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right|\left|\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right|.$$

By Theorem 5.5 we can further bound

$$|\log(\mathbb{I}_{\mathcal{G}}u(x), u(x))| \leq \operatorname{dist}_{L^\infty}(u, \mathbb{I}_{\mathcal{G}}u) \leq Dh\dot{\Theta}_{\infty,1,\Omega}(u)$$

with another constant $D > 0$ independent of h. Putting these estimates together we obtain that

$$\left|\partial_1\exp\left(\mathbb{I}_{\mathcal{G}}u(x), \log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right)\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x) - \frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right| \lesssim hCD\left|\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right| \tag{36}$$

for all $\alpha \in \{1,\ldots,d\}$. Putting together (35) and (36) we obtain

$$(1 - hCD)\left|\frac{d}{dx^\alpha}\mathbb{I}_{\mathcal{G}}u(x)\right| \lesssim \left|\frac{d}{dx^\alpha}u(x)\right| + \left|\frac{D}{dx^\alpha}\log(\mathbb{I}_{\mathcal{G}}u(x), u(x))\right|,$$

which by Theorem 5.5 implies the desired result. $\square$

Remark 5.3. One can generalize these results in terms of the shape functions which are used in the construction for the GFE spaces. Indeed, all approximation error estimates in the present section only use the property that the Lagrange shape functions λ_i are exact on polynomials (24). Therefore the same proofs can be used for any such set of shape functions.

5.3 Retraction Pairs

In certain cases it is computationally expensive to compute the exponential or logarithm function of a given manifold. Then, alternative functions can sometimes be used. This idea is formalized by the concept of *retraction pairs*.

Definition 5.2 (Grohs [23], see also [1, 20])**.** *A pair (P, Q) of smooth functions*
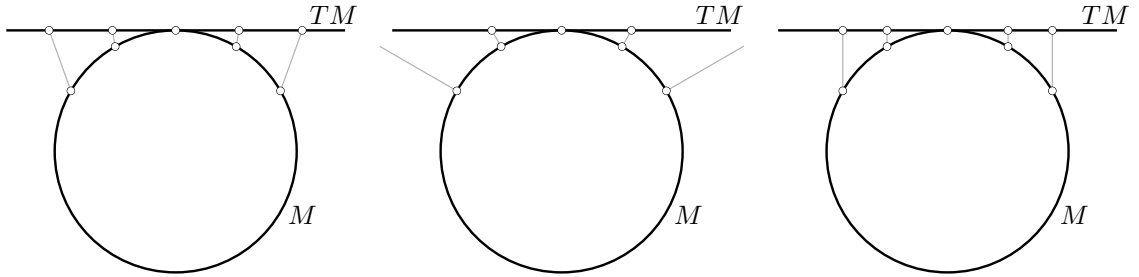
$$P : TM \to M, \quad Q : M \times M \to TM$$

is called a retraction pair *if*

$$P\left(x, Q\left(x, y\right)\right) = y, \qquad \text{for all } x,\ y \in M,$$

and

$$P\left(x, 0\right) = x, \quad \left.\frac{d}{dv}P(x,v)\right|_{v=0} = \operatorname{Id} \qquad \text{for all } x \in M.$$

(a) Retraction pair based on exponential map   (b) Retraction pair based on closest point projection   (c) Retraction pair based on vertical projection

Figure 1: Different retraction pairs for the circle

In general P may only be defined locally around M, and Q around the diagonal of $M \times M$. Certainly, the pair $(\exp, \log)$ satisfies the above assumptions [12], and therefore forms a retraction pair. We refer to [1] for examples of retraction pairs for several manifolds of practical interest. To better illustrate the concept of retraction pairs, Figure 1 shows different pairs for the circle S^1.

Given a retraction pair (P, Q), we can construct generalized geodesic finite elements by using interpolants $\Upsilon^{(P,Q)}$ based on the first order condition (23)

$$\sum_{i=1}^{m} \lambda_i^p(x) \, Q\left(\Upsilon^{(P,Q)}(v_1, \ldots, v_m; x), v_i\right) = 0 \; \in \; T_{\Upsilon^{(P,Q)}(v_1, \ldots, v_m; x)}M. \qquad (37)$$

The results in [23] show that this expression is locally well-defined. We state the following theorem whose correctness can be easily verified by going through the proofs of the results in Section 5.

Theorem 5.6. *All approximation results shown in Section 5 remain valid if we replace the definition of the interpolant Υ by (37) with (P, Q) an arbitrary retraction pair, provided that the function u to be approximated is in $W^{1,\infty}(\Omega, M)$.*

The details are left to the reader.

6 A Priori Error Estimates for Geodesic Finite Elements

We are now in a position to combine the nonlinear Céa Lemma (Theorem 3.3) with the approximation result (Theorem 5.4) to arrive at an a priori error bound for variational problems. For later use we include Dirichlet boundary conditions, and define

$$H^\Phi := \left\{ v \in W^{1,2}(\Omega, M) : \; u|_{\partial\Omega} = \Phi \right\}$$

for a given function $\Phi : \partial\Omega \to M$. Note, however, that all results in this chapter also hold without Dirichlet conditions if the functional $\mathfrak{J}$ has the appropriate ellipticity properties.

We also put

$$H_{K,L}^u(\Omega, M) := H_K(\Omega, M) \cap H_L^u(\Omega, M),$$

where H_K, H_L^u are defined as in Theorem 3.3 for $u \in C(\Omega, M)$, and $K, L > 0$.

We first show a direct consequence of Theorems 3.3 and 5.4. Then we give an alternative proof showing the same optimal error bounds under weaker assumptions on the approximation space.

Theorem 6.1. *Let u be a stationary point of the energy $\mathfrak{J} : H^\Phi \to \mathbb{R}$ w.r.t. variations along geodesic homotopies starting in u, and assume that*

$$u \in H^k(\Omega, M) \cap W^{1,\infty}(\Omega, M)$$

26

for some $k > d/2$.

Let C_4 be the constant from Corollary 5.1, and pick a second constant

$$K \geq C_4 \Theta_{\infty,1,\Omega}(u). \tag{38}$$

With this constant K, and $L > 0$ arbitrary, assume that $\mathfrak{J}$ is elliptic on $H^u_{K,L} \cap H^\Phi$ along geodesic homotopies that start in u.

Let $\mathcal{G}$ be a grid for Ω of width h and order p, $V^M_{p,\mathcal{G}}$ a p-th order GFE space as defined in Section 4, and set

$$V^h := V^M_{p,\mathcal{G}} \cap H^u_{K,L} \cap H^\Phi.$$

Assume that Φ is such that this space is not empty. Finally, denote

$$u^h := \arg\min_{v^h \in V^h} \mathfrak{J}(v^h).$$

Then, whenever $p \geq k - 1$, we have the a priori estimates

$$\|u - u^h\|_{W^{1,2}(\Omega,M)} \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{k,\Omega}(u)$$

(with respect to some embedding) and

$$\mathrm{dist}_{W^{1,2}}(u, u^h) \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{k,\Omega}(u).$$

In these estimates, the implicit constants only depend on d, the ellipticity constants of $\mathfrak{J}$ on $H^u_{K,L} \cap H^\Phi$, the interpolation functions λ_i, $i = 1, \ldots, m$, and the geometry of M.

Proof. Consider the p-th order interpolant $\mathbb{I}_\mathcal{G} u \in V^M_{p,\mathcal{G}}$ of u. By the choice (38), Corollary 5.1, and the assumption on the boundary data, we obtain that $\mathbb{I}_\mathcal{G} u \in H^u_{K,L} \cap H^\Phi$. We can therefore apply the Céa lemma (Theorem 3.3) to get

$$D_{1,2}(u, u^h) \leq C_2^2 \sqrt{\frac{\Lambda}{\lambda}} D_{1,2}(u, \mathbb{I}_\mathcal{G} u),$$

with λ, Λ the ellipticity constants, and C_2 depending only on d, the product KL and the curvature of M. By Theorem 5.4, the term $D_{1,2}(u, \mathbb{I}_\mathcal{G} u)$ is less than $h^{k-1} C_{M,\mathcal{G}} \Theta_{k,\Omega}(u)$ times another constant depending only on the λ_i. On the other hand, Lemmas 2.3 and 2.4 bound $D_{1,2}(u, u^h)$ from below by $\|u - u^h\|_{W^{1,2}(\Omega,M)}$ and $\mathrm{dist}_{W^{1,2}}(u, u^h)$, respectively. Together the assertion follows. $\square$

Theorem 6.1 requires that the discrete solution u^h is obtained by minimizing the energy $\mathfrak{J}$ over the approximation space $V^M_{p,\mathcal{G}} \cap H^u_{K,L} \cap H^\Phi$. The restriction to $H^u_{K,L}$ (i.e., the requirement that the first derivatives of all functions in the approximation space are uniformly bounded by K) is not usually encountered in the geometrically linear theory. It is problematic because the first derivatives of GFE functions deteriorate with decreasing mesh size (Lemma 4.3). In the next theorem we will show that we can dispense with K provided that $u \in H^k$ with k sufficiently large. We do not know whether that result also holds without the additional restriction on k.

Theorem 6.2. *Let u be a stationary point of the energy $\mathfrak{J} : H^\Phi \to \mathbb{R}$ w.r.t. variations along geodesic homotopies starting in u, and assume that*

$$u \in H^k(\Omega, M), \qquad k > \max\left(3, \frac{3}{2}d\right).$$

Suppose that $\mathfrak{J}$ is elliptic along geodesic homotopies starting in u, with ellipticity constants λ, Λ, where, for a geodesic homotopy from u to v, the upper bound Λ may depend on $\max(\Theta_{1,\infty,\Omega}(u), \Theta_{1,\infty,\Omega}(v))$.

Let $\mathcal{G}$ be a grid of width h, and $V^M_{p,\mathcal{G}}$ a p-th order GFE space. Denote $V^h := V^M_{p,\mathcal{G}} \cap H^u_L \cap H^\Phi$ (assuming again that Φ is such that V^h is not empty) with $L > 0$ arbitrary (for M compact set $L = \infty$). Define the discrete minimizer

$$u^h := \arg\min_{v^h \in V^h} \mathfrak{J}(v^h).$$

Then, whenever $p \geq k - 1$, we have the a priori estimates

$$\|u - u^h\|_{W^{1,2}(\Omega,M)} \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{k,\Omega}(u)$$

(with respect to some embedding) and

$$\mathrm{dist}_{W^{1,2}}(u, u^h) \lesssim h^{k-1} C_{M,\mathcal{G}}(u) \Theta_{k,\Omega}(u).$$

Proof. For simplicity we will tacitly assume that the manifold M is embedded into $\mathbb{R}^N$. We proceed in several steps.

Step 1 Using the argument from the proof of Theorem 3.3 we can show that

$$D_{1,2}(u, u^h)^2 \leq \frac{C_2^2}{\lambda} \left(\mathfrak{J}(u^h) - \mathfrak{J}(u) \right) \leq \frac{C_2^2}{\lambda} \left(\mathfrak{J}(\mathbb{I}_\mathcal{G} u) - \mathfrak{J}(u) \right) \leq \frac{C_2^2 \Lambda(u, \mathbb{I}_\mathcal{G} u)}{\lambda} D_{1,2}(u, \mathbb{I}_\mathcal{G} u)^2.$$

The constant C_2 is the one given in (9), and the K appearing there has to be interpreted as an upper bound on $\Theta_{\infty,1,\Omega}$ on the geodesic homotopy from u to u^h.

By Lemma 4.3 we can pick the K such that $\Theta_{\infty,1,\Omega}(w^h) \lesssim K \lesssim h^{-1}$ for any $w^h \in V^h$, where the implicit constant depends on the nodal values of w^h. We therefore obtain

$$D_{1,2}(u, u^h)^2 \lesssim h^{-2} \frac{\Lambda(u, \mathbb{I}_\mathcal{G} u)}{\lambda} D_{1,2}(u, \mathbb{I}_\mathcal{G} u)^2$$

for h small.

Now we use that by our smoothness assumptions $\Theta_{\infty,1,\Omega}(u)$ is finite. Then, by Corollary 5.1 we get

$$\Theta_{\infty,1,\Omega}(\mathbb{I}_\mathcal{G} u) \lesssim \Theta_{\infty,1,\Omega}(u),$$

and the constant is independent of h. Therefore also the quantity $\Lambda(u, \mathbb{I}_\mathcal{G} u)$ is uniformly bounded, independent of h. Using additionally Theorem 5.4 this gives

$$D_{1,2}(u, u^h) \lesssim h^{-1} D_{1,2}(u, \mathbb{I}_\mathcal{G} u) \lesssim h^{k-2} C_{M,\mathcal{G}} \Theta_{k,\Omega}(u), \tag{39}$$

where we have omitted the dependence on the ellipticity constants.

Using Lemma 2.3 we see that that (39) implies that in our embedding we have

$$\|u - u^h\|_{H^1} \lesssim h^{k-2} C_{M,\mathcal{G}} \Theta_{k,\Omega}(u). \tag{40}$$

We need to improve this estimate to the desired order $k - 1$.

Step 2 We now put
$$\lambda_f(t) := \mu\left(\{x \in \Omega : |f(x)| > t\} \right)$$

for any measurable function $f : \Omega \to \mathbb{R}^N$, $t \geq 0$, and μ denoting the Lebesgue measure. This quantity satisfies the well-known Markov inequality

$$\lambda_f(t) \leq \frac{\|f\|_{L^q}^q}{t^q}, \quad t > 0, \quad q \in [1, \infty],$$

see, e.g., [19]. Together with (40) it implies that

$$\lambda_{u-u^h}(h) \lesssim h^{q(k-3)} C_{M,\mathcal{G}} \Theta_{k,\Omega}(u), \tag{41}$$

whenever $H^1(\Omega, \mathbb{R}^N)$ is embedded in $L^q(\Omega, \mathbb{R}^N)$. For the moment, let us assume that we have such an embedding with $q \geq 2$.

Define the set
$$\Omega_h := \left\{ x \in \Omega : |u(x) - u^h(x)| \leq h \right\},$$

which is well-defined by the smoothness of u. Equation (41) then states that

$$\mu(\Omega \setminus \Omega_h) \lesssim h^{q(k-3)} C_{M,\mathcal{G}} \Theta_{k,\Omega}(u). \tag{42}$$

Step 3 Let $\Gamma(t)$ be the geodesic homotopy connecting u and u^h. With the Finsler norm $|\cdot|_{H^1}$ from Definition 2.4 we can write
$$|\dot{\Gamma}(t)|_{H^1}^2 = |\dot{\Gamma}(t)|_{G^b}^2 + |\dot{\Gamma}(t)|_{G^u}^2,$$

where

$$|\dot{\Gamma}(t)|^2_{G^b} := \int_{\Omega_h} |\dot{\Gamma}(x,t)|^2_g \, dx + \int_{\Omega_h} |\nabla_x \dot{\Gamma}(x,t)|^2_g \, dx$$

and

$$|\dot{\Gamma}(t)|^2_{G^u} := \int_{\Omega \setminus \Omega_h} |\dot{\Gamma}(x,t)|^2_g \, dx + \int_{\Omega \setminus \Omega_h} |\nabla_x \dot{\Gamma}(x,t)|^2_g \, dx.$$

Here we have used the shorthand notation $|\nabla_x \dot{\Gamma}(x,t)|^2_g$ introduced in (4).

We now bound the two terms separately. First, note that by construction of Ω_h we have

$$\sup_{x \in \Omega_h} \mathrm{dist}(u(x), u^h(x)) \lesssim h,$$

with a constant that depends only on u and the geometry of M. Also, we have

$$\Theta_{\infty,1,\Omega_h}(u), \; \Theta_{\infty,1,\Omega_h}(u^h) \lesssim h^{-1}.$$

Lemma 2.5 implies that

$$|\dot{\Gamma}(0)|^2_{G^b} = 2 \int_0^1 (1-t)|\dot{\Gamma}(0)|^2_{G^b} \, dt \leq$$

$$2 \left(\sqrt{2} + 2^{d/2} \|\mathrm{Rm}\|_g C_3 \max(\Theta_{\infty,1,\Omega_h}(u), \Theta_{\infty,1,\Omega_h}(u^h)) \, \mathrm{dist}_{L^\infty}(u|_{\Omega_h}, u^h|_{\Omega_h}) \right)^2 \int_0^1 (1-t)|\dot{\Gamma}(t)|^2_{G^b} \, dt.$$

Since $\max(\Theta_{\infty,1,\Omega_h}(u), \Theta_{\infty,1,\Omega_h}(u^h)) \lesssim h^{-1}$ and $\mathrm{dist}_{L^\infty}(u|_{\Omega_h}, u^h|_{\Omega_h}) \lesssim h$ we get that

$$\sqrt{2} + 2^{d/2} \|\mathrm{Rm}\|_g C_3 \max(\Theta_{\infty,1,\Omega_h}(u), \Theta_{\infty,1,\Omega_h}(u^h)) \, \mathrm{dist}_{L^\infty}(u|_{\Omega_h}, u^h|_{\Omega_h}) \lesssim 1,$$

and consequently

$$|\dot{\Gamma}(0)|^2_{G^b} \lesssim \int_0^1 (1-t)|\dot{\Gamma}(t)|^2_{G^b} \, dt \leq \int_0^1 (1-t)|\dot{\Gamma}(t)|^2_{H^1} \, dt. \tag{43}$$

By the ellipticity assumption on $\mathfrak{J}$ it follows that

$$\int_0^1 (1-t)|\dot{\Gamma}(t)|^2_{H^1} \leq \lambda^{-1} \left(\mathfrak{J}(u^h) - \mathfrak{J}(u) \right) \lesssim \mathfrak{J}(\mathbb{I}_{\mathcal{G}} u) - \mathfrak{J}(u).$$

Using the ellipticity of $\mathfrak{J}$ a second time we deduce that

$$\mathfrak{J}(\mathbb{I}_{\mathcal{G}} u) - \mathfrak{J}(u) \lesssim D_{1,2}(u, \mathbb{I}_{\mathcal{G}} u)^2 \lesssim h^{2(k-1)},$$

where we have used Theorem 5.4 in the last inequality. Using the previous two estimates together with (43) yields the bound

$$|\dot{\Gamma}(0)|^2_{G^b} \lesssim h^{2(k-1)}. \tag{44}$$

Note that this is the right order of magnitude compared to the suboptimal bound (39). The reason for this improvement is that the quantity $|\dot{\Gamma}(0)|^2_{G^b}$ is only defined on Ω_h, where we have $\mathrm{dist}_{L^\infty}(u^h, u) \leq h$. Therefore, if we use Lemma 2.5 in Equation (43) with $K = \max\left(\Theta_{1,\infty,\Omega}(u), \Theta_{1,\infty,\Omega}(u^h) \right) \lesssim h^{-1}$ and $L = \mathrm{dist}_\infty(u^h|_{\Omega_h}, u|_{\Omega_h}) \lesssim h$, the product KL is uniformly bounded, independent of h. Since the constant C_2 defined in (9) only depends on the product KL, we gain one order of magnitude in (43).

Step 4 We still need to control the quantity $|\dot{\Gamma}(0)|^2_{G^u}$ to get the desired bound for $D_{1,2}(u, u^h)^2 = |\dot{\Gamma}(0)|^2_{G^b} + |\dot{\Gamma}(0)|^2_{G^u}$. This is not difficult. By

$$\sup_{x \in \Omega} |\nabla_x \dot{\Gamma}(x,t)|_g \lesssim h^{-1},$$

and (42) we have

$$|\dot{\Gamma}(t)|^2_{G^u} \lesssim h^{q(k-3)-2} \tag{45}$$

for all $t \in [0,1]$.

Step 5 Putting together Equations (44) and (45) we arrive at

$$D_{1,2}(u, u^h)^2 = |\dot\Gamma(0)|^2_{H^1} = |\dot\Gamma(0)|^2_{G^b} + |\dot\Gamma(0)|^2_{G^u} \lesssim h^{\min(2(k-1), q(k-3)-2)},$$

whenever $H^1(\Omega, \mathbb{R}^N)$ embeds into $L^q(\Omega, \mathbb{R}^N)$. By the Sobolev embedding theorem this holds for all q smaller than $\frac{2d}{d-2}$ [9]. Picking k so that

$$\min(2(k-1), q(k-3)-2) = 2(k-1)$$

yields the result. $\square$

To summarize, Theorems 6.1 and 6.2 both present extensions of linear a priori error estimates for finite elements. In Theorem 6.1 we require that the approximation spaces and the solution u possess uniformly bounded derivatives. In contrast, Theorem 6.2 does not impose restrictions on the approximation spaces, but poses stronger assumptions on the smoothness of u instead.

7 Examples

To illustrate our results we apply them to a few specific examples. We focus on the harmonic energy and related functionals, and leave the study of more general energies to future work.

Let Ω be a domain and (M, g) a Riemannian manifold. As previously we consider Dirichlet problems only. Boundary values are given in form of a function $\Phi : \partial\Omega \to M$ of sufficient regularity. For such a Φ we write H^Φ for the set of all functions $v : \Omega \to M$ for which $v|_{\partial\Omega} = \Phi$ holds in the sense of traces.

Studying the assumptions of Theorems 6.1 and 6.2 we recall that we can give optimal a priori discretization error bounds for discrete minimizers of an energy functional $\mathfrak{J}$ if $\mathfrak{J}$ is elliptic, and if the minimizer of $\mathfrak{J}$ has sufficient smoothness.

7.1 Harmonic Maps

The prototypical elliptic functional is the harmonic energy

$$\mathfrak{J}^{\mathrm{harm}}(v) = \int_\Omega |\nabla v(x)|^2_{g(v(x))}\, dx.$$

The stationary points of this functional are called harmonic maps, and have been widely studied in the literature (see, e.g., [15]).

There are different approaches to showing ellipticity of the harmonic energy. We first use bounds on the second derivatives along geodesic homotopies. Let K be a positive constant, and H_K as defined in (5).

Lemma 7.1. *The energy $\mathfrak{J}^{\mathrm{harm}}$ is elliptic along geodesic homotopies in $H_K \cap H^\Phi$ in the sense of Definition 3.2 if either*

1. *M has nonpositive sectional curvature, or*

2. *we have $1 - K^2\|\mathrm{Rm}\|_g C_1(\Omega) > 0$,*

where $C_1(\Omega)$ is the Poincaré constant of Ω from Lemma 2.2.

Proof. Let Γ be a geodesic homotopy in $H_K \cap H^\Phi$, and set $f(t) := \mathfrak{J}^{\mathrm{harm}}(\Gamma(t))$. Lemma X.3.2(ii) in [46] tells us that

$$\begin{aligned}
\frac{d^2}{dt^2}f(t) = {}&2\int_\Omega \left\langle \nabla\dot\Gamma(x,t), \nabla\dot\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))}\, dx \\
&- 2\int_\Omega \left\langle \mathrm{Rm}\left(\nabla\Gamma(x,t), \dot\Gamma(x,t)\right)\dot\Gamma(x,t), \nabla\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))}\, dx.
\end{aligned}$$

Now the assertion follows as a direct consequence of the Poincaré inequality in Lemma 2.2. $\square$

Remark 7.1. For positive curvature this ellipticity result is fairly weak. The results in [32] may allow improvements.

Alternatively, one can also directly show the λ-convexity of the harmonic energy functional along geodesic homotopies.

Lemma 7.2. *Let M be simply connected and have nonpositive sectional curvatures. Then the harmonic energy is λ-convex along geodesic homotopies in H^Φ, with λ equal to $1/2$ times the Poincaré constant of Ω.*

Proof. Let u, v be functions in H^Φ, and let Γ be a geodesic homotopy from u to v. Since M is simply connected and has nonpositive curvature it is an NPC-space is the sense of [46, Sec. X.2.1]. For this setting, it is shown in the proof for [46, Thm. X.2.2] that

$$\mathfrak{J}^{\mathrm{harm}}(\Gamma(t)) \leq (1-t)\,\mathfrak{J}^{\mathrm{harm}}(u) + t\,\mathfrak{J}^{\mathrm{harm}}(v) - t(1-t)\int_\Omega |\nabla \operatorname{dist}(u(x), v(x))|^2\, dx.$$

Since u and v fulfill the same Dirichlet boundary conditions we have $\operatorname{dist}(u(x), v(x)) = 0$ on $\partial\Omega$. The assertion then follows with the standard Poincaré inequality. $\square$

Regularity of harmonic maps is a well-studied subject. The following results are derived in [15, 29, 32, 33].

Lemma 7.3. *A harmonic map $u : \Omega \to M$ with continuous boundary data is in C^∞, if one of the following conditions is satisfied:*

1. *M has nonpositive sectional curvature,*

2. *$d \in \{1, 2\}$, or*

3. *the image of u is contained in a convex geodesic ball.*

We remark that in other cases singularities may develop [41].

Using the preliminaries above and Theorem 6.2 we are able to prove the following convergence theorem for harmonic maps.

Theorem 7.1. *Let u be a local minimizer of $\mathfrak{J}^{\mathrm{harm}}$ on $H_K \cap H^\Phi$ for a constant $K > 0$ and continuous boundary data Φ. Also, let u^h be the corresponding minimizer in a p-th order GFE space generated by a grid of width h and order p, and resolving the boundary conditions. If M has positive sectional curvature suppose that $1 - K^2 \|\mathrm{Rm}\|_g C_1(\Omega) > 0$ and that either $d \in \{1, 2\}$ or that the image of u is contained in a convex geodesic ball of M (for M with nonpositive sectional curvature no assumptions are needed). Then*

$$\|u - u^h\|_{W^{1,2}(\Omega, M)} \lesssim h^p \|u\|_{H^{p+1}}^{p+1}$$

(in an embedding), and

$$\operatorname{dist}(u, u^h)_{W^{1,2}} \lesssim h^p \|u\|_{H^{p+1}}^{p+1}.$$

Proof. By Lemma 7.3, u is smooth enough for the smoothness descriptor $\Theta_{p+1,\Omega}(u)$ to be finite, and by Lemma 7.1 the harmonic energy $\mathfrak{J}^{\mathrm{harm}}$ is elliptic. Hence Theorem 6.2 yields bounds in terms of the smoothness descriptor $\Theta_{p+1,\Omega}(u)$, which we bound in turn with Lemma 2.7. $\square$

Remark 7.2. In Theorem 7.1 we have assumed that the boundary data can be represented exactly in the GFE approximation space. This may not always be the case, but a simple approximation argument shows that the same result holds if u^h is interpolating smooth boundary data.

Theorem 7.1 is confirmed by numerical studies in [44] for $M = S^2$. In [43], the same optimal orders were observed for $p = 1$, even though the assumptions of Lemma 7.3 did not hold there.

Remark 7.3. In [6] harmonic maps into spheres $S^2 \in \mathbb{R}^3$ are approximated by minimizing the harmonic energy over piecewise affine finite elements with nodal values on the sphere. It is shown that for $h \to 0$ there exists a subsequence of discrete solutions (more precisely stationary points of the discrete optimization problems) which converges weakly to a harmonic map. This holds even for non-regular solutions and without any ellipticity assumption, which is in contrast to our own results. The latter always assume a certain smoothness of the solution, but, on the other hand, allow to obtain not just weak convergence of a subsequence but strong convergence with optimal rates. We consider it an interesting question whether we can use the approach of Bartels [6] to prove weak convergence of sequences of GFE approximations when the solution is not smooth and/or the harmonic energy is not elliptic.

7.2   Generalizations

We can generalize the discretization error bounds for harmonic maps in a few simple ways. We show only the ellipticity of the different functionals. Together with regularity results available from the literature, optimal discretization error bounds then follow by Theorem 6.2.

7.2.1 F-Harmonic Maps

F-harmonic maps are stationary points of the energy

$$\mathfrak{J}^F(v) := \int_\Omega F\left(x, |\nabla v(x)|^2_{g(v(x))}\right) dx,$$

with a function $F : \Omega \times \mathbb{R}^+ \to \mathbb{R}$. Such energies generalize harmonic maps and include, e.g., p-harmonic maps and exponentially harmonic maps [5]. For notational simplicity we will suppress the dependence of F on x in the following results. The proofs for this case easily carry over to the x-dependent case.

The following result follows from direct calculations.

Lemma 7.4. *Denote*

$$f(t) := \mathfrak{J}^F(\Gamma(t)),$$

where Γ is a geodesic homotopy. Then we have

$$\frac{d^2}{dt^2}f(t) = 2\int_\Omega F'\left(|\nabla\Gamma(x,t)|^2_{g(\Gamma(x,t))}\right)\left\langle \nabla\dot\Gamma(x,t), \nabla\dot\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))} dx$$
$$- 2\int_\Omega F'\left(|\nabla\Gamma(x,t)|^2_{g(\Gamma(x,t))}\right)\left\langle \mathrm{Rm}\left(\nabla\Gamma(x,t), \dot\Gamma(x,t)\right)\dot\Gamma(x,t), \nabla\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))} dx$$
$$+ 4\int_\Omega F''\left(|\nabla\Gamma(x,t)|^2_{g(\Gamma(x,t))}\right)\left\langle \nabla\dot\Gamma(x,t), \nabla\Gamma(x,t)\right\rangle^2_{g(\Gamma(x,t))} dx.$$

Based on this we can prove the following ellipticity result.

Lemma 7.5. *Assume that there are constants w_2, w_2', w_3' such that*

$$w_2' \geq F'(y) \geq w_2 > 0, \quad w_3' \geq F''(y) \geq 0 \quad \forall y \in \mathbb{R}^+,$$

and either

1. *M has nonpositive sectional curvature, or*

2. *$w_2 - w_2' K^2 \|\mathrm{Rm}\|_g C_1(\Omega) > 0$,*

where $C_1(\Omega)$ is the Poincaré constant of Ω from Lemma 2.2. Then the energy $\mathfrak{J}^F$ is elliptic along geodesic homotopies in $H_K \cap H^\Phi$ in the sense of Definition 3.2.

Proof. Let M have nonpositive sectional curvature. Then, using Lemma 7.4, we have

$$\frac{d^2}{dt^2}f(t) \geq 2\int_\Omega F'\left(|\nabla\Gamma(x,t)|^2_{g(\Gamma(x,t))}\right)\left\langle \nabla\dot\Gamma(x,t), \nabla\dot\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))} dx$$
$$\geq 2w_2 \int_\Omega \left\langle \nabla\dot\Gamma(x,t), \nabla\dot\Gamma(x,t)\right\rangle_{g(\Gamma(x,t))} dx$$
$$\geq \frac{2w_2}{1 + C_1(\Omega)}|\dot\Gamma(t)|_{H^1},$$

where $C_1(\Omega)$ is the Poincaré constant of Ω from Lemma 2.2. On the other hand, again by Lemma 7.4, we have

$$\frac{d^2}{dt^2}f(t) \leq 2w_2' \int_\Omega \left(\left|\nabla\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} + K^2\|\mathrm{Rm}\|_g \left|\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} \right) dx$$
$$+ 4w_3'K^2 \int_\Omega \left|\nabla\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} dx.$$

In summary we have ellipticity with $\lambda = \frac{2w_2}{1+C_1(\Omega)}$ and $\Lambda = \max\left(2w_2' + 4w_3'K^2, 2w_2'K^2\|\mathrm{Rm}\|_g\right)$. This proves 1. For the proof of the result under Assumption 2 we estimate

$$\frac{d^2}{dt^2}f(t) \geq 2\int_\Omega \left(w_2 \left|\nabla\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} - w_2'K^2\|\mathrm{Rm}\|_g \left|\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} \right) dx$$
$$\geq 2(w_2 - w_2'K^2\|\mathrm{Rm}\|_g C_1(\Omega)) \int_\Omega \left|\nabla\dot{\Gamma}(x,t)\right|^2_{g(\Gamma(x,t))} dx$$
$$\geq 2\frac{w_2 - w_2'K^2\|\mathrm{Rm}\|_g C_1(\Omega)}{1+C_1(\Omega)} |\dot{\Gamma}(t)|^2_{H^1}.$$

We get ellipticity with $\lambda = 2\frac{w_2-w_2'K^2\|\mathrm{Rm}\|_g C_1(\Omega)}{1+C_1(\Omega)}$ and $\Lambda = \max\left(2w_2' + 4w_3'K^2, 2w_2'K^2\|\mathrm{Rm}\|_g\right)$. $\qquad\square$

7.2.2 Harmonic Maps with Potential

We can also generalize the harmonic energy by adding a source term with potential $G : \Omega \times M \to \mathbb{R}$. We arrive at

$$\mathfrak{J}^{\mathrm{harm},G}(v) = \mathfrak{J}^{\mathrm{harm}}(v) + \int_\Omega G(x,v(x))\,dx,$$

see [17].

Again, for simplicity, in the following we will suppress the dependence of G on its first variable x and assume that $G : M \to \mathbb{R}$. The second derivative of $\mathfrak{J}^{\mathrm{harm},G}$ along a geodesic homotopy Γ splits in the same terms as above for the harmonic energy, plus the Hessian of G. Note that for a point $q \in M$, the Hessian $\mathrm{Hess}\,G : T_qM \times T_qM \to R$ is

$$\mathrm{Hess}(G)(v,w) := \left\langle \frac{D}{dt}\mathrm{grad}\,G(\gamma(s))\Big|_{s=0}, w \right\rangle_{g(q)}, \quad \forall v,w \in T_qM.$$

where $\gamma : \{-\epsilon,\epsilon\} \to M$ is a differentiable path such that $\gamma(0) = q$ and $\dot{\gamma}(0) = v$.

Lemma 7.6. *With Γ a geodesic homotopy and $f(t) := \mathfrak{J}^{\mathrm{harm},G}(\Gamma(t))$ with $G : M \to \mathbb{R}$ we have*

$$\frac{d^2}{dt^2}f(t) = 2\int_\Omega \left\langle \nabla\dot{\Gamma}, \nabla\dot{\Gamma} \right\rangle^2 dx - 2\int_\Omega \left\langle \mathrm{Rm}(\nabla\Gamma,\dot{\Gamma})\dot{\Gamma}, \nabla\Gamma \right\rangle dx + \int_\Omega \mathrm{Hess}(G)(\dot{\Gamma},\dot{\Gamma})\,dx.$$

The potential G influences the ellipticity of $\mathfrak{J}^{\mathrm{harm},G}$ in the following way.

Corollary 7.1. *The energy $\mathfrak{J}^{\mathrm{harm},G} : H_K \cap H^\Phi \to \mathbb{R}$ is elliptic along geodesic homotopies if either*

1. *M has nonpositive sectional curvature, and $\mathrm{Hess}\,G$ is positive semidefinite, or*

2. *we have*
$$1 - K^2\|\mathrm{Rm}\|_g C_1(\Omega) + \inf_{v\in TM} \frac{\mathrm{Hess}(G)(v,v)}{|v|^2_g} > 0.$$

Hence $\mathfrak{J}^{\mathrm{harm},G}$ can be elliptic even if $\mathfrak{J}^{\mathrm{harm}}$ by itself is not, provided that $\mathrm{Hess}\,G$ is sufficiently positive definite.

For various results related to the smoothness of harmonic maps with potential we refer to [17].

7.2.3 Tikhonov Regularization

As a special case of the above, we can choose the source term to be the distance from a given function $w : \Omega \to M$

$$\mathfrak{J}^w(v) := \mathfrak{J}^{\mathrm{harm}}(v) + \int_\Omega \mathrm{dist}(v(x), w(x))^2 \mu(dx).$$

It is useful for applications to allow the source term to be integrated with respect to a general positive measure μ, which may be discrete. Minimizing such an energy $\mathfrak{J}^w$ can be useful in smoothing, denoising or motion planning [54]. For $d = 1$ (by defining μ to be a discrete measure and using boundedness of point evaluations in H^1 for $d = 1$) the framework includes a point-fitting energy

$$\mathfrak{J}^{\mathrm{harm}}(v) + \sum_{i=1}^N \mathrm{dist}(v(x_i), p_i)^2$$

for interpolation points $x_i \in \Omega$ and point values $p_i \in M$.

In the case of nonpositive curvature, ellipticity can be established easily.

Lemma 7.7. *Assume that M has nonpositive sectional curvature. Then the energy $\mathfrak{J}^w : H_K \cap H^\Phi \to \mathbb{R}$ is elliptic along geodesic homotopies.*

Proof. This is a simple consequence of the ellipticity of $\mathfrak{J}^{\mathrm{harm}}$, together with the fact that for a geodesic $\gamma(t)$ in M, we have that

$$\frac{d^2}{dt^2} \mathrm{dist}(\gamma(t), p)^2 \geq 0$$

for all points $p \in M$ if M has nonpositive curvature [51]. Therefore, the functional $\mathfrak{J}^w$ is coercive for any choice of w. $\square$

Observe that the ellipticity of the functional $\mathfrak{J}^w$ holds even without Dirichlet boundary conditions. If M has positive curvature, additional restrictions regarding the diameter of the image $u(\Omega)$ apply.

8 Conclusion

We have provided a generalization of the classical finite element theory to manifold-valued problems by establishing appropriate manifold-valued generalizations of the classical Céa-Lemma and interpolation error bounds for geodesic finite element (GFE) spaces. Along the way we have introduced a number of new technical tools for dealing with the analysis of manifold-valued functions which we expect to be useful beyond this paper. One example application of our theory are high-order numerical schemes for the computation of harmonic maps into manifolds.

Many issues remain for future work. Aside from natural issues such as for instance the investigation of the effects of variational crimes in the spirit of Strang [50], we mention a more thorough study of ellipticity properties for several geometric energies of interest, among them a finer study of the harmonic energy with positively curved target spaces, or the Cosserat energies studied in [38, 42]. Additionally, convexity properties of the energies on the approximation spaces are of interest, because they influence the convergence speed of numerical solvers. Further it will be interesting to study weak convergence properties of GFE discretizations for non-elliptic energies and/or nonsmooth solutions, generalizing results of [6]. Finally we mention further extensions of linear finite element-based methods such as e.g., nonconforming variants of geodesic finite elements and temporal discretizations for nonstationary problems.

Acknowledgments

References

[1] P. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] F. Alouges. A new algorithm for computing liquid crystal stable configurations: the harmonic mapping case. *SIAM Journal on Numerical Analysis*, 34(5):1708–1726, 1997.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser, 2008.

[4] D. Amsallem and C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM Journal on Scientific Computing*, 33(5):2169–2198, 2011.

[5] M. Ara. Geometry of *F*-harmonic maps. *Kodai Mathematical Journal*, 22(2):243–263, 1999.

[6] S. Bartels. Stability and convergence of finite-element approximation schemes for harmonic maps. *SIAM Journal on Numerical Analysis*, 43(1):220–238, 2005.

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

[8] J. Céa. Approximation variationnelle des problemes aux limites. *Ann. Inst. Fourier (Grenoble)*, 14 (2):345–444, 1964.

[9] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM, 2002.

[10] P. Clément. Approximation by finite element function using local regularization. *RAIRO Anal. Numer.*, 2:77–84, 1975.

[11] 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.

[12] M. Do Carmo. *Riemannian geometry*. Birkhäuser Boston, 1992.

[13] N. Dyn, P. Grohs, and J. Wallner. Approximation order of interpolatory nonlinear subdivision schemes. *Journal of computational and applied mathematics*, 233(7):1697–1703, 2010.

[14] J. Eells and J. Sampson. Harmonic mappings of Riemannian manifolds. *American Journal of Mathematics*, 86(1):109–160, 1964.

[15] J. Eels and L. Lemaire. A report on harmonic maps. *Bull. London Math. Soc.*, 10(1):1–68, 1978.

[16] H. I. Eliasson. Geometry of manifolds of maps. *J. Differential Geometry*, 1:169–194, 1967.

[17] A. Fardoun and A. Ratto. Harmonic maps with potential. *Calculus of Variations and Partial Differential Equations*, 5(2):183–197, 1997.

[18] P. T. Fletcher and S. Joshi. Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, 87(2):250–262, 2007.

[19] L. Grafakos. *Classical and Modern Fourier Analysis*. Pearson Education, Inc., 2004.

[20] P. Grohs. Smoothness equivalence properties of univariate subdivision schemes and their projection analogues. *Numerische Mathematik*, 113(2):163–180, 2009.

[21] P. Grohs. A general proximity analysis of nonlinear subdivision schemes. *SIAM Journal on Mathematical Analysis*, 42(2):729–750, 2010.

[22] P. Grohs. Approximation order from stability for nonlinear subdivision schemes. *Journal of Approximation Theory*, 162(5):1085–1094, 2010.

[23] P. Grohs. Geometric multiscale decompositions of dynamic low-rank matrices. 2012. SAM report 2012/03, Seminar for Applied Mathematics, ETH Zürich.

[24] P. Grohs. Quasi-interpolation in Riemannian manifolds. *IMA Journal of Numerical Analysis*, 2012. doi: http://dx.doi.org/10.1093/imanum/drs026.

[25] P. Grohs and J. Wallner. Interpolatory wavelets for manifold-valued data. *Applied and Computational Harmonic Analysis*, 27(3):325–333, 2009.

[26] P. Hajlasz. Sobolev mappings between manifolds and metric spaces. *Sobolev Spaces In Mathematics I*, pages 185–222, 2009.

[27] F. Hang and F. Lin. Topology of Sobolev mappings, II. *Acta mathematica*, 191(1):55–107, 2003.

[28] R. Hardt, D. Kinderlehrer, and F. Lin. Existence and partial regularity of static liquid crystal configurations. *Communications in mathematical physics*, 105(4):547–570, 1986.

[29] F. Hélein. Régularité des applications faiblement harmoniques entre une surface et une variété riemannienne. *CR Acad. Sci. Paris Sér. I Math*, 312, 1991.

[30] M. Hofer and H. Pottmann. Energy-minimizing splines in manifolds. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 284–293. ACM, 2004.

[31] V. Ivancevic. Symplectic rotational geometry in human biomechanics. *SIAM review*, 46(3):455–474, 2004.

[32] W. Jäger and H. Kaul. Uniqueness and stability of harmonic maps and their Jacobi fields. *Manuscripta Mathematica*, 28(1):269–291, 1979.

[33] J. Jost. *Riemannian geometry and geometric analysis*. Springer, 2011.

[34] H. Karcher. Mollifier smoothing and Riemannian center of mass. *Communications on Pure and Applied Mathematics*, 30:509–541, 1977.

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

[36] I. Münch. *Ein geometrisch und materiell nichtlineares Cosserat-Model — Theorie, Numerik und Anwendungsmöglichkeiten*. PhD thesis, Universität Karlsruhe, 2007.

[37] J. Nash. The imbedding problem for Riemannian manifolds. *Annals of Mathematics*, 63(1):20–63, 1956.

[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. Mod. Meth. Appl. Sci.*, 17(3):363–392, 2007.

[39] R. S. Palais. *Foundations of global non-linear analysis*, volume 196. Benjamin New York, 1968.

[40] I. U. Rahman, I. Drori, V. C. Stodden, D. L. Donoho, and P. Schröder. Multiscale representations for manifold-valued data. *Multiscale Modeling & Simulation*, 4(4):1201–1232, 2005.

[41] T. Rivière. Everywhere discontinuous harmonic maps into spheres. *Acta Mathematica*, 175(2): 197–226, 1995.

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

[43] O. Sander. Geodesic finite elements on simplicial grids. *Int. J. Num. Meth. Eng.*, 92(12):999–1025, 2012.

[44] O. Sander. Geodesic finite elements of higher order. 2013. IGPM Preprint 356, RWTH Aachen.

[45] R. Schoen and K. Uhlenbeck. A regularity theory for harmonic maps. *J. Differential Geom.*, 17(2): 307–335, 1982.

[46] R. Schoen and S.-T. Yau. *Lectures on Harmonic Maps.* International Press of Boston, 1997.

[47] J. Simo and D. Fox. On a stress resultant geometrically exact shell model. Part I: Formulation and optimal parametrization. *Comput. Methods Appl. Mech. Engrg.*, 72:267–304, 1989.

[48] 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.

[49] 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(1):21–70, 1990.

[50] G. Strang. Variational crimes in the finite element method. In *The mathematical foundations of the finite element method with applications to partial differential equations*, pages 689–710. Univ. Maryland, Academic Press, New York, 1972.

[51] K. Sturm. Probability measures on metric spaces of nonpositive curvature. *Contemporary mathematics*, 338:357–390, 2003.

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

[53] B. Tang, G. Sapiro, and V. Caselles. Color image enhancement via chromaticity diffusion. *Image Processing, IEEE Transactions on*, 10(5):701–707, 2001.

[54] G. Wahba. *Spline models for observational data*, volume 59. SIAM, 1990.

[55] A. Weinmann. Interpolatory multiscale representation for functions between manifolds. *SIAM Journal on Mathematical Analysis*, 44(1):162–191, 2012.

[56] P. Wriggers and F. Gruttmann. Thin shells with finite rotations formulated in Biot stresses: Theory and finite element formulation. *Int. J. Num. Meth. Engrg.*, 36:2049–2071, 1993.

[57] G. Xie and T. P.-Y. Yu. Approximation order equivalence properties of manifold-valued data subdivision schemes. *IMA Journal of Numerical Analysis*, 32(2):687–700, 2012.

A Taylor Series Remainder Estimates

In this appendix we prove two technical results about certain remainder terms in Taylor series expansions. They are used in the proof of the interpolation error bound in Section 5.1.

Lemma A.1. *For a function $H(x, y) : T_{ref} \times T_{ref} \to \mathbb{R}$ and a multi-index $\vec{k}$ with $|\vec{k}| > d/2$ we have the inequality*

$$\left\| \int_0^1 t^{|\vec{k}|-1} x^{\vec{k}} H(x, tx) \, dt \right\|_{L^2(T_{ref})} \lesssim \left\| \sup_{z \in T_{ref}} |H(z, x)| \right\|_{L^2(T_{ref})},$$

where the integration in the L^2-norms above occur in the variable x. The implicit constant only depends on the diameter of T_{ref}.

Proof. We only treat the case $d = 2$, the general case being similar. Also, for simplicity we assume that T_{ref} is contained in the unit ball. Using polar coordinates $(x_1, x_2) = r s_\varphi := r(\cos(\varphi), \sin(\varphi))$ and the substitution $\tau = rt$, we can write

$$\int_0^1 t^{|\vec{k}|-1} x^{\vec{k}} H(x, tx) \, dt = \int_0^1 t^{|\vec{k}|-1} r^{|\vec{k}|} s_\varphi^{\vec{k}} H(x, trs_\varphi) \, dt = \int_0^r \tau^{|\vec{k}|-1} s_\varphi^{\vec{k}} H(x, \tau s_\varphi) \, d\tau.$$

We need to estimate the L^2-norm of this expression. Since T_{ref} is contained in the unit ball we get

$$\left\| \int_0^1 t^{|\vec{k}|-1} x^{\vec{k}} H(x,tx)\, dt \right\|_{L^2(T_{\text{ref}})}^2 \leq \int_0^1 \int_0^{2\pi} \left(\int_0^r \tau^{|\vec{k}|-1} s_\varphi^{\vec{k}} H(x,\tau s_\varphi)\, d\tau \right)^2 r\, dr\, d\varphi.$$

Using the Cauchy–Schwarz inequality we can bound this expression by

$$\int_0^1 \int_0^{2\pi} \int_0^r H(x,\tau s_\varphi)^2 \tau\, d\tau\, d\varphi \int_0^r \tau^{2|\vec{k}|-4} \tau\, d\tau\, r\, dr$$

which can in turn be bounded by

$$\int_0^1 \int_0^{2\pi} \int_0^1 \sup_{x\in T_{\text{ref}}} H(x,\tau s_\varphi)^2 \tau\, d\tau\, d\varphi \int_0^1 \tau^{2|\vec{k}|-3}\, d\tau\, r\, dr \leq \left\| \sup_{x\in T_{\text{ref}}} |H(x,\cdot)| \right\|_{L^2(T_{\text{ref}})}^2 \int_0^1 \int_0^1 \tau^{2|\vec{k}|-3}\, d\tau\, r\, dr.$$

Since the double integral on the right is no greater than 1, we get the desired expression. □

Lemma A.2. *For a function U defined on T_{ref} and a multi-index $\vec{e}$ with $|\vec{e}| = 1$ we have*

$$\sum_{|\vec{k}|=k} \frac{|\vec{k}|}{\vec{k}!} \partial^{\vec{e}} \int_0^1 t^{|\vec{k}|-1} x^{\vec{k}} \partial^{\vec{k}} U(tx)\, dt = \sum_{|\vec{l}|=k-1} \frac{(-1)^{k-1}}{\vec{l}!} x^{\vec{l}} \partial^{\vec{l}+\vec{e}} U(x). \tag{46}$$

Proof. The term on the left-hand side of (46) is the derivative of the residual term

$$R(x) := U(0) - \sum_{|\vec{l}|<k} (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}} U(x)$$

in the Taylor expansion of U around x and evaluated at zero. Using this interpretation, one can check the statement by direct computation. Indeed, applying the operator $\partial^{\vec{e}}$ to R and using the product rule we get

$$\partial^{\vec{e}} R(x) = -\partial^{\vec{e}} \sum_{|\vec{l}|<k} (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}} U(x)$$

$$= \sum_{|\vec{l}|<k} (-1)^{|\vec{l}|} \frac{x^{\vec{l}-\vec{e}}}{(\vec{l}-\vec{e})!} \partial^{\vec{l}} U(x) + (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}+\vec{e}} U(x)$$

$$= -\sum_{|\vec{l}|<k-1} (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}+\vec{e}} U(x) + \sum_{|\vec{l}|<k} (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}+\vec{e}} U(x)$$

$$= \sum_{|\vec{l}|=k-1} (-1)^{|\vec{l}|} \frac{x^{\vec{l}}}{\vec{l}!} \partial^{\vec{l}+\vec{e}} U(x).$$

In the second line we have used the convention $x^{\vec{l}-\vec{e}} \equiv 0$ whenever $\vec{l} - \vec{e} \in \mathbb{Z}^d$ has a negative entry. □