Approximating the Derivative of Manifold-valued Functions

Ralf Hielscher* Laura Lippert†

We consider the approximation of manifold-valued functions by embedding the manifold into a higher dimensional space, applying a vector-valued approximation operator and projecting the resulting vector back to the manifold. It is well known that the approximation error for manifold-valued functions is close to the approximation error for vector-valued functions. This is not true anymore if we consider the derivatives of such functions. In our paper we give pre-asymptotic error bounds for the approximation of the derivative of manifold-valued function. In particular, we provide explicit constants that depend on the reach of the embedded manifold.

Keywords: nonlinear approximation, manifold-valued functions, embedded manifolds

1. Introduction

Approximating functions $f\colon \Omega \to \mathcal{M}$ with values in some D-dimensional Riemannian manifold $\mathcal{M}$ has attracted lots of interest during the last years. The central challenge is that with $\mathcal{M}$ not being linear, the function spaces over Ω with values in $\mathcal{M}$ are not linear as well and hence, all the well established linear approximation methods do not have a straight forward generalization to manifold-valued functions.

One successful approach to manifold-valued approximation is to consider the problem locally. Either one maps the function values locally to some linear approximation space or one uses local averaging based on the geodesic distance. These approaches allow to generalize subdivision schemes [32, 33, 9, 10, 29], moving least squares [12], quasi-interpolation [11] or splines [30] to the manifold-valued setting.

A different approach is to embed the manifold $\mathcal{M}$ into some higher dimensional linear space $\mathbb{R}^d$ by a map $\mathcal{E}\colon \mathcal{M} \to \mathbb{R}^d$. Note, that according to Nash's embedding theorem [21]

*Faculty of Mathematics and Informatics, TU Bergakademie Freiberg, Germany.
E-mail: ralf.hielscher@math.tu-chemnitz.de
†Faculty of Mathematics, Chemnitz University of Technology, Germany.
E-mail: laura.lippert@math.tu-chemnitz.de

such a mapping always exists and can be guaranteed to be locally isometric provided the dimension d is at least $D(3D+11)/2$, where D denotes the dimension of $\mathcal{M}$. Embedding based approximation methods can be summarized as follows

i) Transfer $f \in C(\Omega, \mathcal{M})$ via the embedding $\mathcal{E}$ into the linear function space $C(\Omega, \mathbb{R}^d)$. Since often a manifold is described by vectors in $\mathbb{R}^d$, we will use the notation f for the function in both function spaces.

ii) Use a linear approximation operator $I_{\mathbb{R}^d} \colon C(\Omega, \mathbb{R}^d) \to C(\Omega, \mathbb{R}^d)$ to find an approximant $I_{\mathbb{R}^d} f$ in the embedding.

iii) Project the resulting $\mathbb{R}^d$-valued function back to the manifold $I_{\mathcal{M}} f = P_{\mathcal{M}} \circ (I_{\mathbb{R}^d} f)$ using some projection operator $P_{\mathcal{M}} \colon \mathbb{R}^d \to \mathcal{M}$.

Because of its generality and simplicity this approach has already been widely investigated [11, 7] and applied [20, 27]. In particular it has been shown in [7] that the approximation order of the embedding based approximation operator $I_{\mathcal{M}} f$ is the same as the approximation order of its linear counterpart $I_{\mathbb{R}^d}$. It is important to note, that the projection operator $P_{\mathcal{M}}$ is in general only defined in some neighborhood $U \supset \mathcal{M}$ of the manifold. Hence, the pre-asymptotic behavior of the approximation operator $I_{\mathcal{M}}$ strongly depends on the size of this neighborhood which is directly related to the so-called reach of the embedded manifold.

The aim of this paper is to analyze the pre-asymptotic behavior of the approximation operator $I_{\mathcal{M}}$ with respect to the reach of the embedded manifold. While for the error $I_{\mathcal{M}} f - f$ the reach only controls the required linear approximation error $I_{\mathbb{R}^d} f - f$ that allows for a meaningful approximation $I_{\mathcal{M}} f$, the situation is completely different for the error of the derivatives $\mathrm{d}(I_{\mathcal{M}} f) - \mathrm{d}f$. In this case the chain rule has to be used for the derivative $\mathrm{d}(I_{\mathcal{M}} f)$ of a concatenation of approximation in $\mathbb{R}^d$ and projection on the manifold $\mathcal{M}$. The derivative of the projection on $\mathcal{M}$ leads to pre-asymptotic constants, which depend on the reach of the manifold.

Our paper is organized as follows. In section 2 we will first show some general differential geometric properties of submanifolds of $\mathbb{R}^d$. Most importantly, we identify in Lemma 2.1 the projection operator $P_{\mathcal{M}}$ with an orthogonal projection in the normal bundle over the manifold $\mathcal{M}$. This is only possible within some tubular neighborhood of the manifold $\mathcal{M}$ that is controlled by the reach of $\mathcal{M}$. The relationship between the reach of the manifold $\mathcal{M}$ and its curvature or second fundamental form is addressed in section 2.2. In Theorem 2.7 we make use of these relationships to describe the differential of the projection operator $P_{\mathcal{M}}$ in terms of the second fundamental form. In section 2.4 we end up with the main results of this chapter, that is we show in Theorem 2.10 that the derivative $dP_{\mathcal{M}}(\boldsymbol{x})$ of the projection operator at some point $\boldsymbol{x} \in \mathbb{R}^d$ satisfies a Lipschitz-condition with respect to $\boldsymbol{x}$. As our Lipschitz bound is with respect to the Euclidean distance in the embedding it is more sharp then the bound reported in [3] that relies on the geodesic distance.

Section 3 is dedicated to manifold-valued approximation. We show that the approach of using a linear approximation operator on an embedded manifold $\mathcal{M}$ in $\mathbb{R}^d$ and then

projecting back on the manifold inherits the approximation order of the linear approximation. Our main result is stated in Theorem 3.2 and gives a pre-asymptotic bound for the approximation error of the first derivative that relies exclusively on the reach of the embedded manifold. This result is illustrated in Theorem 3.4 for a specific approximation operator, the Fourier partial sum operator.

In the final section 4 we consider two real world examples for approximating manifold valued data. The first example deals with functions from the two-sphere into the two-dimensional projective space that describe the dependency between the propagation direction and the polarization directions of seismic waves. The second example is from crystallographic texture analysis where the local alignment of the atom lattice is described by a map with values in the quotient $\mathrm{SO}(3)/\mathcal{S}$ of the rotation group $\mathrm{SO}(3)$ modulo some finite symmetry group $\mathcal{S}$. The derivative of this map has important connections microscopic and macroscopic material properties.

2. Submanifolds

In this section we will consider smooth compact Riemannian submanifolds $\mathcal{M}$ of $\mathbb{R}^d$. We will show some differential geometric properties of submanifolds as well as some estimations for the projection $P_{\mathcal{M}}$ and the differential of this projection. We will use these results for estimating some approximation errors in section 3.

2.1. The Projection Operator

Throughout our work we denote by $\mathcal{M} \subset \mathbb{R}^d$ a smooth compact Riemannian submanifold of $\mathbb{R}^d$. For every point $\boldsymbol{m} \in \mathcal{M}$ we denote the tangent space $T_{\boldsymbol{m}}\mathcal{M}$ as well as the normal space $N_{\boldsymbol{m}}\mathcal{M}$. Furthermore, we denote by $P_{\mathcal{M}} \colon \mathbb{R}^d \to \mathcal{M}$ the projection operator onto $\mathcal{M}$ defined as the solution of the minimization problem

$$P_{\mathcal{M}}(\boldsymbol{x}) = \operatorname*{argmin}_{\boldsymbol{m} \in \mathcal{M}} \|\boldsymbol{x} - \boldsymbol{m}\|_2 . \tag{2.1}$$

In general, this minimization problem does not posses a unique solution for every $\boldsymbol{x} \in \mathbb{R}^d$, since there is an ambiguity to which branch of the manifold the point should be attributed. However, if we restrict the domain of the definition of $P_{\mathcal{M}}$ to some open neighborhood $U \subset \mathbb{R}^d$ of $\mathcal{M}$ uniqueness can be granted.

In order to find such a neighborhood U we define on the normal bundle

$$N\mathcal{M} = \{(\boldsymbol{m}, \boldsymbol{v}) \in \mathbb{R}^d \times \mathbb{R}^d : \boldsymbol{m} \in \mathcal{M}, \boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}\}$$

of $\mathcal{M}$ the smooth map

$$E \colon N\mathcal{M} \to \mathbb{R}^d, \quad E(\boldsymbol{m}, \boldsymbol{v}) = \boldsymbol{m} + \boldsymbol{v},$$

that maps every normal space $N_{\boldsymbol{m}}\mathcal{M}$ to an affine linear subspace through $\boldsymbol{m} \in \mathbb{R}^d$. Since we assumed $\mathcal{M}$ to be compact and smooth, there exist a maximum constant $\tau > 0$ such that the mapping E restricted to the open subset

$$V = \{(\boldsymbol{m}, \boldsymbol{v}) \in N\mathcal{M} : \|\boldsymbol{v}\|_2 < \tau\}$$

of the normal bundle is injective, cf. [17, 6.24]. Setting $U = E(V)$ defines the so-called *tubular neighborhood* of $\mathcal{M}$ and the restriction $E\colon V \to U$ becomes a diffeomorphism. The constant τ is commonly called *reach* and its inverse $1/\tau$ is the *condition number* of the manifold. The reach τ is affected by two factors: the curvature of the manifold and the width of the narrowest bottleneck-like structure of $\mathcal{M}$, which quantifies how far $\mathcal{M}$ is from being self-intersecting. An estimate on the relationship between the reach and the curvature of the manifold $\mathcal{M}$ will be given in Lemma 2.4.

Using the mapping E we may now give an explicit definition of the projection operator $P_{\mathcal{M}}$.

Lemma 2.1. *Let $\boldsymbol{u} \in U$ and let $\pi\colon N\mathcal{M} \to \mathcal{M}$, $(\boldsymbol{m}, \boldsymbol{n}) \mapsto \boldsymbol{m}$ be the canonical projection operator. Then*

$$P_{\mathcal{M}}(\boldsymbol{u}) = \pi \circ E^{-1}(\boldsymbol{u})$$

is the unique solution of the minimization problem (2.1).

Proof. Let $\boldsymbol{u} \in U$ and $P_{\mathcal{M}}(\boldsymbol{u}) = \boldsymbol{m} \in \mathcal{M}$. We show that $\boldsymbol{u} - \boldsymbol{m} \in N_{\boldsymbol{m}}\mathcal{M}$. We assume the opposite and decompose $\boldsymbol{u} - \boldsymbol{m}$ in one part in $N_{\boldsymbol{m}}\, M$ and a part $\boldsymbol{t}$ in $T_{\boldsymbol{m}}M$. Then there is a curve $\gamma(s)$ in $\mathcal{M}$ with $\gamma(0) = \boldsymbol{m}$ and $\dot{\gamma}(0) = \boldsymbol{t}$. If we go along this curve, we obtain for sufficient small $\epsilon > 0$ that $\boldsymbol{u} - \gamma(\epsilon) < \boldsymbol{u} - \gamma(0) = \boldsymbol{u} - \boldsymbol{m}$. That is a contradiction to the definition of $P_{\mathcal{M}}$. Since the projection $P_{\mathcal{M}}$ should be unique, we have to show that $\pi \circ E^{-1}$ is also unique. For that reason we assume that for $\boldsymbol{u} \in U$ there holds $\pi \circ E^{-1} = \boldsymbol{m} \in \mathcal{M}$ and $\pi \circ E^{-1} = \boldsymbol{m}' \in \mathcal{M}$. This would imply $\boldsymbol{u} = \boldsymbol{m} + \boldsymbol{v} = \boldsymbol{m}' + \boldsymbol{v}'$ with $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ and $\boldsymbol{v} \in N_{\boldsymbol{m}'}\mathcal{M}$. That is a contradiction to the uniqueness of E^{-1} in the tubular neighborhood U. ∎

Let us illustrate this by a simple example.

Example 2.2. Let the manifold $\mathcal{M}$ be the $(d-1)$-dimensional unit sphere, embedded in $\mathbb{R}^d$. These manifolds can be described by

$$\mathbb{S}^{d-1} = \left\{ \boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_2 = 1 \right\}.$$

The projection $P_{\mathbb{S}^{d-1}}$ easily reads

$$P_{\mathbb{S}^{d-1}} : \mathbb{R}^d \backslash \{\boldsymbol{0}\} \to \mathbb{S}^{d-1}, \qquad P_{\mathbb{S}^{d-1}}(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}.$$

This map is well-defined and smooth.

2.2. Curvature and Reach of Submanifolds

For any point $\boldsymbol{m} \in \mathcal{M} \subset \mathbb{R}^d$ on the manifold we can decompose $\mathbb{R}^d$ as the direct sum $\mathbb{R}^d = T_{\boldsymbol{m}}\mathcal{M} \oplus N_{\boldsymbol{m}}\mathcal{M}$ of the tangential space $T_{\boldsymbol{m}}\mathcal{M}$ and the normal space $N_{\boldsymbol{m}}\mathcal{M}$. Let us denote by $P_T\colon \mathbb{R}^d \to T_{\boldsymbol{m}}\mathcal{M}$ and $P_N\colon \mathbb{R}^d \to N_{\boldsymbol{m}}\mathcal{M}$ the corresponding orthogonal projections. Then the canonical connection ∇ on $\mathbb{R}^d$ defines a connection $\nabla^{\mathcal{M}}$ on $\mathcal{M}$ by

$$\nabla^{\mathcal{M}}_{\boldsymbol{X}} = P_T \nabla_{\boldsymbol{X}}(P_T \boldsymbol{Y}) + P_N \nabla_{\boldsymbol{X}}(P_N \boldsymbol{Y}), \tag{2.2}$$

where $\boldsymbol{X}\colon \mathcal{M} \to T\mathcal{M}$ is a tangential and $\boldsymbol{Y}\colon \mathcal{M} \to \mathbb{R}^d$ a general vector field on $\mathcal{M}$.

If $\boldsymbol{Y}$ is a tangential vector field as well, the first summand $P_T \nabla_{\boldsymbol{X}}(P_T \boldsymbol{Y}) = P_T \nabla_{\boldsymbol{X}} \boldsymbol{Y}$ in (2.2) is just the Levi-Cevita-connection on $\mathcal{M}$, whereas its orthogonal complement

$$\mathrm{II}(\boldsymbol{X}, \boldsymbol{Y}) = P_N(\nabla_{\boldsymbol{X}} \boldsymbol{Y})$$

is the *second fundamental form* on $\mathcal{M}$.

We call a vector field $\boldsymbol{Y}\colon \mathcal{M} \to \mathbb{R}^d$ parallel along a curve γ if $\nabla^{\mathcal{M}}_{\dot\gamma} \boldsymbol{Y} = 0$. For a geodesic γ with $\gamma(0) = \boldsymbol{m}$, $\dot\gamma(0) = \boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}$ and an arbitrary vector $\boldsymbol{y} \in T_{\boldsymbol{m}}\mathcal{M} \oplus N_{\boldsymbol{m}}\mathcal{M} = \mathbb{R}^d$ we shall use the abbreviation

$$\nabla_{\boldsymbol{t}} \boldsymbol{y} = \nabla_{\boldsymbol{t}} \boldsymbol{Y}(0)$$

where $\boldsymbol{Y}$ is the parallel transport of the vector $\boldsymbol{y}$ along the curve γ.

For a fixed point $\boldsymbol{m} \in \mathcal{M}$ and a normal direction $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$ we define the operator $\boldsymbol{B}_{\boldsymbol{n}}\colon T_{\boldsymbol{m}}\mathcal{M} \to T_{\boldsymbol{m}}\mathcal{M}$ on the tangent space by

$$\langle \boldsymbol{B}_{\boldsymbol{n}} \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{n}, \nabla_{\boldsymbol{x}} \boldsymbol{y} \rangle, \quad \boldsymbol{x}, \boldsymbol{y} \in T_{\boldsymbol{m}}\mathcal{M}. \tag{2.3}$$

We may also express $B_{\boldsymbol{n}}\boldsymbol{x}$ as the tangential part of the covariant derivative of $\boldsymbol{n}$ in direction $\boldsymbol{x}$.

Lemma 2.3. *Let $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$ be a normal and $\boldsymbol{x} \in T_{\boldsymbol{m}}\mathcal{M}$ a tangential vector. Then*

$$\boldsymbol{B}_{\boldsymbol{n}} \boldsymbol{x} = -P_T \nabla_{\boldsymbol{x}} \boldsymbol{n}.$$

Proof. Let γ be a geodesics in $\mathcal{M}$ with $\gamma(0) = \boldsymbol{m}$ and $\dot\gamma(0) = \boldsymbol{x}$ and let $\boldsymbol{N}$ be the parallel transport of $\boldsymbol{n}$ along γ. Let furthermore, $\boldsymbol{Y}$ be an arbitrary tangent vector field parallel along γ. Then we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}s} \langle \boldsymbol{N}(s), \boldsymbol{Y}(s) \rangle|_{s=0} = \langle \nabla_{\boldsymbol{x}} \boldsymbol{N}(0), \boldsymbol{Y}(0) \rangle + \langle \boldsymbol{N}(0), \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{Y}(0) \rangle.$$

This yields

$$\langle \boldsymbol{B}_{\boldsymbol{n}} \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{n}, \nabla_{\boldsymbol{x}} \boldsymbol{y} \rangle = -\langle \nabla_{\boldsymbol{x}} \boldsymbol{n}, \boldsymbol{y} \rangle.$$

Since the vector field $\boldsymbol{Y}$ was arbitrary, this yields the assertion. ∎

The operator $\boldsymbol{B}_{\boldsymbol{n}}$ describes the extrinsic curvature of the manifold in the point $\boldsymbol{m}$ and the normal direction $\boldsymbol{n}$. Its norm is bounded by the condition number $\frac{1}{\tau}$ of $\mathcal{M}$. More precisely the following result is shown in [22, Proposition 6.1].

Lemma 2.4. *Let τ be the reach of $\mathcal{M}$, $\boldsymbol{m} \in \mathcal{M}$ be an arbitrary point on the manifold and $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$ be a normal vector. Then the operator $B_{\boldsymbol{n}}$ defined in* (2.3) *is symmetric and bounded by $\frac{1}{\tau}$, i.e., we have for tangential vectors $\boldsymbol{x}, \boldsymbol{y} \in T_{\boldsymbol{m}}\mathcal{M}$ the inequality*

$$\langle \boldsymbol{B}_{\boldsymbol{n}} \boldsymbol{x}, \boldsymbol{y} \rangle \leq \frac{1}{\tau} \|\boldsymbol{n}\| \, \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|. \tag{2.4}$$

The next lemma bounds the covariant derivative of parallel vector fields by the condition number $\frac{1}{\tau}$ of the manifold.

Lemma 2.5. *Let $\boldsymbol{Y}$ be a parallel vector field along a geodesic γ in $\mathcal{M}$. Then its covariant derivative in $\mathbb{R}^d$ is bounded by*

$$\|\nabla_{\dot\gamma}\boldsymbol{Y}\|_2 \le \tfrac{1}{\tau}\,\|\boldsymbol{Y}\|_2\,\|\dot\gamma\|_2\,.$$

Proof. Let $\boldsymbol{Y} = \boldsymbol{T} + \boldsymbol{N}$ be the decomposition of $\boldsymbol{Y}$ into a tangent vector field $\boldsymbol{T}$ and a normal vector field $\boldsymbol{N}$. Since $\boldsymbol{Y}$ is parallel along γ we have

$$0 = \nabla^{\mathcal{M}}_{\dot\gamma}\boldsymbol{Y} = P_T\nabla_{\dot\gamma}\boldsymbol{T} + P_N\nabla_{\dot\gamma}\boldsymbol{N}$$

and, hence,

$$\nabla_{\dot\gamma}\boldsymbol{Y} = P_N\nabla_{\dot\gamma}\boldsymbol{T} + P_T\nabla_{\dot\gamma}\boldsymbol{N}.$$

Let $\boldsymbol{n} = P_N\nabla_{\dot\gamma(s)}\boldsymbol{T}(s)$. Then we obtain by Lemma 2.4

$$\|\boldsymbol{n}\|^2 = \big\langle \boldsymbol{n}, \nabla_{\dot\gamma(s)}\boldsymbol{T}(s)\big\rangle = \langle \boldsymbol{B}_{\boldsymbol{n}}\dot\gamma(s), \boldsymbol{T}(s)\rangle \le \tfrac{1}{\tau}\,\|\dot\gamma\|\,\|\boldsymbol{n}\|\,\|\boldsymbol{T}(s)\|\,.$$

For the tangential part we have by Lemma 2.3

$$\|P_T\nabla_{\dot\gamma(s)}\boldsymbol{N}(s)\|_2 = \|\boldsymbol{B}_{\boldsymbol{N}(s)}\dot\gamma(s)\|_2 \le \tfrac{1}{\tau}\,\|\dot\gamma\|\,\|\boldsymbol{N}(s)\|\,.$$

The assertion follows now from Parsevals inequality. ∎

2.3. The Differential of the Projection Operator.

The differential $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m})\colon \mathbb{R}^d \to T_{\boldsymbol{m}}\mathcal{M}$ of the projection $P_{\mathcal{M}}\colon \mathbb{R}^d \to \mathcal{M}$ is especially easy to compute at points $\boldsymbol{m} \in \mathcal{M}$ on the manifold. In this case it is simply the linear projection $P_{T_{\boldsymbol{m}}\mathcal{M}}\colon \mathbb{R}^d \to T_{\boldsymbol{m}}\mathcal{M}$ onto the tangential space attached to $\boldsymbol{m}$, i.e.

$$\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) = P_{T_{\boldsymbol{m}}\mathcal{M}}. \tag{2.5}$$

We can verify this by observing that for normal vectors $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$ we have

$$\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m})\,\boldsymbol{n} = \lim_{h\to 0}\frac{P_{\mathcal{M}}(\boldsymbol{m}+h\boldsymbol{n}) - P_{\mathcal{M}}(\boldsymbol{m})}{h} = \lim_{h\to 0}\frac{\boldsymbol{m}-\boldsymbol{m}}{h} = \boldsymbol{0},$$

while for tangent vectors $\boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}$ we obtain

$$\begin{aligned}\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m})\,\boldsymbol{t} &= \lim_{h\to 0}\frac{P_{\mathcal{M}}(\boldsymbol{m}+h\boldsymbol{t}) - P_{\mathcal{M}}(\boldsymbol{m})}{h} = \lim_{h\to 0}\frac{\boldsymbol{m}+\exp(h\boldsymbol{t}) - \boldsymbol{m}}{h}\\ &= \lim_{h\to 0}\frac{\exp(h\boldsymbol{t})}{h} = \boldsymbol{t},\end{aligned}$$

where exp denotes the exponential map to the manifold.

The differential $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v})$, $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ at a point not in the manifold is a little bit more tricky. We start by observing that the tangential $T_{(\boldsymbol{m},\boldsymbol{v})}N\mathcal{M} \subset \mathbb{R}^{2d}$ of the normal bundle at a point $(\boldsymbol{m},\boldsymbol{v}) \in N\mathcal{M}$ is

$$\begin{aligned}T_{(\boldsymbol{m},\boldsymbol{v})}N\mathcal{M} &= \{(\boldsymbol{0},\boldsymbol{n}) \mid \boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}\} \oplus \{(\boldsymbol{t}, \nabla_{\boldsymbol{t}}\boldsymbol{v}) \mid \boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}\}\\ &= \{(\boldsymbol{t},\boldsymbol{u}) \mid \boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}, P_T\boldsymbol{u} = \nabla_{\boldsymbol{t}}\boldsymbol{v}\}.\end{aligned}$$

The following lemma describes the differential $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v})$.

Lemma 2.6. *Let $\boldsymbol{m} \in \mathcal{M}$ be an arbitrary point on the manifold $\mathcal{M}$ and $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ be a normal vector with $\|\boldsymbol{v}\|_2 < \tau$, i.e. $\boldsymbol{m} + \boldsymbol{v}$ is in the tubular neighborhood of $\mathcal{M}$. Then the derivative $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})$ satisfies for every tangent direction $\boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}$,*

$$\left(\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})\right)(\boldsymbol{t} + \nabla_{\boldsymbol{t}}\boldsymbol{v}) = \boldsymbol{t}.$$

while it vanishes for any normal direction $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$, i.e.

$$\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})\,\boldsymbol{n} = \boldsymbol{0}.$$

Proof. According to Lemma 2.1 we have $P_{\mathcal{M}} = \pi \circ E^{-1}$, where $\pi\colon N\mathcal{M} \to \mathcal{M}$, $(\boldsymbol{m}, \boldsymbol{v}) \mapsto \boldsymbol{m}$ is the projection operator. Its differential at the point $(\boldsymbol{m}, \boldsymbol{v}) \in N\mathcal{M}$ is the projection

$$\mathrm{d}\pi(\boldsymbol{m}, \boldsymbol{v})\colon T_{(\boldsymbol{m},\boldsymbol{v})}N\mathcal{M} \to T_{\boldsymbol{m}}\mathcal{M}, \quad (\boldsymbol{t}, \boldsymbol{u}) \mapsto \boldsymbol{t}.$$

The differential of the mapping $E\colon N\mathcal{M} \to \mathbb{R}^d$ in a point $(\boldsymbol{m}, \boldsymbol{v}) \in N\mathcal{M}$ is given by

$$\mathrm{d}E(\boldsymbol{m}, \boldsymbol{v})\colon T_{(\boldsymbol{m},\boldsymbol{v})}N\mathcal{M} \to \mathbb{R}^d, \quad (\boldsymbol{t}, \boldsymbol{u}) \mapsto \boldsymbol{t} + \boldsymbol{u}.$$

Since $\boldsymbol{m} + \boldsymbol{v}$ is within the tubular neighborhood of $\mathcal{M}$, E is invertible in some neighborhood of $\boldsymbol{m} + \boldsymbol{v}$. Then $\mathrm{d}E(\boldsymbol{m}, \boldsymbol{v})$ is invertible as well and we have for any normal vector $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$

$$\mathrm{d}E^{-1}(\boldsymbol{m} + \boldsymbol{v})\,\boldsymbol{n} = (\boldsymbol{0}, \boldsymbol{n})$$

and for any tangent vector $\boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}$

$$\mathrm{d}E^{-1}(\boldsymbol{m} + \boldsymbol{v})\,(\boldsymbol{t} + \nabla_{\boldsymbol{t}}\boldsymbol{v}) = (\boldsymbol{t}, \nabla_{\boldsymbol{t}}\boldsymbol{v}).$$

Together with the chain rule this implies the assertion. ■

The image of $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})$ is contained in the tangential space $T_{\boldsymbol{m}}\mathcal{M}$, especially $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})$ is the projection $P_{T_{\boldsymbol{m}}\mathcal{M}}$ up to a factor matrix. We will write this linear operator $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})$ in another way, to see the difference to the linear operator $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m})$.

Theorem 2.7. *Let $\boldsymbol{m} \in \mathcal{M}$ be a point on the manifold, let $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ be a normal vector with $\|\boldsymbol{v}\| < \tau$ and let $\boldsymbol{B}_{\boldsymbol{v}}\colon T_{\boldsymbol{m}}\mathcal{M} \to T_{\boldsymbol{m}}\mathcal{M}$ be the symmetric operator defined in (2.3), extended to $\boldsymbol{B}_{\boldsymbol{v}}\colon \mathbb{R}^d \to \mathbb{R}^d$ by $\boldsymbol{B}_{\boldsymbol{v}}\boldsymbol{n} = \boldsymbol{0}$ for all normal vectors $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$. Then the derivative of the projection operator $P_{\mathcal{M}}$ satisfies*

$$\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v}) = P_{T_{\boldsymbol{m}}\mathcal{M}}(\boldsymbol{I} - \boldsymbol{B}_{\boldsymbol{v}})^{-1} = \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) - \boldsymbol{B}_{\boldsymbol{v}}\left(\boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{v}}\right)^{-1},$$

where $\boldsymbol{I}\colon \mathbb{R}^d \to \mathbb{R}^d$ is the identity.

Proof. Using Lemma 2.6 we obtain for all tangential vectors $\boldsymbol{t} \in T_{\boldsymbol{m}}\mathcal{M}$,

$$\begin{aligned} P_{T_{\boldsymbol{m}}\mathcal{M}}\boldsymbol{t} = \boldsymbol{t} &= \left(\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})\right)(\boldsymbol{t} + \nabla_{\boldsymbol{t}}\boldsymbol{v}) \\ &= \left(\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})\right)(\boldsymbol{t} - \boldsymbol{B}_{\boldsymbol{v}}\boldsymbol{t}) = \left(\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v})\right)(\boldsymbol{I} - \boldsymbol{B}_{\boldsymbol{v}})\,\boldsymbol{t}. \end{aligned}$$

and for all normal vectors $\boldsymbol{n} \in N_{\boldsymbol{m}}\mathcal{M}$,

$$\mathbf{0} = P_{T_{\boldsymbol{m}}\mathcal{M}}\boldsymbol{n} = \left(\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v})\right)\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right)\boldsymbol{n}.$$

Consequently, we have

$$P_{T_{\boldsymbol{m}}\mathcal{M}} = \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v})\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right).$$

By our assumption and (2.4) we have $\|\boldsymbol{B}_{\boldsymbol{v}}\| \leq \frac{1}{\tau}\|\boldsymbol{v}\| < 1$ and hence, the operator $\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}$ is invertible. This yields the first part of the assertion. For the second part we use (2.5) and compute

$$\begin{aligned}
\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v}) &= P_{T_{\boldsymbol{m}}\mathcal{M}}\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right)^{-1} \\
&= P_{T_{\boldsymbol{m}}\mathcal{M}}\left(\boldsymbol{I}+\boldsymbol{B}_{\boldsymbol{v}}\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right)^{-1}\right) \\
&= \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) + P_{T_{\boldsymbol{m}}\mathcal{M}}\boldsymbol{B}_{\boldsymbol{v}}\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right)^{-1} \\
&= \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) + \boldsymbol{B}_{\boldsymbol{v}}\left(\boldsymbol{I}-\boldsymbol{B}_{\boldsymbol{v}}\right)^{-1},
\end{aligned}$$

where the last equality follows from the fact that the image of $\boldsymbol{B}_{\boldsymbol{n}}$ is in the tangent space $T_{\boldsymbol{m}}\mathcal{M}$, so the projection on $T_{\boldsymbol{m}}\mathcal{M}$ is unnecessary. ■

We consider again the manifold from example 2.2.

Example 2.8. For $\mathcal{M} = \mathbb{S}^{d-1} \subset \mathbb{R}^d$ any normal vector $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathbb{S}^{d-1}$ has the representation $\boldsymbol{v} = v\,\boldsymbol{m}$. Let $\{\boldsymbol{t}_i\}_{i=1}^{d-1} \subset T_{\boldsymbol{m}}\mathbb{S}^{d-1}$ be an orthonormal basis of $T_m\mathbb{S}^{d-1}$. Then $\nabla_{\boldsymbol{t}_i}\boldsymbol{m} = \boldsymbol{t}_j$ and hence

$$\boldsymbol{B}_{\boldsymbol{v}} = -v\sum_{i=1}^{d-1}\boldsymbol{t}_i\boldsymbol{t}_i^{\top}.$$

By Theorem 2.7 and the orthonormality of $\{\boldsymbol{m}\} \cap \{\boldsymbol{t}_i\}_{i=1}^{d-1}$ of we obtain for $v > -1$ and $\boldsymbol{x} = \boldsymbol{m} + v\boldsymbol{m}$,

$$\mathrm{d}P_{\mathbb{S}^{d-1}}(\boldsymbol{x}) = \frac{1}{1+v}\sum_{i=1}^{d-1}\boldsymbol{t}_i\boldsymbol{t}_i^{\top} = \frac{1}{\|\boldsymbol{x}\|_2}\left(\boldsymbol{I}_{d\times d} - \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\right)^{\top}\right).$$

2.4. Deviation of the Projection Operator

In this section we are interested in the change of the derivative $\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m})$ of the projection operator for small deviations of $\boldsymbol{m}$. We shall show that for two points $\boldsymbol{m}$ and $\boldsymbol{z}$ on $\mathcal{M}$ and $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ with $\|\boldsymbol{v}\|_2 < \tau$ we can bound the difference $\|\,\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}+\boldsymbol{v}) - \mathrm{d}P_{\mathcal{M}}(\boldsymbol{z})\|_2$ by a multiple of the Euclidean distance $\|\boldsymbol{m}+\boldsymbol{v}-\boldsymbol{z}\|_2$.

As usual we start with the case that both points are on the manifold. According to [3, Lemma 6] the difference of the differentials is then bounded by

$$\|\,\mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) - \mathrm{d}P_{\mathcal{M}}(\boldsymbol{z})\|_2 \leq \frac{1}{\tau}\,d(\boldsymbol{m},\boldsymbol{z}),$$

where $d(\boldsymbol{m},\boldsymbol{z})$ denotes the geodesic distance between the points $\boldsymbol{m},\boldsymbol{z} \in \mathcal{M}$. If the Euclidean distance between the two points is bounded by $\|\boldsymbol{m}-\boldsymbol{z}\|_2 \leq 2\tau$ we have by

[3, Lemma 3] and the fact that $\arcsin(x) \leq \frac{\pi}{2}x$ for $0 \leq x \leq 1$, the following estimate between geodesic distance and Euclidean distance in the embedding

$$d(\boldsymbol{m}, \boldsymbol{z}) \leq \frac{\pi}{2} \|\boldsymbol{m} - \boldsymbol{z}\|_2 , \tag{2.6}$$

which leads to the local estimate

$$\| \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) - \mathrm{d}P_{\mathcal{M}}(\boldsymbol{z})\|_2 \leq \frac{\pi}{2\tau} \|\boldsymbol{m} - \boldsymbol{z}\|_2 .$$

In the following Theorem we prove a sharper and global bound for this difference.

Theorem 2.9. *For all $\boldsymbol{m}, \boldsymbol{z} \in \mathcal{M}$ the difference between the projection operators $P_{T_{\boldsymbol{m}}\mathcal{M}}$ and $P_{T_{\boldsymbol{z}}\mathcal{M}}$ onto the respective tangential spaces is bounded by*

$$\|P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}}\|_2 \leq \frac{1}{\tau} \|\boldsymbol{m} - \boldsymbol{z}\|_2 .$$

Proof. First of all we note that for $\|\boldsymbol{m} - \boldsymbol{z}\|_2 \geq 2\,\tau$ the assertion is immediately satisfied since $\|P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}}\|_2 \leq 2$ independently of $\boldsymbol{m}, \boldsymbol{z} \in \mathcal{M}$. We may therefore assume $\|\boldsymbol{m} - \boldsymbol{z}\|_2 < 2\,\tau$ for the rest of the proof.

In order to estimate the difference between the two projection operators we consider a geodesic γ with $\gamma(0) = \boldsymbol{m}$, $\gamma(t) = \boldsymbol{z}$ and $\|\dot{\gamma}\|_2 = 1$. Furthermore, we consider an orthonormal basis $\{\boldsymbol{t}_i\}_{i=1}^{D}$ in $T_{\boldsymbol{m}}\mathcal{M}$ and an orthonormal basis $\{\boldsymbol{n}_j\}_{j=1}^{d-D}$ in $N_{\boldsymbol{m}}\mathcal{M}$. The parallel transport of these basis vectors along γ defines a rotation $\boldsymbol{R} \in \mathrm{SO}(d)$ that maps the tangent space $T_{\boldsymbol{m}}\mathcal{M}$ onto the tangent space $T_{\boldsymbol{z}}\mathcal{M}$. Using the rotation $\boldsymbol{R}$ we may rewrite the difference between the projection operators as

$$P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}} = P_{T_{\boldsymbol{m}}\mathcal{M}} - \boldsymbol{R}P_{T_{\boldsymbol{m}}\mathcal{M}}\boldsymbol{R}^T.$$

By Lemma A.1 in the appendix we obtain

$$\|P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}}\|_2 = \|P_{T_{\boldsymbol{m}}\mathcal{M}}\boldsymbol{R} - \boldsymbol{R}P_{T_{\boldsymbol{z}}\mathcal{M}}\|_2 \leq \|\boldsymbol{I} - \boldsymbol{R}\|_2 \tag{2.7}$$

and hence, it suffices to bound for any normalized $\boldsymbol{x} \in \mathbb{R}^d$

$$\|(\boldsymbol{I} - \boldsymbol{R})\boldsymbol{x}\|_2^2 = 2 - 2\,\langle \boldsymbol{x}, \boldsymbol{R}\boldsymbol{x}\rangle . \tag{2.8}$$

By definition $\boldsymbol{R}\boldsymbol{x}$ is the result of the parallel transport of $\boldsymbol{x}$ along the curve γ in $\gamma(t) = \boldsymbol{z}$. Let us denote by $\boldsymbol{X}(s)$ the parallel transport of $\boldsymbol{x}$ along γ for all times $s \in [0, t]$. Viewing $s \mapsto \boldsymbol{X}(s)$ as a curve on $\mathbb{S}^{d-1}$ with velocity bounded according to Lemma 2.5 by $\|\dot{\boldsymbol{X}}(s)\| = \|\nabla_{\dot{\gamma}(s)}\boldsymbol{X}(s)\|_2 \leq \frac{1}{\tau}$, we conclude that

$$\angle(\boldsymbol{X}(\eta), \boldsymbol{X}(\xi)) \leq \frac{1}{\tau}|\eta - \xi|, \quad \eta, \xi \in [0, t]. \tag{2.9}$$

Since γ is a geodesic we can set in (2.9), $\boldsymbol{X} = \dot{\gamma}$. As $|\eta - \xi| \leq t$ and t is the geodesic distance between $\boldsymbol{z}$ and $\boldsymbol{m}$ we can use (2.6) and our assumption $\|\boldsymbol{m} - \boldsymbol{z}\|_2 < 2\tau$ to bound the right hand side of (2.9) by

$$\angle(\dot{\gamma}(\xi), \dot{\gamma}(\eta)) \leq \frac{1}{\tau}\,|\eta - \xi| \leq \frac{t}{\tau} \leq \frac{\pi}{2\tau} \|\boldsymbol{z} - \boldsymbol{m}\|_2 \leq \pi.$$

Making use of the monotonicity of the cosine this implies

$$\cos \angle(\dot\gamma(\xi), \dot\gamma(\eta)) > \cos \tfrac{\xi-\eta}{\tau}, \quad \xi, \eta \in [0, t]. \tag{2.10}$$

Considering again the general vector field $\boldsymbol{X}$ we use (2.9) and (2.10) to bound (2.8) by

$$\begin{aligned}
2 - 2 \langle \boldsymbol{X}(0), \boldsymbol{X}(t) \rangle &= 2 - 2 \cos(\angle(\boldsymbol{X}(0), \boldsymbol{X}(t))) \\
&\leq 2 - 2 \cos \tfrac{t}{\tau} = \frac{1}{\tau^2} \int_0^t \int_0^t \cos \tfrac{\xi-\eta}{\tau} \, \mathrm{d}\eta \, \mathrm{d}\xi \\
&\leq \frac{1}{\tau^2} \int_0^t \int_0^t \cos \angle(\dot\gamma(\xi), \dot\gamma(\eta)) \, \mathrm{d}\eta \, \mathrm{d}\xi \\
&= \frac{1}{\tau^2} \int_0^t \int_0^t \langle \dot\gamma(\xi), \dot\gamma(\eta) \rangle \, \mathrm{d}\eta \, \mathrm{d}\xi = \frac{1}{\tau^2} \left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2^2 .
\end{aligned}$$

In combination with (2.7) and (2.8) this proves

$$\left\| P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}} \right\|_2 \leq \frac{1}{\tau} \left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2 . \qquad \blacksquare$$

Using the example of the unit circle it can be easily verified that our new bound is sharp.

So far we bounded the variation of the projection operator for points on the manifold. For the general case that only one point is on the manifold we have the following result.

Theorem 2.10. *Let $\boldsymbol{m}, \boldsymbol{z} \in \mathcal{M}$ and $\boldsymbol{v} \in N_{\boldsymbol{m}}\mathcal{M}$ with $\left\| \boldsymbol{v} \right\|_2 < \tau$. Then*

$$\begin{aligned}
\left\| \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m} + \boldsymbol{v}) - \mathrm{d}P_{\mathcal{M}}(\boldsymbol{z}) \right\|_2 &\leq \frac{1}{\tau} \left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2 + \frac{1}{\tau - \left\| \boldsymbol{v} \right\|_2} \left\| \boldsymbol{v} \right\|_2 \\
&\leq \left(\frac{2}{\tau} + \frac{1}{\tau - \left\| \boldsymbol{v} \right\|_2} \right) \left\| \boldsymbol{m} + \boldsymbol{v} - \boldsymbol{z} \right\|_2 .
\end{aligned}$$

Proof. Using Theorem 2.7 and Theorem 2.9 we find

$$\begin{aligned}
\left\| \mathrm{d}P_M(\boldsymbol{m} + \boldsymbol{v}) - \mathrm{d}P_M(\boldsymbol{z}) \right\|_2 &\leq \left\| \mathrm{d}P_M(\boldsymbol{m} + \boldsymbol{v}) - \mathrm{d}P_M(\boldsymbol{m}) \right\|_2 + \left\| P_{T_{\boldsymbol{m}}\mathcal{M}} - P_{T_{\boldsymbol{z}}\mathcal{M}} \right\|_2 \\
&\leq \left\| \boldsymbol{B}_{\boldsymbol{v}} \left(\boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{v}} \right)^{-1} \right\|_2 + \tfrac{1}{\tau} \left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2 .
\end{aligned}$$

From Lemma 2.4 we know that $\left\| \boldsymbol{B}_{\boldsymbol{v}} \right\| \leq \frac{\left\| \boldsymbol{v} \right\|_2}{\tau} < 1$. This allows us to bound the second term by

$$\left\| \boldsymbol{B}_{\boldsymbol{v}} \left(\boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{v}} \right)^{-1} \right\|_2 \leq \frac{\left\| \boldsymbol{B}_{\boldsymbol{v}} \right\|}{1 - \left\| \boldsymbol{B}_{\boldsymbol{v}} \right\|} \leq \frac{\left\| \boldsymbol{v} \right\|_2}{\tau - \left\| \boldsymbol{v} \right\|_2},$$

which implies the first inequality of the theorem

$$\left\| \mathrm{d}P_M(\boldsymbol{m} + \boldsymbol{v}) - \mathrm{d}P_M(\boldsymbol{z}) \right\|_2 \leq \frac{1}{\tau} \left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2 + \frac{1}{\tau - \left\| \boldsymbol{v} \right\|_2} \left\| \boldsymbol{v} \right\|_2 . \tag{2.11}$$

Since $\left\| \boldsymbol{v} \right\|_2 < \tau$ the point $\boldsymbol{m} + \boldsymbol{v}$ is within the tubular neighborhood of $\mathcal{M}$ and, hence $\left\| \boldsymbol{v} \right\|_2 \leq \left\| \boldsymbol{m} + \boldsymbol{v} - \boldsymbol{z} \right\|_2$. Together with the triangle inequality this gives us $\left\| \boldsymbol{m} - \boldsymbol{z} \right\|_2 \leq 2 \left\| \boldsymbol{m} + \boldsymbol{v} - \boldsymbol{z} \right\|_2$. Including these two inequalities into (2.11) we obtain the assertion. $\blacksquare$

We observe that the constants in Theorem 2.10 become large if either the reach of the manifold becomes small or the point $\boldsymbol{m} + \boldsymbol{v}$ is close to the boundary of the tubular neighborhood of $\mathcal{M}$.

3. Manifold-valued Approximation

In this section we generalize arbitrary approximation operators for vector valued functions to approximation operators for manifold-valued functions. To this end we consider for an arbitrary domain Ω a generic approximation operator $I_{\mathbb{R}^d}\colon C(\Omega,\mathbb{R}^d)\to C(\Omega,\mathbb{R}^d)$. For an embedded manifold $\mathcal{M}\subset\mathbb{R}^d$ with reach τ and projection operator

$$P_{\mathcal{M}}\colon U\to\mathcal{M},\quad U=\{\boldsymbol{y}\in\mathbb{R}^d\mid \min_{\boldsymbol{m}\in\mathcal{M}}\|\boldsymbol{y}-\boldsymbol{m}\|<\tau\},$$

we define the approximation operator $I_{\mathcal{M}}\colon C(\Omega,\mathcal{M})\to C(\Omega,\mathcal{M})$ for manifold-valued functions as

$$I_{\mathcal{M}}f=P_{\mathcal{M}}\circ I_{\mathbb{R}^d}f.$$

It is important to note that $I_{\mathcal{M}}$ is not defined for all functions $f\in C(\Omega,\mathcal{M})$, but only for those for which $I_{\mathbb{R}^d}f(x)$ is within the reach of the manifold $\mathcal{M}$, i.e., $\|f(\boldsymbol{x})-I_{\mathbb{R}^d}f(x)\|_2\leq\tau$ for all $\boldsymbol{x}\in\Omega$.

It is straight forward to see that operator $I_{\mathcal{M}}$ has the same order of approximation as $I_{\mathbb{R}^d}$, c.f. [7].

Theorem 3.1. *Let $f\in C(\Omega,\mathcal{M})$ be a continuous $\mathcal{M}$-valued function such that for all $\boldsymbol{x}\in\Omega$, $I_{\mathbb{R}^d}f(\boldsymbol{x})$ is contained in the reach of $\mathcal{M}$. We then have for all $\boldsymbol{x}\in\Omega$*

$$\|I_{\mathcal{M}}f(\boldsymbol{x})-f(\boldsymbol{x})\|_2\leq 2\,\|I_{\mathbb{R}^d}f(\boldsymbol{x})-f(\boldsymbol{x})\|_2\,. \tag{3.1}$$

Proof. Since f has function values on $\mathcal{M}$, it follows from the definition of $P_{\mathcal{M}}$ in equation (2.1) for all $\boldsymbol{x}\in\Omega$ that

$$\|I_{\mathcal{M}}f(\boldsymbol{x})-I_{\mathbb{R}^d}f(\boldsymbol{x})\|_2\leq\|f(\boldsymbol{x})-I_{\mathbb{R}^d}f(\boldsymbol{x})\|_2\,.$$

Because of the triangle inequality and the definition of $I_{\mathcal{M}}$ we have

$$\|I_{\mathcal{M}}f(\boldsymbol{x})-f(\boldsymbol{x})\|_2\leq\|I_{\mathcal{M}}f(\boldsymbol{x})-I_{\mathbb{R}^d}f(\boldsymbol{x})\|_2+\|I_{\mathbb{R}^d}f(\boldsymbol{x})-f(\boldsymbol{x})\|_2\,,\quad \boldsymbol{x}\in\Omega.$$ ■

As we will see later, considering the error of the differential, things become more complicated.

3.1. Approximation Order of the Differential

In this section we are interested in the approximation error $\|\,\mathrm{d}I_{\mathcal{M}}f-\mathrm{d}f\|_2$ between the differential of the manifold-valued approximation $\mathrm{d}I_{\mathcal{M}}f$ and the original differential $\mathrm{d}f$. To this end we assume from now on that both, $f\colon\Omega\to\mathbb{R}^d$ and the vector-valued approximation $\tilde{f}=I_{\mathbb{R}^d}f$, are differentiable.

While the error bound for $I_{\mathcal{M}}f$ is independent of the geometry of the manifold $\mathcal{M}$, we will see that this is not true for the differential $\mathrm{d}I_{\mathcal{M}}f$ of the manifold-valued approximation. Moreover, it is not sufficient to ensure that $\tilde{f}$ is contained in the reach of $\mathcal{M}$, but instead, it must be bounded away from the reach by some positive constant.

Theorem 3.2. *Let τ be the reach of the manifold $\mathcal{M}$, $\varepsilon < \tau$ and $f \in C^1(\Omega, \mathcal{M})$, such that $\tilde{f}(\boldsymbol{x}) = I_{\mathbb{R}^d} f(\boldsymbol{x})$ satisfies for all $\boldsymbol{x} \in \Omega$,*

$$\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2 \leq \varepsilon$$

and, consequently, is contained in the ε-tubular neighborhood of $\mathcal{M}$. Then we have for all $\boldsymbol{x} \in \Omega$ the following upper bound on the approximation error of the differential $\mathrm{d}I_{\mathcal{M}} f$,

$$\|\,\mathrm{d}I_{\mathcal{M}} f(\boldsymbol{x}) - \,\mathrm{d}\, f(\boldsymbol{x})\|_2 \leq \|\,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2 + C\,\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2 \tag{3.2}$$

where C is given by

$$C = \left(\frac{2}{\tau} + \frac{1}{\tau - \varepsilon}\right)\left(\|\,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2 + \|\,\mathrm{d}f(\boldsymbol{x})\|_2\right).$$

Proof. By the chain rule we obtain for all $\boldsymbol{x} \in \Omega$,

$$\mathrm{d}I_{\mathcal{M}} f(\boldsymbol{x}) = \mathrm{d}P_{\mathcal{M}}(\tilde{f}(\boldsymbol{x})) \circ \,\mathrm{d}\tilde{f}(\boldsymbol{x})$$

and from $P_{\mathcal{M}} f = f$,

$$\mathrm{d}f(\boldsymbol{x}) = \,\mathrm{d}(P_{\mathcal{M}} f)(\boldsymbol{x}) = \mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x})) \circ \,\mathrm{d}f(\boldsymbol{x}).$$

Using the expansion

$$\begin{aligned}\mathrm{d}(I_{\mathcal{M}} f)(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x}) =& \mathrm{d}P_{\mathcal{M}}(\tilde{f}(\boldsymbol{x})) \circ \,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x})) \circ \,\mathrm{d}f(\boldsymbol{x}) \\ =& (\mathrm{d}P_{\mathcal{M}}(\tilde{f}(\boldsymbol{x})) - \mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x}))) \circ \,\mathrm{d}\tilde{f}(\boldsymbol{x}) \\ &+ \mathrm{d}P_{\mathcal{M}} f(\boldsymbol{x}) \circ (\,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x}))\end{aligned}$$

we conclude that

$$\begin{aligned}\|\,\mathrm{d}I_{\mathcal{M}} f(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2 \leq& \|\mathrm{d}P_{\mathcal{M}}(\tilde{f}(\boldsymbol{x})) - \mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x}))\|_2\,\|\,\mathrm{d}\tilde{f}(\boldsymbol{x})\|_2 \\ &+ \|\mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x}))\|_2\,\|\,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2.\end{aligned}$$

Since $f(\boldsymbol{x}) \in \mathcal{M}$ and $\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2 < \varepsilon$ we have by Theorem 2.10

$$\|\mathrm{d}P_{\mathcal{M}}(\tilde{f}(\boldsymbol{x})) - \mathrm{d}P_{\mathcal{M}}(f(\boldsymbol{x}))\| \leq \left(\frac{2}{\tau} + \frac{1}{\tau - \varepsilon}\right)\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2.$$

Together with the fact that $\|\mathrm{d}P_{\mathcal{M}} f(\boldsymbol{x})\|_2 = 1$ we obtain

$$\begin{aligned}\|\,\mathrm{d}(I_{\mathcal{M}} f)(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2 \leq& \|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2 \left(\frac{2}{\tau} + \frac{1}{\tau - \varepsilon}\right)\|\,\mathrm{d}\tilde{f}(\boldsymbol{x})\|_2 \\ &+ \|\,\mathrm{d}\tilde{f}(\boldsymbol{x}) - \,\mathrm{d}f(\boldsymbol{x})\|_2.\end{aligned}$$

This implies the assertion by triangle inequality. ∎

Asymptotically, as $\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2 \to 0$ we have $C \to \frac{3}{\tau} \|\,\mathrm{d}f(x)\|_2$. Since for most approximation methods the decay of the differential $\|\,\mathrm{d}f(\boldsymbol{x}) - \mathrm{d}\tilde{f}(\boldsymbol{x})\|_2$ is one order slower than the decay of $\|f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})\|_2$, we conclude that the right hand bound in (3.2) is dominated by the first summand and, hence, the approximation error of the differential $\mathrm{d}I_{\mathcal{M}}f$ of the manifold-valued approximant coincides asymptotically with the approximation error of the vector-valued approximant, as it was already reported in [7]. However, the pre-asymptotic behavior depends strongly on the reach of the embedding of the manifold $\mathcal{M}$.

3.2. Fourier Interpolation

In this section we want to illustrate Theorem 3.2 using Fourier-Interpolation as the approximation operator $I_{\mathbb{R}^d}$. More precisely, we define for a function $f \in C(\mathbb{T}, \mathbb{R}^d)$ on the torus $\mathbb{T}$ the Fourier partial sum

$$I_{\mathbb{R}^d} f(t) = S_n f(t) = \sum_{k=-n}^{n} c_k(f) e^{2\pi i k t}$$

with the vector-valued Fourier coefficients

$$c_k(f) = \int_0^1 f(x) e^{-2\pi i k x}\,\mathrm{d}x.$$

The Fourier-Interpolation satisfies the following well known approximation inequalities, cf. [26].

Theorem 3.3. *Let $r \in \mathbb{N}$ with $r \geq 2$ and $f \in C^r(\mathbb{T}, \mathbb{R}^d)$. Then*

$$\|f(x) - S_n f(x)\|_2 \leq \frac{\sqrt{2d}}{(2\pi)^r} \frac{\sqrt{n}}{n^r} \|f^{(r)}\|_{L^2(\mathbb{T}),2}\,,$$
$$\|\,\mathrm{d}f(x) - \mathrm{d}(S_n f)(x)\|_2 \leq \frac{\sqrt{2d}}{(2\pi)^{r-1}} \frac{\sqrt{n}}{n^{r-1}} \|f^{(r)}\|_{L^2(\mathbb{T}),2}\,,$$

with the norm $\|f\|^2_{L^2(\mathbb{T}),2} = \int_{\mathbb{T}} \|f(x)\|_2^2\,\mathrm{d}x$.

Proof. The first bound can be found in [4, Theorem 4.3]. The second bound follows from $\mathrm{d}(S_n f) = S_n(\mathrm{d}f)$ and the fact that the regularity of $\mathrm{d}f$ is one less than the regularity of f. The factor $\sqrt{d}$ comes from the fact that the function f maps in the d-dimensional space. ∎

In [26, Theorem 1.39] a similar bound for the $L^\infty(\mathbb{T})$-norm can be found. Using the Fourier partial sum operator as the approximation operator $I_{\mathbb{R}^d}$ Theorem 3.2 becomes the following.

Theorem 3.4. *Let $\mathcal{M} \subset \mathbb{R}^d$ be a submanifold with reach $\tau > 0$ and $f \in C^r(\mathbb{T}, \mathcal{M})$ an $r \geq 2$ times differentiable function with values in $\mathcal{M}$. Let, furthermore, $\varepsilon < \tau$ be an auxiliary constant and the bandwidth n of the Fourier partial sum $S_n f$ at least such that $\|f(x) - S_n f(x)\|_2 \leq \varepsilon$ for all $x \in \mathbb{T}$, i.e.,*

$$n^{r-\frac{1}{2}} \geq \frac{C_1 \, \|f^{(r)}\|_{L^2(\mathbb{T}),2}}{\varepsilon} \text{ with } C_1 = \frac{\sqrt{2d}}{(2\pi)^r}. \tag{3.3}$$

Then the projection $P_{\mathcal{M}} \circ S_n f$ of the Fourier partial sum satisfies for all $x \in \mathbb{T}$,

$$\|P_{\mathcal{M}} \circ S_n f(x) - f(x)\|_2 \leq 2\, C_1 \, n^{\frac{1}{2}-r} \, \|f^{(r)}\|_{L^2(\mathbb{T}),2}\,, \tag{3.4}$$

whereas for its differential $\mathrm{d}(P_{\mathcal{M}} \circ S_n f)$ we obtain

$$\begin{aligned} \|\,\mathrm{d}(P_{\mathcal{M}} \circ S_n f)(x) - \,\mathrm{d}f(x)\|_2 \leq 2\pi\, C_1 \, n^{\frac{3}{2}-r} \, \|f^{(r)}\|_{L^2(\mathbb{T}),2} \\ + C_2\, n^{\frac{1}{2}-r} \, \|f^{(r)}\|^2_{L^2(\mathbb{T}),2}\,, \end{aligned} \tag{3.5}$$

with the constant

$$C_2 = C_1 \, (\tfrac{2}{\tau} + \tfrac{1}{\tau-\varepsilon}) \, (1 + 2\pi\, C_1^2 \, n^{\frac{3}{2}-r}).$$

Proof. Together with Theorem 3.3 condition (3.3) ensures that

$$\|f(x) - S_n f(x)\|_2 \leq \varepsilon < \tau, \quad x \in \mathbb{T}$$

and, hence, $S_n f(x)$ has distance less than the reach to $\mathcal{M}$ for all $x \in \mathbb{T}$. This allows us to apply Theorem 3.1 in conjunction with Theorem 3.3 to conclude (3.4).

For the approximation error of the derivative we have by Theorem 3.2

$$\|\,\mathrm{d}(P_{\mathcal{M}} \circ S_n f)(x) - \,\mathrm{d}f(x)\|_2 \leq \|\,\mathrm{d}(S_n f)(x) - \,\mathrm{d}f(x)\|_2 + C\, \|S_n f(x) - f(x)\|_2$$

with

$$C = \left(\tfrac{\tau}{2} + \tfrac{1}{\tau-\varepsilon}\right) \, (\|\,\mathrm{d}(S_n f)(x) - \,\mathrm{d}f(x)\|_2 + \|\,\mathrm{d}f(x)\|_2)\,.$$

Together with Theorem 3.3 this yields

$$\|\,\mathrm{d}(P_{\mathcal{M}} \circ S_n f)(x) - \,\mathrm{d}f(x)\|_2 \leq 2\pi\, C_1 \, n^{\frac{3}{2}-r} \, \|f^{(r)}\|_{L^2(\mathbb{T}),2} + C\, C_1 \, n^{\frac{1}{2}-r} \, \|f^{(r)}\|_{L^2(\mathbb{T}),2}$$

and

$$\begin{aligned} C &\leq (\tfrac{\tau}{2} + \tfrac{1}{\tau-\varepsilon}) \, (2\pi\, C_1 \, n^{\frac{3}{2}-r} \, \|f^{(r)}\|_{L^2(\mathbb{T}),2} + \|\,\mathrm{d}f(x)\|_2) \\ &\leq (\tfrac{\tau}{2} + \tfrac{1}{\tau-\varepsilon}) \, (2\pi\, C_1 \, n^{\frac{3}{2}-r} + 1) \, \|f^{(r)}\|_{L^2(\mathbb{T}),2}\,, \end{aligned}$$

where we have used the fact that $\|\,\mathrm{d}f(x)\|_2 \leq \|f^{(r)}\|_{L^2(\mathbb{T}),2}$ for periodic functions. Setting

$$C_2 = C_1(\tfrac{\tau}{2} + \tfrac{1}{\tau-\varepsilon})(1 + 2\pi\, C_1 \, n^{\frac{3}{2}-r})$$

yields the assertion.

■

Since for $n \to \infty$ we have $C_2 \to C_1(\frac{2}{\tau} + \frac{1}{\tau-\varepsilon})$. Theorem 3.4 states that the approximation order of the differential of the manifold-valued Fourier partial sum operator coincides with the approximation order of the vector-valued operator. In the preasymptotic setting, however, also the second summand with the faster rate $n^{\frac{1}{2}-r}$ is relevant. The constant of this second summand becomes large if the point-wise approximation error is close to the reach τ of the embedding.

4. Examples

In this section we apply our findings to two real world examples of manifold-valued approximation. Both examples are related to the analysis of crystalline materials. In the first example we consider functions that relate propagation directions of waves to polarization directions and in the second example we consider functions that relate points within crystalline specimen to the local orientation of its crystal lattice. Both examples have been realized using Matlab Toolbox `MTEX 5.8`, cf. [5]. The corresponding scripts and data files can be found at `https://github.com/mtex-toolbox/mtex-paper/tree/master/manifoldValuedApproximation`.

4.1. Wave Velocities

In crystalline materials the propagation velocity and polarization direction of waves is often isotropic, i.e., it depends on the propagation direction relative to the crystal lattice. This posses an important issue in seismology where one analyzes the distribution of earthquake waves in order to get a deeper understanding of the core of the earth, cf. [18]. Each earthquake wave decomposes into a p-wave and two perpendicular shear-wave components. The polarization vectors of p-wave components as well as of the two s-wave components depend on the propagation direction of the wave relative to the crystal, cf. [24, 5]. Mathematically, the directional dependency of the polarization directions from the propagation direction is modeled as function

$$f \colon \mathbb{S}^2 \to \mathbb{R}P^2$$

from the two-sphere $\mathbb{S}^2$ into the two–dimensional projective space $\mathbb{R}P^2$. Our goal is to approximate this function from finite measurements $\boldsymbol{y}_n = f(\boldsymbol{x}_n) \in \mathbb{R}P^2$, $n = 1, \dots, N$.

To this end, we identify the two dimensional projective space $\mathbb{R}P^2$ with the quotient $\mathbb{S}^2/\!\sim$ with respect to the equivalence relation $\boldsymbol{x} \sim -\boldsymbol{x}$ and consider the embedding $\mathcal{E} \colon \mathbb{S}^2/\!\sim \to \mathbb{R}^{3\times 3}$, $\mathcal{E}(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{x}^\top$. The reach of this embedding is $\tau = \frac{1}{\sqrt{2}}$ as we show in the following lemma.

Lemma 4.1. *The two dimensional projective space $\mathbb{R}P^2$ embedded into the space of symmetric 3×3 matrices $\mathcal{E}(\mathbb{S}^2/\!\sim) \subset \mathbb{R}^{3\times 3}$ has the reach $\tau = \frac{1}{\sqrt{2}}$.*

Proof. Following [1, Thm. 2.2], we can estimate the reach by the following infimum

$$\tau = \inf_{\boldsymbol{x} \neq \boldsymbol{y} \in \mathbb{S}^2/\!\sim} \frac{\|\mathcal{E}(\boldsymbol{x}) - \mathcal{E}(\boldsymbol{y})\|_2^2}{2\, d(\mathcal{E}(\boldsymbol{x}) - \mathcal{E}(\boldsymbol{y}), T_{\boldsymbol{y}}\mathbb{S}^2/\!\sim)}. \tag{4.1}$$

Since in our setting in both spaces, $\mathbb{S}^2/\!\sim$ and $\mathbb{R}^{3\times 3}$ the metric is invariant with respect to the action of SO(3), it suffices to take the infimum for $\boldsymbol{y} = \boldsymbol{e}_1 = (1,0,0)^\top$. We define the other canonical basis vectors in $\mathbb{R}^3$ as $\boldsymbol{e}_2 = (0,1,0)^\top$ and $\boldsymbol{e}_3 = (0,0,1)^\top$ The tangent vector space in $\boldsymbol{e}_1$ is then given by these tangent vectors:

$$T_{(1,0,0)^\top}\mathbb{S}^2/\!\sim \;= \operatorname{span}\left\{\frac{1}{\sqrt{2}}(\boldsymbol{e}_2\boldsymbol{e}_1^\top + \boldsymbol{e}_1\boldsymbol{e}_2^\top), \frac{1}{\sqrt{2}}(\boldsymbol{e}_3\boldsymbol{e}_1^\top + \boldsymbol{e}_1\boldsymbol{e}_3^\top)\right\}.$$

Hence, we write $\boldsymbol{v} = \mathcal{E}(\boldsymbol{x}) - \mathcal{E}(\boldsymbol{e}_1) = \boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{e}_1\boldsymbol{e}_1^\top$ and therefore we can calculate the reach by

$$\begin{aligned}
\tau &= \inf_{\boldsymbol{x}\neq\boldsymbol{e}_1\in\mathbb{S}^2/\!\sim} \frac{\|\boldsymbol{v}\|_2^2}{2\,\|\boldsymbol{v} - \langle \boldsymbol{v}, \frac{1}{\sqrt{2}}\boldsymbol{e}_2\boldsymbol{e}_2^\top\rangle - \langle \boldsymbol{v}, \frac{1}{\sqrt{2}}\boldsymbol{e}_3\boldsymbol{e}_3^\top\rangle\|_2} \\
&= \inf_{\boldsymbol{x}\neq\boldsymbol{e}_1\in\mathbb{S}^2/\!\sim} \frac{2-2\,x_1^2}{2\,\sqrt{2-2\,x_1^2-2x_1^2x_2^2-2x_1^2x_3^2}} \\
&= \inf_{\boldsymbol{x}\neq\boldsymbol{e}_1\in\mathbb{S}^2/\!\sim} \frac{1-x_1^2}{\sqrt{2}\,\sqrt{(1-x_1^2)^2}} = \frac{1}{\sqrt{2}},
\end{aligned}$$

which finishes the proof. ∎

The calculation of the reach gives us the constants in Theorem 3.1 and 3.2 for this specific manifold.

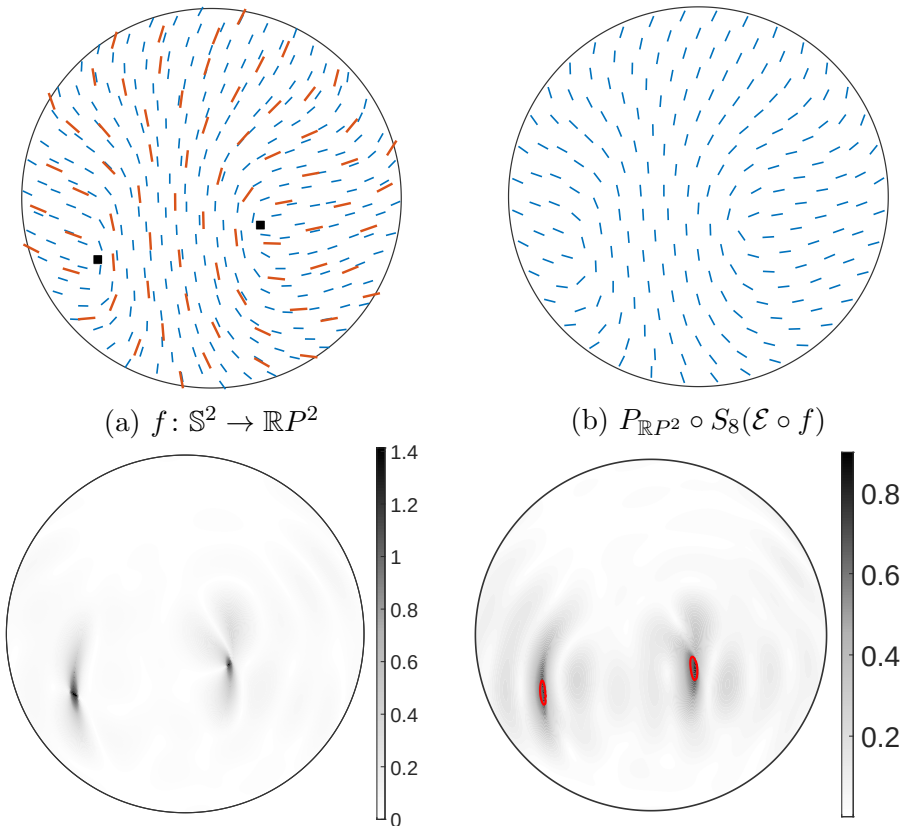
Let $\mathcal{E}\circ f\colon \mathbb{S}^2 \to \mathbb{R}^{3\times 3}$ be the embedded function. A common method, cf. [19], of approximating the spherical function $\mathcal{E}\circ f$ is by linear combinations of spherical harmonics $Y_{\ell,k}$, $\ell = 0,\dots,L$, $k = -\ell,\dots,\ell$ up to a fixed bandwidth L,

$$S_L(\mathcal{E}\circ f)(\boldsymbol{x}) = \sum_{\ell=0}^{L}\sum_{k=-\ell}^{\ell} c_{\ell,k}Y_{\ell,k}(\boldsymbol{x}),$$

where the coefficients $c_{\ell,k}\in\mathbb{R}^{3\times 3}$, $k=-\ell,\dots,\ell$, $\ell=0,\dots,L$ are elements of the embedding space. This expansion in spherical harmonics coincides with the linear Fourier operator in Section 3.2 defined for functions over the sphere. In our little example we simply assume that the measurement points $\boldsymbol{x}_n$ together with some weights ω_n form a spherical quadrature rule up to degree $2L$ which allows us to determine the Fourier coefficients $c_{\ell,k}$, by

$$c_{k,\ell} = \sum_{n=1}^{N}\omega_n\mathcal{E}(\boldsymbol{y}_n)\overline{Y_{\ell,k}(\boldsymbol{x}_n)}.$$

Fig. 1a displays the theoretical polarization directions of an Olivine crystal in dependency of the propagation direction. We observe the points of singularity, marked by the black squares. In order to approximate this non-smooth function we fixed the bandwidth $L = 8$ and used 144 Chebyshev quadrature nodes $\boldsymbol{x}_1,\dots,\boldsymbol{x}_{144}\in\mathbb{S}^2$ as sampling points, cf. [8]. These quadrature nodes are approximately equispaced and are displayed as red lines in Fig. 1a.



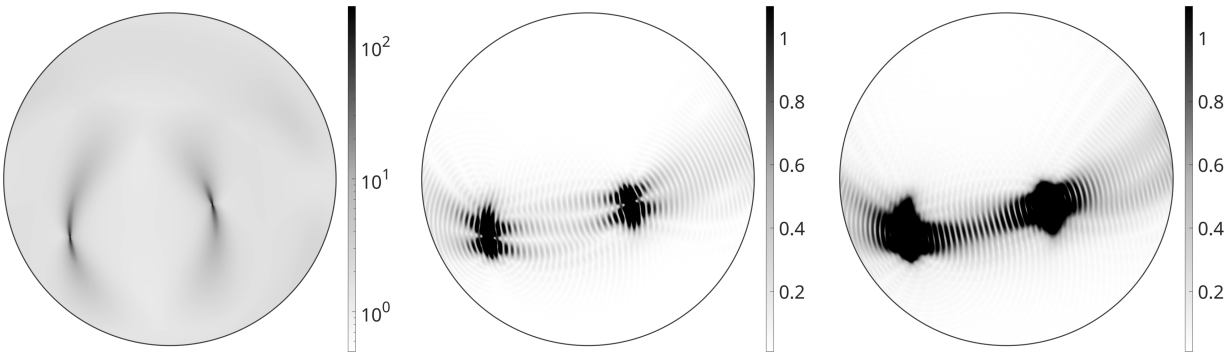
(a) $f\colon \mathbb{S}^2 \to \mathbb{R}P^2$ (b) $P_{\mathbb{R}P^2} \circ S_8(\mathcal{E} \circ f)$

(c) $\|f - P_{\mathbb{R}P^2} \circ S_8(\mathcal{E} \circ f)\|_2$. (d) $\|f - S_8(\mathcal{E} \circ f)\|_2$.

Figure 1.: Polarization directions of the fastest shear wave with respect to the propagation direction $\boldsymbol{x} \in \mathbb{S}^2$ plotted as vector fields on the upper hemisphere. The left upper image (a) displays the true polarization directions $f(\boldsymbol{x}) \in \mathbb{R}P^2$. The right upper image (b) is the harmonic approximation $P_{\mathbb{R}P^2} \circ S_8(\mathcal{E} \circ f)(\boldsymbol{x})$ using the sampling points marked red in (a). The lower left image (c) displays the norm of the point-wise residual. The lower right image (d) displays the error of the linear approximation, i.e. exactly half of the upper bound from Theorem 3.1. We marked red the regions where this residual is bigger than the reach of the manifold.

The approximated function $P_{\mathbb{R}P^2} \circ S_8(\mathcal{E} \circ f)$ is depicted in Fig. 1b and shows good approximation with the original function away from the singularity points. This is supported by a plot of the point-wise error $\|f(\boldsymbol{x}) - P_{\mathbb{R}P^2} \circ S_8(\mathcal{E} \circ f)(\boldsymbol{x})\|_2$ in Fig. 1c. Note that we measure the error in the Euclidean norm of the 9-dimensional embedding space $\mathbb{R}^9$, which is for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}P^2$ equal to the Frobenius norm $\|\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{y}\boldsymbol{y}^\top\|_F$. Compared to this, Figure 1d shows the error of the linear approximation $\|f(\boldsymbol{x}) - S_8(\mathcal{E} \circ f)(\boldsymbol{x})\|$, which is half of the error bound from 3.1. Additionally, we marked the areas where the residual is bigger than the reach blue. In this regions our Theorems are not applicable, since there the assumption $I_{\mathbb{R}^d} f$ is within the reach is not met.

We determined the derivatives of f numerically by choosing a basis $\boldsymbol{t}_1, \boldsymbol{t}_2$ in the tangent



(a) $\|df(x)\|_2$ (b) $\|df(x) - dP_{\mathbb{R}P^2} \circ \tilde{f}\|_2$ (c) Bound from Thm. 3.2.

Figure 2.: The left image (a) shows the point-wise norm of the differential $\mathrm{d}f$. The middle image (b) depicts the approximation error between the differential of the original function f and the differential of its harmonic approximation $P_{\mathbb{R}P^2} \circ \tilde{f} = P_{\mathbb{R}P^2} \circ S_{64}(\mathcal{E} \circ f)$. The right images (c) gives the upper bound for (b) from Theorem 3.2.

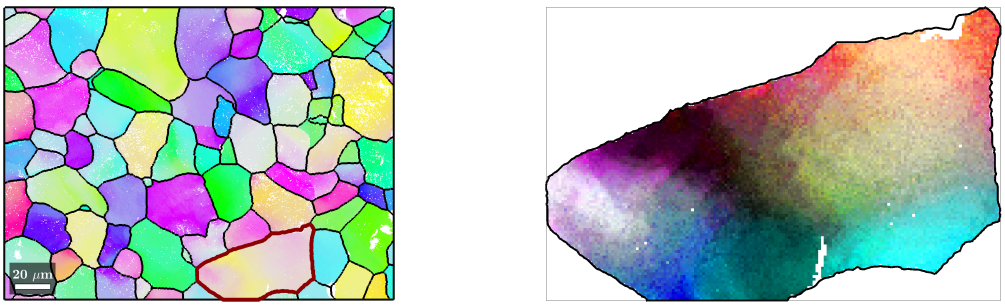
space $T_{\boldsymbol{x}}\mathbb{S}^2$ and approximating the columns $\boldsymbol{v}_i \in \mathbb{R}^9$, $i = 1, 2$ of $\mathrm{d}f(\boldsymbol{x}) \in \mathbb{R}^{9\times 2}$ by the difference quotients

$$\boldsymbol{v}_i = \frac{1}{h}\left(f\left(\tfrac{\boldsymbol{x}+h\,\boldsymbol{t}_i}{\|\boldsymbol{x}+h\,\boldsymbol{t}_i\|_2}\right) - f(\boldsymbol{x})\right),$$

with $h = 10^{-6}$. The norm of the derivative is depicted in Fig. 2a and clearly shows the position of the singularities. In Fig. 2b the error between $\mathrm{d}f(\boldsymbol{x})$ and the differential of harmonic approximation $\mathrm{d}(P_{\mathbb{R}P^2} \circ S_{64})(\mathcal{E} \circ f)(\boldsymbol{x})$ is plotted as a function of the propagation direction $\boldsymbol{x} \in \mathbb{S}^2$. Since the differential $\mathrm{d}f(\boldsymbol{x})$ is a matrix in $\mathbb{R}^{9\times 2}$, we consider here the spectral norm of the error matrix. In order to illustrate our theoretical result of Theorem 3.2 we plotted our theoretical upper bound on that approximation error of the derivative in Fig. 2c. It should be noted that for the differential we needed to increased the polynomial degree to $L = 64$ with 21000 sample points in order to obtain a reasonable approximation at some distance to the singularities.

4.2. Electron Back Scatter Diffraction

The subject of crystallographic texture analysis is the microstructure of polycrystalline materials. Locally the microstructure is described by the orientation of the atom lattice with respect to some specimen fixed reference frame. More specifically, one describes the local orientation of the atom lattice by a coset $[\boldsymbol{R}]_{\mathcal{S}} \in \mathrm{SO}(3)/\mathcal{S}$ of the rotation group $\mathrm{SO}(3)$ modulo the finite subgroup $\mathcal{S} \subset \mathrm{SO}(3)$, called point group. The point group of a crystal consists of all symmetries of its atom lattice and is either one of the cyclic groups C_1, C_2, C_3, C_4, C_6, the dihedral groups D_2, D_3, D_4, D_6, the tetragonal group T or the octahedral group O. Assuming a monophase material, i.e., a material consisting only of a single type of crystals, the variation of the local orientation of the atom lattice at the



(a) full map with a global color key (b) single grain with a local color key

Figure 3.: The raw EBSD data. Each of the 410 x 547 pixels corresponds to a single orientation measurement at the surface of the specimen. The color is computed by the procedure described in [23]. The 5% white pixels in Fig. (a) correspond to corrupted measurements with no data. The data has been segmented into 92 grains as outlined by the black boundaries.

surface $\Omega \subset \mathbb{R}^2$ of the specimen is modeled by the map

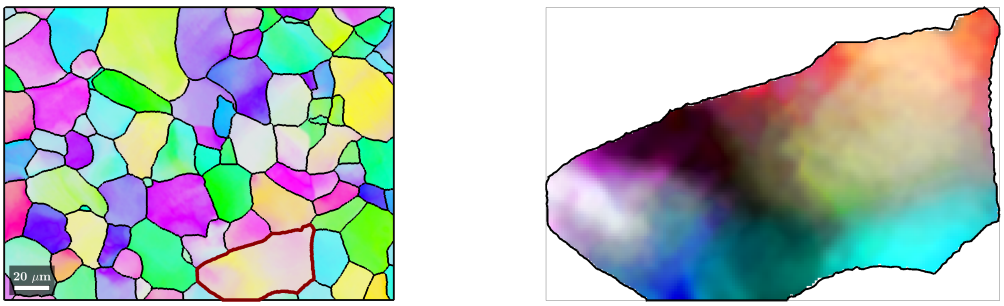
$$f \colon \Omega \to \mathrm{SO}(3)/\mathcal{S}.$$

The gradient of the function f, also called lattice curvature tensor $\boldsymbol{\kappa}(\boldsymbol{x})$, is closely related to elastic and plastic deformations the specimen has been exposed to. More specifically, it is related via the Nye equation to the dislocation density tensor $\boldsymbol{\alpha}(\boldsymbol{x})$, [25, 15] that describes how many lattice dislocations are geometrically necessary in order to preserve the compatibility of the lattice for a given deformation. Hence, estimating f and its derivatives from experimental data is a central problem in material science.

Electron back scatter diffraction (EBSD) is an experimental technique [2, 16] for determining the local lattice orientations $f(\boldsymbol{x}_\ell) \in \mathrm{SO}(3)/\mathcal{S}$ at discrete sampling points $\boldsymbol{x}_{i,j} \in \Omega$. An example of such EBSD data is the $\mathrm{SO}(3)/\mathcal{S}$ - valued image displayed in Fig. 3. It describes the variation of lattice orientation at the surface of an Aluminum alloy of size $200\mu m \times 150\mu m$ at an resolution of $0.4\mu m$. The symmetry group in this case is the octahedral group O.

The data is displayed with respect to two different color keys. In Fig. 3a the colors are assigned globally to the cosets $f(\boldsymbol{x}) \in \mathrm{SO}(3)/O$ as described in [23]. Regions of similar lattice orientation form so-called grains as outlined by the black boundaries. In Fig. 3b only the single grain outlined by the red boundary in Fig. 3a is displayed. For this grain we computed an average lattice orientation $[\boldsymbol{M}] \in \mathrm{SO}(3)/O$ and selected for each coset $f(\boldsymbol{x}) \in \mathrm{SO}(3)/O$ the rotation $\boldsymbol{R}(x) \in [\boldsymbol{M}^{-1} f(\boldsymbol{x})]$ with the smallest rotational angle. Next we associated a color to $\boldsymbol{R}(x)$ according to a spherical color representation where the rotational angle of $\boldsymbol{R}(x)$ determines the saturation and the rotational axis hue and value. More details on this orientation coloring can be found in [31].

Estimating the derivative from such a noisy map of lattice orientations is usually not a good idea as we will see later. Reducing the noise by means of local approximation methods has been discussed in [14, 28]. In order to demonstrate our embedding based



(a) full map of the approximated data (b) single grain with a local color key

Figure 4.: EBSD map from Fig. 3 after the embedding based approximation approach.

approximation approach we make use of the locally isometric embedding $\mathcal{E}_O\colon \mathrm{SO}(3)/O \to \mathbb{R}^9$ described in [13] and proceed as follows

i) Compute an $\mathbb{R}^9$-valued image $\boldsymbol{u}_{i,j} = \mathcal{E}_O(f(\boldsymbol{x}_{i,j}))$.

ii) Approximate the $\mathbb{R}^9$-valued image using a cosine series $\tilde{u}\colon \Omega \to \mathbb{R}^9$ computed by robust, penalized least squares [6].

iii) Evaluate the function $\tilde{u}$ at the grid points $\boldsymbol{x}_{i,j}$ to obtain a noise reduced $\mathbb{R}^9$-valued image $\tilde{\boldsymbol{u}}_{i,j}$.

iv) Compute the projection of $\tilde{\boldsymbol{u}}_{i,j}$ onto the embedding $\mathcal{E}_O(\mathrm{SO}(3)/O)$ of the quotient and apply the inverse map $\mathcal{E}_O^{-1}$ to end up with a noise reduced $\mathrm{SO}(3)/O$-valued image $\tilde{f}(\boldsymbol{x}_{i,j})$.

The resulting image is depicted in Figure 4. We observe that all no data pixels have been inpainted and that the magnified part 4b is much less noisy in comparison to Fig. 3b.

For the computation of the lattice curvature tensor $\boldsymbol{\kappa}$ we use the skew symmetric matrices

$$\boldsymbol{s}^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{s}^{(2)} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{s}^{(3)} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

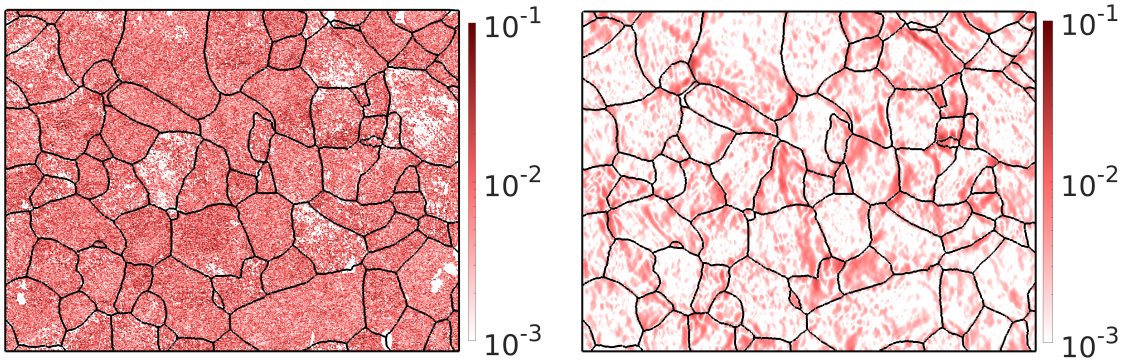
to fix the basis $\boldsymbol{R}\boldsymbol{s}^{(1)}, \boldsymbol{R}\boldsymbol{s}^{(2)}, \boldsymbol{R}\boldsymbol{s}^{(3)}$ in the tangential space $T_{\boldsymbol{R}}\mathrm{SO}(3)/O$ at some rotation $\boldsymbol{R} \in \mathrm{SO}(3)$. With respect to this basis the differential $D\mathcal{E}(\boldsymbol{R})\colon T_{\boldsymbol{R}}\mathrm{SO}(3)/O \to \mathbb{R}^9$ of the embedding $\mathcal{E}\colon \mathrm{SO}(3)/O \to \mathbb{R}^9$ can be represented as a full rank 3×9 matrix. Furthermore, we obtain for the differential $D\tilde{u}\colon \mathbb{R}^2 \to \mathbb{R}^9$ of the embedded image $\tilde{u} = \mathcal{E} \circ \tilde{f}\colon \Omega \to \mathbb{R}^9$ at some point $\boldsymbol{x} \in \Omega$ the matrix product $D\tilde{u}(\boldsymbol{x}) = D\mathcal{E}(\tilde{f}(\boldsymbol{x}))Df(\boldsymbol{x})$. Hence, the lattice curvature tensor $\tilde{\boldsymbol{\kappa}}$ of the noise reduced EBSD map evaluates to

$$\tilde{\boldsymbol{\kappa}}(\boldsymbol{x}) = D\tilde{f}(\boldsymbol{x}) = \left(D\mathcal{E}(\tilde{f}(\boldsymbol{x}))D\mathcal{E}(\tilde{f}(\boldsymbol{x}))^\top\right)^{-1} D\mathcal{E}(\tilde{u}f(\boldsymbol{x}))^\top D\tilde{u}(\boldsymbol{x}).$$

The map of the first component $\tilde{\kappa}_{1,1}$ of the lattice curvature tensor obtained from the approximating function $\tilde{u}$ is depicted in Fig. 5b. For comparison we plotted in Fig. 5a a finite difference approximation

$$\kappa(\boldsymbol{x}_{i,j})_{1,1} = \frac{\log_{f(\boldsymbol{x}_{i,j})}(f(\boldsymbol{x}_{i+1,j}))}{[x_{i+1,j} - x_{i,j}]_1},$$

derived from the discrete data $f(\boldsymbol{x}_{i,j})$. Here, we denoted by $\log_{\boldsymbol{R}}\colon \mathrm{SO}(3)/O \to T_{\boldsymbol{R}}\mathrm{SO}(3)/O$ the logarithmic mapping with respect to the base point $\boldsymbol{R} \in \mathrm{SO}(3)/O$. As expected, we observe that the lattice curvature tensor $\tilde{\kappa}$ derived from the approximated map $\tilde{u}$ is much less noisy.



(a) $\kappa_{1,1}$ of noisy map.

(b) $\kappa_{1,1}$ of noise reduced map.

Figure 5.: First coefficient $\tilde{\kappa}_{1,1}(\boldsymbol{x})$ of the lattice curvature tensor of the $SO(3)/O$-valued map depicted in Fig. 3 (left) and Fig. 4 (right).

Conclusion and further directions

We proposed a method for approximating a manifold-valued function using an embedding approach and a generic approximation operator $I_{\mathbb{R}^d}$ into the Euclidean space. Our main result are the Theorems 3.1 and 3.2 which give upper bounds on the approximation error for the function values as well as for the derivatives. The central requirement of the Theorems is that generic approximation $I_{\mathbb{R}^d} f(\boldsymbol{x})$ has distance less than τ to the manifold $\mathcal{M}$. For the approximation error of the function values it is only important that $I_{\mathbb{R}^d} f(\boldsymbol{x})$ is within the reach of the manifold $\mathcal{M}$ and has no impact on the convergence rate or the constants. However, for the approximation error of the derivatives the constant of the upper bound becomes arbitrary large for $I_{\mathbb{R}^d} f(\boldsymbol{x})$ close to the reach. This stresses the importance of finding embeddings with a large reach.

The basis of our approximation approach is to find a suitable embedding of the manifold $\mathcal{M}$ into $\mathbb{R}^d$. For an arbitrary manifold this can be a difficult challenge. However, for many important manifolds low dimensional embeddings are well known. A further challenge is the numerical realization of the projection operator $P_{\mathcal{M}}$ on the manifold needed for our embedding based approximation method. This leads to a problem of manifold optimization.

So far we did not deal with noisy data. The main challenge here is to guaranty that $I_{\mathbb{R}^d} f(\boldsymbol{x})$ is sufficiently close to the manifold even for noisy data. Up to this point it is not clear how strongly noise that keeps the data on the manifold can increase the distance of $I_{\mathbb{R}^d} f(\boldsymbol{x})$ to the manifold.

Acknowledgments

The authors would like to thank Prof. Dr. Philipp Reiter for the nice hint for completing Theorem 2.9. Furthermore, we thank the anonymous reviewers for providing helpful comments and suggestions to improve this article. The second author acknowledges funding by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 416228727 - SFB 1410.

A. Bound for the commutator

To bound the term $\| \mathrm{d}P_{\mathcal{M}}(\boldsymbol{m}) - \mathrm{d}P_{\mathcal{M}}(\boldsymbol{z}) \|_2$ in section 2.4 we need a lemma, which is based on linear algebra.

Lemma A.1. *Let $\boldsymbol{T}$ be a projection matrix and $\boldsymbol{R}$ be a rotation matrix. Then there holds for the commutator*

$$\|\boldsymbol{TR} - \boldsymbol{RT}\|_2 \leq \|\boldsymbol{I} - \boldsymbol{R}\|_2 ,$$

where again $\|\cdot\|_2$ denotes the spectral norm.

Proof. Since the spectral norm doesn't change under change of basis, we choose a matrix representation where the projection matrix has the form

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

where $\boldsymbol{I}$ is the identity matrix of dimension D. Then we also write the rotation matrix $\boldsymbol{R}$ in these blocks:

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} \end{pmatrix}.$$

Simple matrix multiplication yields because of the orthogonality of $\boldsymbol{R}$

$$\begin{aligned}(\boldsymbol{TR} - \boldsymbol{RT})^\top (\boldsymbol{TR} - \boldsymbol{RT}) &= \boldsymbol{R}^\top \boldsymbol{TR} - \boldsymbol{R}^\top \boldsymbol{TRT} - \boldsymbol{TR}^\top \boldsymbol{TR} + \boldsymbol{T} \\ &= \begin{pmatrix} \boldsymbol{I} - \boldsymbol{R}_{11}^\top \boldsymbol{R}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_{12}^\top \boldsymbol{R}_{12} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_{12}^\top \boldsymbol{R}_{12} \end{pmatrix}\end{aligned}$$

On the other hand there holds, again with help of the orthogonality of $\boldsymbol{R}$,

$$\begin{aligned}(\boldsymbol{I} - \boldsymbol{R})^\top (\boldsymbol{I} - \boldsymbol{R}) = 2\boldsymbol{I} - \boldsymbol{R} - \boldsymbol{R}^\top &= \begin{pmatrix} 2\boldsymbol{I} - \boldsymbol{R}_{11} - \boldsymbol{R}_{11}^\top & -\boldsymbol{R}_{12} - \boldsymbol{R}_{21}^\top \\ -\boldsymbol{R}_{21} - \boldsymbol{R}_{12}^\top & 2\boldsymbol{I} - \boldsymbol{R}_{22} - \boldsymbol{R}_{22}^\top \end{pmatrix} \\ &= \begin{pmatrix} (\boldsymbol{I} - \boldsymbol{R}_{11}^\top)(\boldsymbol{I} - \boldsymbol{R}_{11}) + \boldsymbol{I} - \boldsymbol{R}_{11}^\top \boldsymbol{R}_{11} & -\boldsymbol{R}_{12} - \boldsymbol{R}_{21}^\top \\ -\boldsymbol{R}_{21} - \boldsymbol{R}_{12}^\top & (\boldsymbol{I} - \boldsymbol{R}_{22}^\top)(\boldsymbol{I} - \boldsymbol{R}_{22}) + \boldsymbol{I} - \boldsymbol{R}_{22}^\top \boldsymbol{R}_{22} \end{pmatrix} \\ &= \begin{pmatrix} (\boldsymbol{I} - \boldsymbol{R}_{11}^\top)(\boldsymbol{I} - \boldsymbol{R}_{11}) + \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} & -\boldsymbol{R}_{12} - \boldsymbol{R}_{21}^\top \\ -\boldsymbol{R}_{21} - \boldsymbol{R}_{12}^\top & (\boldsymbol{I} - \boldsymbol{R}_{22}^\top)(\boldsymbol{I} - \boldsymbol{R}_{22}) + \boldsymbol{R}_{12}^\top \boldsymbol{R}_{12} \end{pmatrix}.\end{aligned}$$

The spectral norm of a matrix $\boldsymbol{A}$, i.e., the largest absolute value of the eigenvalues can be written as

$$\|\boldsymbol{A}\|_2^2 = \max_{\|\boldsymbol{x}\|_2 = 1} \|\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A} \boldsymbol{x}\|_2 .$$

For that reason we choose the vector $\boldsymbol{x}$ as the eigenvector of the matrix $(\boldsymbol{TR} - \boldsymbol{RT})^\top (\boldsymbol{TR} - \boldsymbol{RT})$. Since the eigenvalues and eigenvectors of a block-diagonal matrix are the union of the eigenvalues and eigenvectors, i.e., there holds $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{0} \end{pmatrix}^\top$ or $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{x}_2 \end{pmatrix}^\top$. We assume the first case, the other one is analog. Hence, there holds

$$\|\boldsymbol{TR} - \boldsymbol{RT}\|_2^2 = \boldsymbol{x}^\top \begin{pmatrix} \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_{12}^\top \boldsymbol{R}_{12} \end{pmatrix} \boldsymbol{x} = \boldsymbol{x}_1^\top \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} \boldsymbol{x}_1.$$

If we look at the norm of the matrix $\boldsymbol{I} - \boldsymbol{R}$, we get

$$\begin{aligned}\|\boldsymbol{I} - \boldsymbol{R}\|_2^2 &\geq \boldsymbol{x}^\top (\boldsymbol{I} - \boldsymbol{R})^\top (\boldsymbol{I} - \boldsymbol{R}) \boldsymbol{x} = \boldsymbol{x}_1^\top \left((\boldsymbol{I} - \boldsymbol{R}_{11}^\top)(\boldsymbol{I} - \boldsymbol{R}_{11}) + \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} \right) \boldsymbol{x}_1 \\ &\geq \boldsymbol{x}_1^\top \boldsymbol{R}_{21}^\top \boldsymbol{R}_{21} \boldsymbol{x}_1,\end{aligned}$$

since the eigenvalues of $(\boldsymbol{I} - \boldsymbol{R}_{11}^\top)(\boldsymbol{I} - \boldsymbol{R}_{11})$ are positive. Putting this together and taking the square root, yields the assertion. ■

References

[1] E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, and L. Wasserman. Estimating the reach of a manifold. *Electron. J. Statist.*, 13(1):1359–1399, 2019.

[2] B. L. Adams, S. I. Wright, and K. Kunze. Orientation imaging: The emergence of a new microscopy. *Metall. Mater. Trans. A*, 24:819–831, 1993.

[3] J.-D. Boissonnat, A. Lieutier, and M. Wintraecken. The reach, metric distortion, geodesic convexity and the variation of tangent spaces. *J. Appl. Comput. Topol.*, 3:1–30, 07 2019.

[4] A. Constantin. *Fourier Analysis: Volume 1, Theory*, volume 85. Cambridge University Press, 2016.

[5] D.Mainprice, R. Hielscher, and H. Schaeben. Calculating anisotropic physical properties from texture data using the MTEX open source package. *Geological Society, London, Special Publications*, 360:175–192, 2011.

[6] D. Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput. Statist. Data Anal.*, 54(4):1167–1178, 2010.

[7] E. S. Gawlik and M. Leok. Embedding-based interpolation on the special orthogonal group. *SIAM J. Sci. Comput.*, 40(2):A721–A746, 2018.

[8] M. Gräf. Quadrature rules on manifolds. `http://homepage.univie.ac.at/manuel.graef/quadrature.php`, 2013.

[9] P. Grohs. Smoothness of interpolatory multivariate subdivision in lie groups. *IMA J. Numer. Anal.*, 29(3):760–772, 2009.

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

[11] P. Grohs and M. Sprecher. Projection-based quasiinterpolation in manifolds. 2013. Technical Report 2013-23, Seminar for Applied Mathematics, ETH Zürich, Switzerland.

[12] P. Grohs, M. Sprecher, and T. Yu. Scattered manifold-valued data approximation. *Numer. Math.*, 135:987–1010, 2017.

[13] R. Hielscher and L. Lippert. Locally isometric embeddings of quotients of the rotation group modulo finite symmetries. *J. Multivar. Anal.*, 185:104764, 2021.

[14] R. Hielscher, C. Silbermann, E. Schmiedl, and J. Ihlemann. Denoising of crystal orientation maps. *J. Appl. Cryst.*, 52, 2019.

[15] P. Konijnenberg, S. Zaefferer, and D. Raabe. Assessment of geometrically necessary dislocation levels derived by 3d ebsd. *Acta Materialia*, 99:402 – 414, 2015.

[16] K. Kunze, S. I. Wright, B. L. Adams, and D. J. Dingley. Advances in automatic EBSP single orientation measurements. *Textures and Microstructures*, 20:41–54, 1993.

[17] J. Lee. *Introduction to Smooth Manifolds*. Springer-Verlag New York, second edition, 2012.

[18] D. Mainprice. Seismic Anisotropy of the Deep Earth from a Mineral and Rock Physics Perspective. In *Treatese of Geophysics, vol. 2*, pages 437–491. Elsevier, 2007.

[19] V. Michel. *Lectures on Constructive Approximation: Fourier, Spline, and Wavelet Methods on the Real Line, the Sphere, and the Ball*. Birkhäuser, New York, 2013.

[20] M. Moakher. Means and averaging in the group of rotations. *SIAM J. Matrix Anal. Appl*, 24, 04 2002.

[21] J. Nash. C1 isometric imbeddings. *Ann. Math.*, 60(3):383–396, 1954.

[22] P. Niyogi, S. Smale, and S. Weinberger. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete Comput. Geom.*, 39:419–441, 03 2008.

[23] G. Nolze and R. Hielscher. Orientations – perfectly colored. *J. Appl. Cryst.*, 49:1786–1802, 2016.

[24] J. F. Nye. *Physical Properties of Crystals: Their Representation by Tensors 852 and Matrices*. Oxford Univ. Press, England, 2nd ed. edition, 1985.

[25] W. Pantleon. Resolving the geometrically necessary dislocation content by conventional electron backscattering diffraction. *Scripta Materialia*, 58:994–997, 2008.

[26] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier Analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2018.

[27] A. Sarlette and R. Sepulchre. Consensus optimization on manifolds, 2008.

[28] A. Seret, C. Moussa, M. Bernacki, J. Signorelli, and N. Bozzolo. Estimation of geometrically necessary dislocation density from filtered EBSD data by a local linear adaptation of smoothing splines. *J. Appl. Crystallogr.*, 52(3):548–563, Jun 2019.

[29] N. Sharon and U. Itai. Approximation schemes for functions of positive-definite matrix values. *IMA J. Numer. Anal.*, 33(4):1436–1468, 04 2013.

[30] T. Shingel. Interpolation in special orthogonal groups. *IMA J. Numer. Anal.*, 29(3):731–745, 07 2008.

[31] K. Thomsen, K. Mehnert, P. W. Trimby, and A. Gholinia. Quaternion-based disorientation coloring of orientation maps. *Ultramicroscopy*, 182:62–67, 2017.

[32] J. Wallner and N. Dyn. Convergence and C1 analysis of subdivision schemes on manifolds by proximity. *Comput. Aided Geom. Des.*, 22(7):593–622, 2005. Geometric Modelling and Differential Geometry.

[33] G. Xie and T. P.-Y. Yu. Smoothness equivalence properties of manifold-valued data subdivision schemes based on the projection approach. *SIAM J. Numer. Anal.*, 45:1200–1225, 2007.