Bending energy regularization on shape spaces: a class of iterative methods on manifolds and applications to inverse obstacle problems

Julian Eckhardt

Bending energy regularization on shape spaces: a class of iterative methods on manifolds and applications to inverse obstacle problems

Dissertation

zur Erlangung des
mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium"

im Promotionsstudiengang *Mathematical Sciences*
der Georg-August University School of Science (GAUSS)

vorgelegt von
Julian Eckhardt
aus Kassel

Göttingen, 2019

Betreuungsausschuss

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Prof. Dr. Max Wardetzky
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission

Referent:
Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Koreferent:
Prof. Dr. Max Wardetzky
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission

Jun.-Prof. Dr. Madeleine Jotz Lean
Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Russell Luke
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Dr. Frank Werner
Institut für Mathematische Stochastik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 11. September 2019

Abstract

In applications such as nondestructive testing, geophysical exploration or medical imaging one often aims to reconstruct the boundary curve of a smooth bounded domain from indirect measurements. As a typical example we concentrate here on inverse obstacle scattering problems.

We introduce a class of shape manifolds for describing admissible obstacles and we allow the reconstruction of general, not necessarily star-shaped, curves. By applying the bending energy as regularization term the Tikhonov regularization gain independence of the parameterization.

Moreover, the structure of the shape manifold is investigated. It turns out to be a infinite-dimensional Riemannian manifold and therefore, geometry provides several tools, such as Levi-Civita connection, geodesics, Riemannian exponential map, Riemannian Hessian of a functional and parallel transport. One construction, we focus on, is the second fundamental form for which we give explicit formulas and prove local bounds.

Furthermore, we introduce an iteratively regularized Gauss-Newton method on Riemannian manifolds. In each step we compute an update direction as an element in the tangent space using the derivative of the forward operator, the gradient and the Hessian of a regularizing functional. This update direction is mapped by the Riemannian exponential map onto the manifold. Under a general framework we prove convergence rates of this algorithm for exact and perturbed data. The assumptions appearing in the proof are discussed and mostly verified for inverse obstacle scattering problems.

Numerical simulations demonstrate the benefits of the geometrical approach by using shape manifolds and bending-energy-based regularization to reconstruct non-star-shaped obstacles.

Acknowledgments

First of all, I want to give special thanks to my advisor Thorsten Hohage for introducing me to the world of inverse problems and guiding me through it. This road was fascinating and challenging, but I never want to miss it. I also thank Max Wardetzky for supervising and helping me to get a deeper insight into geometry. I am grateful that I could always ask Henrik Schumacher for help whenever I was once again struggling with the technical details in Riemannian geometry.

Furthermore, I acknowledge the financial support of the German Research Foundation (DFG) by the Research Training Group (RTG) 2088 *Discovering structure in complex data: Statistics meets Optimization and Inverse Problems*. This provided me the possibilities to attend international conferences and workshops for broadening my scientific horizon.

Moreover, I thank my current and former colleagues from the research group *Inverse Problems* and of course from the *Institute for numerical and applied mathematics* for making this working environment incredible and unforgettable. I especially like to thank Simon, Frederic, Anna and Henrik for unbelievable comprehensive proofreading this thesis.

I am eternally grateful to all my friends, who supported me unconditionally, suffer from my moaning and were always patient with me all over the last years and especially while finalizing this thesis. This work would not have been possible without you.

Contents

1 Inverse obstacle scattering problems

"Begin at the beginning," the King said, very gravely,
"and go on till you come to the end: then stop."
— Lewis Carroll, *Alice's Adventures in Wonderland*

This thesis is motivated by inverse obstacle problems, which is a class problems arising in real life, where one seeks to reconstruct information about an obstacle from physical measurements. Here we only consider scattering problems as a generic example, but the principles investigated in the following should also be applicable to other inverse obstacle problems, for example in potential theory (see e.g. [22, 50]).

In inverse obstacle scattering problems one tries to reconstruct the shape of an obstacle from measurements of scattered waves. Such problems, occurring for example in nondestructive testing, geophysical exploration, structural health monitoring or medical imaging, have been studied intensively, see the monographs [8, 9, 32, 48] and references therein.

In practice objects in $\mathbb{R}^3$ are more relevant, but nevertheless we consider only the two-dimensional case in this work. Both it is a more simpler test problem and in parts it can be seen as the limit case of an obstacle in $\mathbb{R}^3$. One example is a three-dimensional shape, which does not change in one coordinate direction. Then a cross section approximates such a cylindric obstacle in $\mathbb{R}^2$.

We start by briefly recalling the physical background for the mathematical formulation of the direct and inverse scattering problems. Assume that the object can be described by a bounded, connected, and simply connected Hölder $C^{1,\beta}$-smooth domain Ω_{int} ($\beta > 0$). Then its unbounded complement $\Omega := \mathbb{R}^2 \setminus \overline{\Omega_{\text{int}}}$ is connected, and the boundary curve will be denoted by $\Gamma = \partial\Omega = \partial\Omega_{\text{int}}$. The propagation of an acoustic wave in a homogeneous, isotropic and inviscid fluid is approximately described by a velocity potential $U(x, t)$ that satisfies the wave equation

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = \Delta U,$$

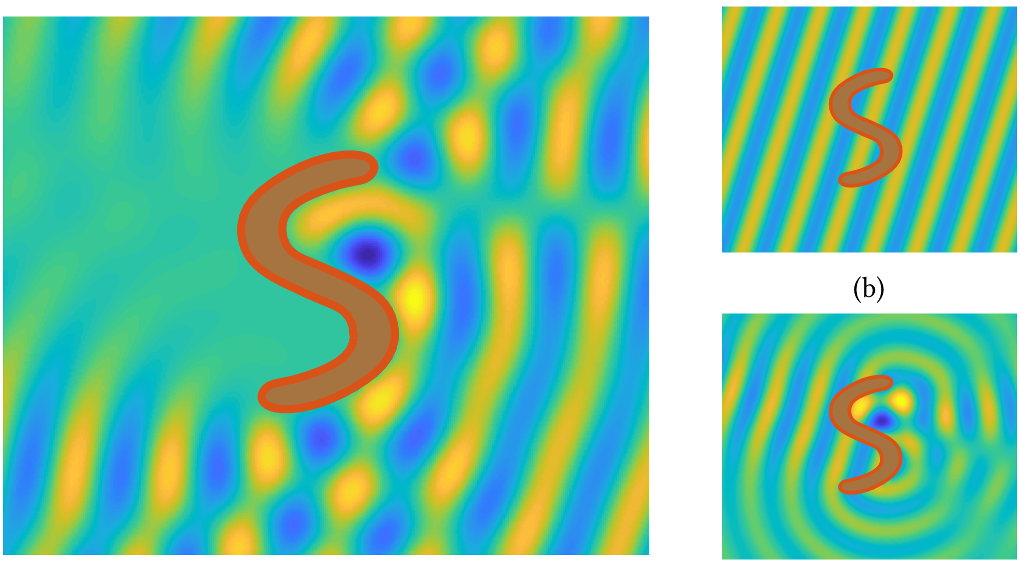
where c denotes the speed of sound and $p = -\frac{\partial U}{\partial t}$ the pressure. In the monograph [9] one can find more information about the physical background. We consider only the time-harmonic case, i.e.

$$U(x, t) = \operatorname{Re}\left(u(x)\, \mathrm{e}^{-\mathrm{i}\,\omega\, t}\right),$$

for $\omega > 0$, and the spatial complex valued function u satisfies the *Helmholtz equation*

$$\Delta u + k^2 u = 0 \qquad \text{in } \Omega, \tag{1.1}$$

where $k := \frac{\omega}{c}$ is the *wave number*. Depending on the physical context the total field u satisfies some boundary condition at Γ. We consider in our generic problem only so-called *sound-soft obstacles*, i.e. the pressure p vanishes on the boundary Γ, which



(a) (b) (c)

Figure 1.1: Example for a total wave (a) with incident wave (b) and scattered wave (c) which is scattered by the obstacle indicated by the red line. For more details of the numerical simulations we refer to the Section 6.2.

can be written as a homogeneous *Dirichlet boundary condition*

$$u = 0 \qquad \text{on } \Gamma. \tag{1.2}$$

Motivated by experimental setups we are interested in case, where we decompose the total field u into a known incident plane wave $u_{\mathrm{i}}(x) = \mathrm{e}^{\mathrm{i}\,k\,\langle x,d\rangle}$ coming from the direction $d \in \mathbb{S}^1$, and a scattered wave $u_{\mathrm{s}} \in H^2_{\mathrm{loc}}(\Omega)$, i.e. the total wave $u := u_{\mathrm{i}} + u_{\mathrm{s}}$ solves (1.1) with (1.2). Additionally the scattered wave satisfies the *Sommerfield radiation condition*

$$\lim_{|x|\to 0} \sqrt{|x|}\Big(\frac{\partial u_{\mathrm{s}}(x)}{\partial |x|} - \mathrm{i}\,k\,u_{\mathrm{s}}(x)\Big) = 0 \tag{1.3}$$

uniformly for all directions. This condition is on the one hand reasonable to assume from the physical point of view since it means that energy is carried away from the scatterer and on the other hand guarantees uniqueness of u_{s}. An example for an total, incident and scattered wave is shown in Figure 1.1.

We recall that solutions to the Helmholtz equation which satisfy the Sommerfield radiation condition (1.3) have the asymptotic behavior

$$u_{\mathrm{s}}(x) = \frac{\mathrm{e}^{\mathrm{i}k|x|}}{\sqrt{|x|}}\Big(u_\infty\big(\tfrac{x}{|x|}, d\big) + O\big(\tfrac{1}{|x|}\big)\Big), \qquad |x| \to \infty \tag{1.4}$$

(see [9, Sect. 2.2 and 3.4]). The function $u_\infty(\cdot, d)$ is analytic on $\mathbb{S}^1$ and known as the *far field pattern* of the scattered wave u_{s}. Often the far field pattern $u_\infty \in L^2(\mathbb{S}^1 \times \mathbb{S}^1)$ can only be measured on some submanifold $\mathbb{M} \subset \mathbb{S}^1 \times \mathbb{S}^1$, e.g. $\mathbb{M} = \mathbb{S}^1 \times \{d\}$ for one incident field or $\mathbb{M} = \{(d, -d) \colon d \in \mathbb{S}^1\}$ for backscattering data. In Figure 1.2 we illustrate two full far field patterns of the obstacle shown in Figure 1.1 for two different wavelengths.

Using the notions introduced above, we can state the *direct problem of obstacle scattering*:

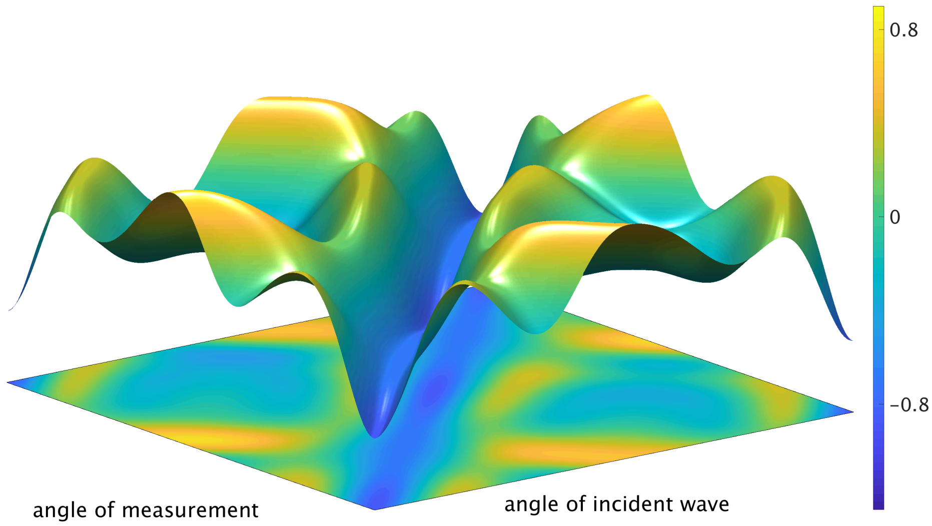
Problem 1.1 (Direct obstacle scattering problem)**.** *For given Γ and u_{i} compute the far field pattern u_∞ of the scattered wave u_{s}, such that $u = u_{\mathrm{i}} + u_{\mathrm{s}}$ solves the Helmholtz equation (1.1) with boundary condition (1.2) and the scattered wave satisfies the Sommerfield radiation condition (1.3).*

This problem is well studied (see e.g. [40]). Numerically, it can be solved for example using a boundary integral equation ansatz (see [9]).

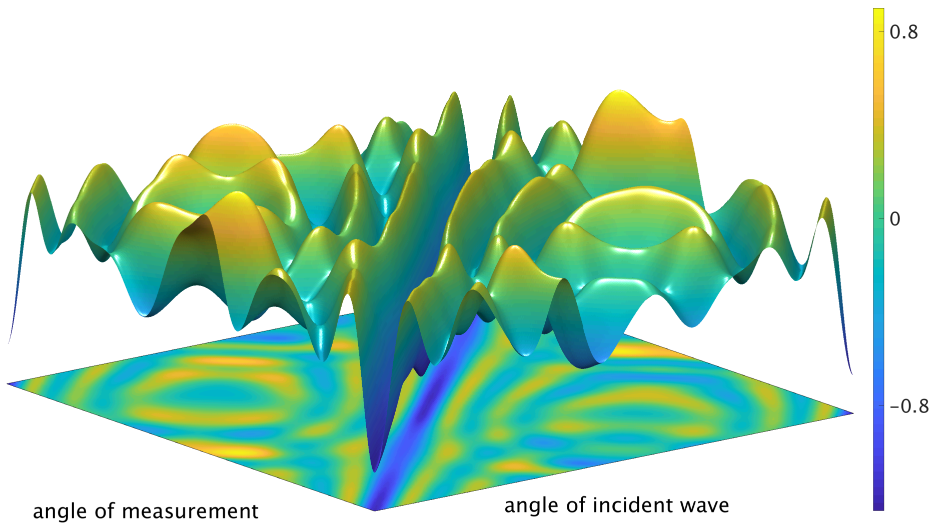
However, in applications the boundary Γ is usually unknown while the far field pattern can be obtained from measurements. Therefore more relevant task is the corresponding *inverse problem*:

Problem 1.2 (Inverse obstacle scattering problem)**.** *Given the incident wave* u_{i} *and either the exact far field pattern* u_∞ *or perturbed measured data* u_∞^δ*, compute the corresponding boundary* Γ *of an obstacle or an approximation to it.*

Solving this problem is challenging by the nonlinearity and ill-posedness. In this thesis we concentrate on so-called *parameterization-based methods*. For these one seeks approximate parameterizations of the unknown shape within a chosen class of boundary curves. One frequently used class in the literature are star-shaped



(a)



(b)

Figure 1.2: The real parts of the far field pattern for the scattered wave in Figure 1.1 under two different wavelength - in (a) we used a shorter one and in (b) the same one as in Figure 1.1. For more details of the numerical simulations we refer to the Section 6.2.

obstacles with respect to a known point such that the boundary can be described by a positive, periodic radial function. In this manner, one can formulate the inverse problem as an operator equation in Hilbert spaces, see Chapter 2. The problem can be solved for example using a Tikhonov regularization penalizing Sobolev norms of the parameterization, see e.g. [9, 24].

To understand this state-of-the-art approach in detail, we will recall some known facts from the literature concerning regularization theory in Chapter 2. In particular we highlight the theory of convergence rates in a general setting both for linear and nonlinear problems. At the end of the chapter we point out how to apply this general setup to inverse obstacle scattering problems.

Note the following: The Sobolev norms or any other regularizing term defined in terms of the radial functions crucially depend on the choice of the parameterization and thus disregard the geometry of the shape to be reconstructed. Indeed, a single curve $\mathbb{S}^1 \to \mathbb{R}^2$ admits a continuum of possible parameterizations and therefore, parameterization-dependent norms break symmetry in an unnatural manner. Moreover, the assumption of star-shaped obstacles is severely restrictive.

In Chapter 3 therefore we introduce a space describing the set of admissible boundaries as a shape manifold from a purely geometric point of view. In the context of shape spaces it is mandatory to consider a set of curves as geometric objects, independent of any particular parameterization. It is shown by Michor and Mumford in [41] that the space of closed and sufficiently regular curves carries a Riemannian manifold structure. Our shape manifolds carry also a natural Riemannian metric and in the rest of the chapter we investigate this structure and state new explicit formulas for the curvature of these shape manifolds.

The following Chapter 4 introduces the scaling invariant bending energy of boundary curves as a regularizing term. This continues the purely geometric approach, which guarantees *independence of the choice of any parameterization.* The bending energy is a second order curvature-based formulation. It has been considered in order to get physically plausible simulations of thin elastic rods and threads [3, 5, 52].

In particular the use of shape manifolds in combination with the bending energy as regularization is preferable over parameterization-dependent methods to overcome the restriction to a subclass of boundary curves and the dependencies arising within. In Chapter 4 we focus on Tikhonov regularization and show that a penalty functional based on shape manifolds and the bending energy is regularizing. However the resulting optimization problem is nonconvex due to the nonlinearity of the forward operator and the bending energy. Hence we do not get a numerically robust algorithm.

This motivates to consider iteratively regularized Gauss-Newton methods as an alternative. Thus we introduce a general algorithmic framework to solve a nonlinear ill-posed operator equation on an infinite-dimensional Riemannian manifold by such an approach in Chapter 5. The main contribution is to prove convergence rates of the algorithm if a source condition is satisfied.

Afterwards we discuss in how far the required assumptions for the general convergence result can be verified for the considered inverse obstacle scattering problems.

Finally in Chapter 6 we give a discretization of the setting in order to solve the problem numerically. The shapes are discretized using polygonal (piecewise straight) curves. Convergence in Hausdorff distance of the resulting minimizers (under suitable boundary conditions and a length constraint) to their smooth counterparts was recently proven in [55], which shows the validity of the chosen discretization approach.

In the second part of the chapter we illustrate practical benefits of our geometrical approach for solving inverse obstacle scattering problems by numerical simulations.

The necessary background from infinite-dimensional Riemannian geometry is provided in the Appendix A.

2 On regularization theory in Hilbert spaces

The formulation of the problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill.
— Albert Einstein, *The Evolution of Physics*

2.1 Inverse problems and regularization

The concept of calling two problems inverse to each other was introduced in [31]. Following this definition one calls a problem *inverse* to another one if the first problem contains the solution to the second one. Such kind of problems arise in various applications in physics. There one usually wants to determine a cause from an observation. In most cases the other problem - predict the observation from a cause - is better understood and hence it is called the *direct problem.*

Hadamard suggested (see [19]) the following definition to classify problems.

Definition 2.1. A problem is called *well-posed* if

1. there exists a solution.
2. the solution is unique.
3. the solution depends continuously on the data.

A problem that is not well-posed is called *ill-posed.*

Either for historical reasons or because of a better understanding of one problem, usually this one is chosen to be the direct problem. In applications it turns out that this one is well-posed most of the time while a corresponding inverse problem is ill-posed. Concerning the ill-posed character of a problem the existence and the uniqueness of a problem is often controllable by either shrinking or enlarging the set of possible solutions, whereas the discontinuous dependence, often referred to the term of stability, is the crucial issue of solving an ill-posed problem. In practice every kind of measurement is affected by noise and therefore data errors have to be taken into account for reconstructing a cause.

In the following we always assume that the forward problem can be written as an operator equation

$$F(f) = g, \tag{2.1}$$

where $F\colon \operatorname{dom}(F) \subset \mathbb{X} \to \mathbb{Y}$ is called the *forward operator* between the topological spaces $\mathbb{X}$, $\mathbb{Y}$. Using this formulation the criteria for well-posedness can be written as

1. F is surjective,
2. F is injective,
3. F^{-1} is continuous.

If F^{-1} is not continuous, small perturbations of g may lead to arbitrarily large perturbations of f. Theoretically one could choose either a finer topology in $\mathbb{Y}$ or a coarser topology in $\mathbb{X}$ to gain continuity of F^{-1}, but in actual applications this does not solve the problem of instability. By choosing different topologies one has to measure the corresponding errors in $\mathbb{X}$ and $\mathbb{Y}$ with respect to these constructed topologies. In applications, the natural choice usually arises from a metric or even a norm. Boundedness of the error can usually be only obtained in certain metrics or norms determined by the measurements process or model. Stronger metrics - which lead to finer topologies - on the other hand lead to unbounded errors.

Therefore one inverts the operator F approximately by a family of continuous operators $\{R_\alpha\}_{\alpha>0}$ such that

$$R_\alpha(g) \to F^{-1}(g) \quad \text{as} \quad \alpha \to 0$$

if the right-hand side is well-defined. For a more general notion of an inverse operator, we use the following definition.

Definition 2.2. Let $F\colon \mathcal{A} \to \mathcal{B}$ be a mapping between sets. A mapping $G\colon \mathcal{B} \to \mathcal{A}$ is called *generalized inverse* of F if

$$F \circ G \circ F = F, \tag{2.2a}$$

$$G \circ F \circ G = G. \tag{2.2b}$$

This enables the definition of a regularization method as follows.

Definition 2.3. Let $(\mathbb{X}, d_{\mathbb{X}})$, $(\mathbb{Y}, d_{\mathbb{Y}})$ be metric spaces, $F\colon \operatorname{dom}(F) \subset \mathbb{X} \to \mathbb{Y}$ a forward operator and $F^\dagger\colon \operatorname{dom}(F^\dagger) \subset \mathbb{Y} \to \operatorname{dom}(F)$ a generalized inverse of F. A pair $(\{R_\alpha\}_{\alpha>0}, \boldsymbol{\alpha})$ of a family of continuous operators $R_\alpha\colon \mathbb{Y} \to \mathbb{X}$ and a *parameter choice rule* $\boldsymbol{\alpha}\colon (0,\infty) \times \mathbb{Y} \to \mathbb{R}_+$ is called a *(deterministic) regularization method* if

$$\limsup_{\delta\to 0} \left\{ d_{\mathbb{X}}\big(R_{\boldsymbol{\alpha}(\delta, g^\delta)}(g^\delta), F^\dagger(g)\big) \,\middle|\, g^\delta \in \mathbb{Y},\ d_{\mathbb{Y}}(g^\delta, g) \le \delta \right\} = 0 \tag{2.3}$$

for all $g \in \operatorname{dom}(F^\dagger)$. $\boldsymbol{\alpha}$ is called an *a-priori parameter choice rule* if $\boldsymbol{\alpha}$ only depends on δ, otherwise it is called an *a-posteriori choice rule.*

In the definition above (2.3) is assumed to hold true for each $g \in \operatorname{dom}(F^\dagger)$. The next theorem shows that (2.3) cannot be formulated uniformly and there is no uniform convergence rate. This means the convergence in (2.3) can be arbitrarily slow. For a proof see for example [13, Prop. 3.11].

Theorem 2.4. *Let $F\colon \mathbb{X} \to \mathbb{Y}$ be injective. Assume for a regularization method $(\{R_\alpha\}_{\alpha>0}, \boldsymbol{\alpha})$ of $F^\dagger$ that there is a continuous function $\varphi\colon [0,\infty) \to [0,\infty)$ with $\varphi(0) = 0$ such that*

$$\sup \left\{ d_{\mathbb{X}}\big(R_{\boldsymbol{\alpha}(\delta, g^\delta)}(g^\delta), F^\dagger(g)\big) \,\middle|\, g^\delta \in \mathbb{Y},\ d_{\mathbb{Y}}(g^\delta, g) \le \delta \right\} \le \varphi(\delta) \tag{2.4}$$

for all $\delta > 0$. Then $F^\dagger$ is continuous.

One of the most prominent examples of a regularization method is Tikhonov regularization. Since it is also the starting point for our approach it will be studied briefly in the following.

2.2 Tikhonov regularization

In this part $\mathbb{X}$, $\mathbb{Y}$ are assumed to be Hilbert spaces and the forward operator to be an injective linear map. To highlight this restriction - also in notation - the forward operator is denoted by T, whenever F is assumed to be linear. The operator equation, to be solved is given by

$$Tf^\dagger = g^\dagger \tag{2.5}$$

with the assumption $g^\dagger \in \operatorname{ran}(T)$. Here $f^\dagger$ always denotes the exact solution in $\mathbb{X}$ and $g^\dagger$ its corresponding data. In this setting the ill-posedness, i.e. discontinuous dependence on the data, is characterized by $\|T^{-1}\| = \infty$. Throughout this thesis the data is assumed to be affected by noise. The standard model is to assume the observed data g^δ satisfies

$$\|g^\delta - g^\dagger\|_{\mathbb{Y}} \leq \delta, \tag{2.6}$$

where $\delta \geq 0$ is some error bound. Solving the equation (2.5) with perturbed data is equivalent to finding the minimum

$$\widehat{f} \in \operatorname*{argmin}_{f \in \mathbb{X}} \|Tf - g^\delta\|^2_{\mathbb{Y}}.$$

This problem is clearly still ill-posed. It corresponds to solving the normal equation $T^*T\widehat{f} = T^*g^\delta$. A well-known idea to stabilize this problem is to add a penalty term in the variational formulation, which goes back to Tikhonov and was proposed in [57, 58]:

$$\widehat{f_\alpha} \in \operatorname*{argmin}_{f \in \mathbb{X}} \mathcal{J}_{g^\delta,\alpha}(f), \quad \text{with} \quad \mathcal{J}_{g^\delta,\alpha}(f) := \|Tf - g^\delta\|^2_{\mathbb{Y}} + \alpha\|f - f_0\|^2_{\mathbb{X}}. \tag{2.7}$$

Today the functional $\mathcal{J}_{g^\delta,\alpha}$ is known as *Tikhonov functional* and α as the *regularization parameter*. The first term, also called the *data fidelity term*, measures how close one approximates the data, whereas the second term, known as *regularizing term*, imposes a-priori knowledge of the problem - here one computes a smooth solution that is close to a given point f_0. The minimizer $\widehat{f_\alpha}$ is only an approximate solution to the problem (2.5), but it turns out that for $\alpha \to 0$ the sequence of minimizers can converge to the exact solution $f^\dagger$.

The rest of this section is concerned with proving that the approach (2.7) forms a regularization method.

The following theorem is well-known in the literature and states that the approach (2.7) is well-defined and a stable approximation. A proof can be found for example in [13, Thm. 5.1].

Theorem 2.5. *The Tikhonov function $\mathcal{J}_{g^\delta,\alpha}$ has a unique minimum $\widehat{f_\alpha}$ for all $\alpha > 0$, $g^\delta \in \mathbb{Y}$ and $f_0 \in \mathbb{X}$, and it is given by*

$$\widehat{f_\alpha} = \left(\alpha I + T^*T\right)^{-1}(T^*g^\delta + \alpha f_0). \tag{2.8}$$

Furthermore $\widehat{f_\alpha}$ depends continuously on the data g^δ.

The next theorem verifies that this forms indeed a regularization method in the sense of Definition 2.3 with the following generalized inverse. The operator $T^\dagger\colon \operatorname{dom}(T^\dagger) \subset \mathbb{Y} \to \mathbb{X}$ is called the *Moore-Penrose inverse* of T if it is a generalized inverse of T and satisfies

$$(TT^\dagger)^* = TT^\dagger, \tag{2.9a}$$

$$(T^\dagger T)^* = T^\dagger T. \tag{2.9b}$$

A proof for the theorem can be found in [13, Thm. 5.2].

Theorem 2.6. *Let $g^\dagger \in \mathrm{dom}(T^\dagger) \subset \mathbb{Y}$ be given, where $T^\dagger$ is the Moore-Penrose inverse of T, and denote $f^\dagger = T^\dagger g^\dagger$. Let $(\delta_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$ be a given sequence such that $\lim_{n\to\infty} \delta_n = 0$ and $(g^{\delta_n})_{n\in\mathbb{N}} \subset \mathbb{Y}$ a corresponding sequence of perturbed data satisfying (2.6). Assume that the regularization parameters are chosen such that*

$$\lim_{n\to\infty} \alpha_n = 0 \qquad \textit{and} \qquad \lim_{n\to\infty} \frac{\delta_n}{\sqrt{\alpha_n}} = 0.$$

Then for $\widehat{f}_n := \widehat{f}_{\alpha_n}$ the minimizer of $\mathcal{J}_{g^{\delta_n},\alpha_n}$ it holds that

$$\lim_{n\to\infty} \|T\widehat{f}_n - g^\dagger\|_{\mathbb{Y}} = 0 \qquad \textit{and} \qquad \lim_{n\to\infty} \|\widehat{f}_n - f^\dagger\|_{\mathbb{X}} = 0 \tag{2.10}$$

and therefore (2.7) forms a regularization method.

In recent years various kinds of approaches came up to generalizes Tikhonov regularization, for example with different regularizing or data fidelity terms as well as extensions to more general spaces such as Banach spaces or even in parts locally convex vector spaces.

For this thesis we want to point out one variation of Tikhonov regularization, since it will play a role in the later work: using weighted norms. Consider the minimization problem

$$\widehat{f}_{\alpha,W} \in \operatorname*{argmin}_{f\in\mathrm{dom}(W)} \mathcal{J}^W_{g^\delta,\alpha}(f), \quad \text{with} \quad \mathcal{J}^W_{g^\delta,\alpha}(f) := \|Tf - g^\delta\|^2_{\mathbb{Y}} + \alpha\|W(f-f_0)\|^2_{\mathbb{W}} \tag{2.11}$$

for a linear operator W on $\mathbb{X}$ into a Hilbert space $\mathbb{W}$. One standard choice for W is a differentiation operator, for example the gradient or Laplace operator. Already from the example it highlights that W does not need to have a trivial null space, which makes this approach challenging. In [13] the theoretical foundations can be found. A necessary assumption is the so-called *complementation condition* (see [43]): There is a constant $c > 0$ such that

$$\|Tf\|^2_{\mathbb{Y}} + \|Wf\|^2_{\mathbb{W}} \geq c\|f\|^2_{\mathbb{X}}$$

for all $f \in \mathbb{X}$. Further the minimizer for (2.11) has the form

$$\widehat{f}_{\alpha,W} = \left(\alpha\, W^*W + T^*T\right)^{-1}(T^*g^\delta + \alpha\, W^*W f_0). \tag{2.12}$$

2.3 Convergence rates theory

In Theorem 2.4 it was shown that for an ill-posed operator equation there can be no uniform convergence rate in the noise level δ. Therefore the approximation by a regularization method, i.e. the convergence in (2.10), can become arbitrarily bad without additional a-priori information.

This motivates to investigate properties of elements $f^\dagger$ in the solution space to the equation (2.5) for which one can control the rate of convergence.

Definition 2.7. A function $\varphi\colon [0,\infty) \to [0,\infty)$ that is continuous and monotonically increasing with $\varphi(0) = 0$ is called an *index function*.

An element $f^\dagger \in \mathbb{X}$ is said to satisfy a *spectral source condition* if there is an index function φ and an element $v \in \mathbb{X}$ such that

$$f^\dagger = \varphi(T^*T)v. \tag{2.13}$$

Here $\varphi(T^*T)$ is defined by the spectral calculus, see e.g. [49]. In literature the two most commonly used index functions are on the one hand

$$\varphi_\nu(\lambda) = \lambda^\nu \tag{2.14}$$

with $\nu > 0$ is called *Hölder source conditions* and on the other hand for the *logarithmic source conditions* the index functions

$$\varphi_p(\lambda) = \begin{cases} \big(-\log\max\{\lambda,\lambda_0\}\big)^{-p} & \lambda > 0, \\ 0 & \lambda = 0, \end{cases} \tag{2.15}$$

with $p > 0$ and some $\lambda_0 \in (0,1)$.

If $\mathbb{X}$ is a function space the condition (2.13) together with (2.14) or (2.15) can often be interpreted as a smoothness condition on $f^\dagger$. In [26] with the forward operator T solving a heat equation it was shown that $f^\dagger \in \operatorname{ran}(\varphi_p(T^*T))$ if and only if $f^\dagger$ is in a Sobolev space of smoothness $2p$ for all $p > 0$. There are also other examples of operators for which one can prove such a statement, see for example [39].

The following notion defines a quasiorder on the set of index functions.

Definition 2.8. An index function φ_0 *covers* an index function φ if there is a constant $C > 0$ such that

$$C\frac{\varphi_0(\alpha)}{\varphi(\alpha)} \le \inf_{\alpha\le\lambda\le 1} \frac{\varphi_0(\lambda)}{\varphi(\lambda)}, \qquad \text{for all } \alpha \in (0,1].$$

In this case one writes $\varphi \succcurlyeq \varphi_0$.

In this way one sorts the index functions by their behavior around zero, i.e. φ_0 covers φ if φ_0 decays faster to zero than φ.

The next theorem is the main result of this section. Under a source condition one gets a rate of convergence of the approximation to the exact solution. The proof can be found in [39, Thm. 5.2].

Theorem 2.9. *Let $f^\dagger$ satisfy (2.13) for some index function φ and $v \in \mathbb{X}$ with $\|v\|_{\mathbb{X}} \leq \rho$. Assume $\varphi \succcurlyeq \mathrm{id}$ and denote $\Psi(\lambda) := \sqrt{\lambda}\varphi(\lambda)$. If $\boldsymbol{\alpha}$ is chosen by the parameter choice rule*

$$\boldsymbol{\alpha} := \Psi^{-1}\left(\frac{\delta}{\rho}\right), \tag{2.16}$$

then there is a constant $C > 0$ such that the minimizer $\widehat{f_{\boldsymbol{\alpha}}}$ of the Tikhonov functional (2.7) with $f_0 = 0$ satisfies the error estimate

$$\|\widehat{f_{\boldsymbol{\alpha}}} - f^\dagger\|_{\mathbb{X}} \leq C\,\rho\,\varphi\left(\Psi^{-1}\left(\frac{\delta}{\rho}\right)\right) \tag{2.17}$$

for all $\delta \in (0, \rho]$.

The main idea how to prove this is splitting the error into

$$\|\widehat{f_\alpha} - f^\dagger\|_{\mathbb{X}} \leq \|\widehat{f_\alpha} - f_\alpha\|_{\mathbb{X}} + \|f_\alpha - f^\dagger\|_{\mathbb{X}}$$

where f_α is the minimizer for the noiseless Tikhonov functional $\mathcal{J}_{g^\dagger,\alpha}$. The terms on the right-hand side are called the *propagated data noise error* and the *approximation error* respectively. One can estimate these two by

$$\|\widehat{f_\alpha} - f_\alpha\|_{\mathbb{X}} \leq C\frac{\delta}{\sqrt{\alpha}}$$

and

$$\|f_\alpha - f^\dagger\|_{\mathbb{X}} \leq C\rho\varphi(\alpha)$$

under the assumptions of Theorem 2.9. Then the parameter choice rule $\boldsymbol{\alpha}$ from above balances the two terms, which yields the estimate above.

Note that the index function φ_ν in (2.14) satisfies $\varphi_\nu \succcurlyeq \mathrm{id}$ if and only if $\nu \leq 1$. Recall that in the case $\nu > 1$ the additional smoothness will not improve convergence rates by the saturation limit of Tikhonov regularization (see [46]). If one wants to get faster convergence rates, other methods have to been used, which are not covered here.

Concerning the logarithmic source conditions φ_p from (2.15) the covering $\varphi_p \succcurlyeq \mathrm{id}$ is true for all $p > 0$ and $\lambda_0 \in (0, 1)$.

From Theorem 2.9 the natural question arises whether all elements $f^\dagger \in \mathbb{X}$ satisfy a source condition. In the case of injective operators the proof for the following theorem can be found in [23].

Theorem 2.10. *Let T be injective and $f^\dagger \in \mathbb{X}$ be given. Then there exists an index function φ and an element $v \in \mathbb{X}$, such that (2.13) holds true and $\|v\|_{\mathbb{X}} \le 2\|f^\dagger\|_{\mathbb{X}}$.*

To conclude this section we shortly come back to the case of Tikhonov regularization with a weighted norm (2.11) with $f_0 = 0$. Assume W is self-adjoint on a dense domain $\operatorname{dom}(W) \subset \mathbb{X}$ and there is a positive constant $c > 0$ such that

$$\|Wf\|_{\mathbb{W}}^2 \ge c\|f\|_{\mathbb{X}}^2$$

for all $f \in \operatorname{dom}(W)$. Then we define the weighted forward operator

$$L := TW^{-1} \tag{2.18}$$

and write for the minimizer $\widehat{f}_{\alpha,W}$ of the Tikhonov functional (2.11) with $f_0 = 0$

$$\widehat{f}_{\alpha,W} = W^{-1}\left(\alpha I + L^*L\right)^{-1} L^* g^\delta. \tag{2.19}$$

We can prove an analog to Theorem 2.9.

Theorem 2.11. *Let $f^\dagger$ satisfy a source condition*

$$f^\dagger = W^{-1}\varphi(L^*L)v = W^{-1}\varphi(W\,T^*T\,W)v \tag{2.20}$$

for some index function φ and $v \in \mathbb{X}$ with $\|v\|_{\mathbb{X}} \le \rho$. Assume $\varphi \succcurlyeq \mathrm{id}$ and denote $\Psi(\lambda) := \sqrt{\lambda}\varphi(\lambda)$. If $\boldsymbol{\alpha}$ is chosen by the parameter choice rule

$$\boldsymbol{\alpha} := \Psi^{-1}\Big(\frac{\delta}{\rho}\Big), \tag{2.21}$$

then there is a constant $C > 0$ such that $\widehat{f}_{\boldsymbol{\alpha},W}$ satisfies the error estimate

$$\|\widehat{f}_{\boldsymbol{\alpha},W} - f^\dagger\|_{\mathbb{X}} \le C\,\rho\,\varphi\Big(\Psi^{-1}\Big(\frac{\delta}{\rho}\Big)\Big) \tag{2.22}$$

for all $\delta \in (0,\rho]$.

Proof. First denote

$$f_{\alpha,W} := W^{-1}\left(\alpha I + L^*L\right)^{-1}(L^*Tf^\dagger + W\alpha f_0)$$

and decompose into approximation and noise error

$$\|\widehat{f}_{\boldsymbol{\alpha},W} - f^\dagger\|_{\mathbb{X}} \le \|\widehat{f}_{\boldsymbol{\alpha},W} - f_{\alpha,W}\|_{\mathbb{X}} + \|f_{\alpha,W} - f^\dagger\|_{\mathbb{X}}.$$

For the noise error one can estimate using $\|(\alpha I + L^*L)^{-1}L^*\| \leq \frac{1}{2\sqrt{\alpha}}$

$$\|\widehat{f}_{\alpha,W} - f_{\alpha,W}\|_{\mathbb{X}} = \left\|W^{-1}\left(\alpha I + L^*L\right)^{-1}L^*\left(g^\delta - Tf^\dagger\right)\right\|_{\mathbb{X}} \leq \|W^{-1}\|\,\frac{\delta}{2\sqrt{\alpha}}.$$

Concerning the approximation error denote $q_\alpha(\lambda) := \frac{1}{\alpha+\lambda}$ and $r_\alpha(\lambda) := 1 - \lambda q_\alpha(\lambda)$. By the assumption $\varphi \succcurlyeq \mathrm{id}$ it follows (see [39]) that there is a constant $C_\varphi < \max\{1, \frac{1}{C}\}$, where C is given by the covering $\varphi \succcurlyeq \mathrm{id}$, such that

$$\sup_{\lambda\in[0,\Lambda]} |r_\alpha(\lambda)|\,\varphi(\lambda) \leq C_\varphi\,\varphi(\alpha) \qquad \text{for all } 0 < \alpha \leq \Lambda. \tag{2.23}$$

Therefore with the functional calculus it follows that

$$\begin{aligned}
\|f_{\alpha,W} - f^\dagger\|_{\mathbb{X}} &= \|W^{-1}\left(\alpha I + L^*L\right)^{-1}L^*\,Tf^\dagger - f^\dagger\|_{\mathbb{X}} \\
&= \|W^{-1}\left(q_\alpha(L^*L)\,L^*L\,Wf^\dagger - Wf^\dagger\right)\|_{\mathbb{X}} \\
&= \|W^{-1}\,r_\alpha(L^*L)\,\varphi(L^*L)v\|_{\mathbb{X}} \\
&\leq C_\varphi\,\|W^{-1}\|\,\|v\|_{\mathbb{X}}\,\varphi(\alpha).
\end{aligned}$$

The application of the parameter choice rule (2.21) yields the assertion. □

2.4 Iteratively regularized Gauß-Newton method

In this section it will be discussed how one can extend some tools and parts of the theory for linear operators T to nonlinear forward maps $F\colon \mathbb{X} \to \mathbb{Y}$.

Whenever the forward operator is differentiable a widely used approach to solve the operator equation (2.1) is via an iterative linearization. In one step one solves the linear equation

$$F(f) \approx F(f_k) + \mathrm{D}F(f_k)\,(f - f_k) = g^\delta$$

to get $f_{k+1} = f$ with $\mathrm{D}F(f)\colon \mathbb{X} \to \mathbb{Y}$ the Fréchet derivative of F at f. This linearized equation is in general still ill-posed and can be treated as (2.5).

If one applies Tikhonov regularization with parameters $\alpha_k > 0$ to this linearized equation one obtains iterations $(f_k)_{k\in\mathbb{N}}$ given by

$$f_{k+1} = f_k + \left(\alpha_k I + \mathrm{D}F(f_k)^*\,\mathrm{D}F(f_k)\right)^{-1}\mathrm{D}F(f_k)^*\left(g^\delta - F(f_k)\right). \tag{2.24}$$

By reformulating it as a minimization problem one can write it as

$$f_{k+1} = \operatorname*{argmin}_{f\in\mathbb{X}}\ \mathcal{J}^{f_k,f_k}_{g^\delta,\alpha_k}(f) \tag{2.25a}$$

with

$$\mathcal{J}_{g^\delta,\alpha_k}^{f_k,f_k}(f) := \|F(f_k) + \mathrm{D}F(f_k)(f - f_k) - g^\delta\|_{\mathbb{Y}}^2 + \alpha_k\|f - f_k\|_{\mathbb{X}}^2. \tag{2.25b}$$

This method is known as *Levenberg-Marquardt algorithm.*

A related algorithm, which was suggested first by Bakushinskii in [4], is given by

$$f_{k+1} = \operatorname*{argmin}_{f\in\mathbb{X}} \mathcal{J}_{g^\delta,\alpha_k}^{f_k,f_0}(f) \tag{2.26a}$$

with

$$\mathcal{J}_{g^\delta,\alpha_k}^{f_k,f_0}(f) = \|F(f_k) + \mathrm{D}F(f_k)(f - f_k) - g^\delta\|_{\mathbb{Y}}^2 + \alpha_k\|f - f_0\|_{\mathbb{X}}^2. \tag{2.26b}$$

Today this algorithm is known as *iteratively regularized Gauss-Newton method.* To ensure convergence the regularization parameters α_k must tend to zero as $k \to \infty$.

One further ingredient to prove convergence or even rates is the notion of a *stopping rule* $\boldsymbol{k}(\delta, g^\delta) = \boldsymbol{k}$. The presence of noise in the data makes it appropriate to stop the iteration at some point, since otherwise the impact of the noise would become dominant. Therefore if one deals with the noise-free case, i.e. $\delta = 0$, usually $K = \infty$. If K does not depend on g^δ it is called an *a-priori stopping rule* and otherwise an *a-posteriori stopping rule.*

The spectral source condition that is typically assumed for the above linearized problems has the form

$$f^\dagger - f_0 = \varphi\big(\mathrm{D}F(f^\dagger)^*\,\mathrm{D}F(f^\dagger)\big)v \tag{2.27}$$

for some $v \in \mathbb{X}$.

In order to prove convergence one needs a nonlinearity condition for the forward operator to guarantee that the local linearizations describe the problem well enough. In the literature there are several approaches. In principle it depends also on the index function in the source condition how restrictive the nonlinearity condition has to be. Using general source conditions of the form (2.27) one needs the following:

Assumption 2.12. *Let* $\widehat{f}, f \in B_\varrho(f^\dagger) = \{f \in \mathbb{X} \mid \|f - f^\dagger\|_{\mathbb{X}} \le \varrho\}$. *There are operators* $S(\widehat{f}, f) \in L(\mathbb{Y})$ *and* $Q(\widehat{f}, f) \in L(\mathbb{X}, \mathbb{Y})$ *and constants* $C_S, C_Q > 0$ *such that*

$$\mathrm{D}F(\widehat{f}) = S(\widehat{f}, f)\,\mathrm{D}F(f) + Q(\widehat{f}, f), \tag{2.28}$$

$$\|I - S(\widehat{f}, f)\|_{\mathbb{Y}} \le C_S, \tag{2.29}$$

$$\|Q(\widehat{f}, f)\| \le C_Q\|\mathrm{D}F(f^\dagger)(\widehat{f} - f)\|_{\mathbb{Y}}. \tag{2.30}$$

In principle one assumes that the derivative of the forward operator at two points differs only up to an operator S, which is not far away from the identity and an additional deviation operator Q, which will be assumed to be sufficiently small.

One can formulate simpler and less restrictive nonlinearity conditions if one deals with less general source conditions. For example if one only assumes a Hölder source condition (2.14) with $\nu \geq \frac{1}{2}$ then it suffices to assume Lipschitz continuity of $\mathrm{D}F$ instead of Assumption 2.12.

The next theorem contains the main convergence statement for the iteratively regularized Gauss-Newton method. The proof can be found in [37].

Theorem 2.13. *Let $f^\dagger - f_0$ satisfy (2.27) for some concave index function $\varphi \succcurlyeq \sqrt{\mathrm{id}}$ and an element $v \in \mathbb{X}$. The sequence of regularization parameters are chosen such that*

$$1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq C_\alpha, \qquad \lim_{k\to\infty} \alpha_k = 0, \qquad \alpha_k > 0 \tag{2.31}$$

with some $C_\alpha > 1$. Moreover the a-priori stopping rule $\boldsymbol{k}(\delta, g^\delta)$ is given by the first index $\boldsymbol{k}$ such that

$$\tau\sqrt{\alpha_{\boldsymbol{k}}}\varphi(\alpha_{\boldsymbol{k}}) \leq \delta \leq \tau\sqrt{\alpha_k}\varphi(\alpha_k), \qquad 0 \leq k < \boldsymbol{k}, \tag{2.32}$$

for some $\tau > 0$. Then in the noiseless case we have

$$\|f_k - f^\dagger\|_{\mathbb{X}} = \mathcal{O}\big(\varphi(\alpha_k)\big), \qquad k \to \infty. \tag{2.33}$$

Furthermore in the case of perturbed data one gets the convergence rate

$$\|f_{\boldsymbol{k}} - f^\dagger\|_{\mathbb{X}} = \mathcal{O}\Big(\Psi^{-1}\Big(\frac{\delta}{\tau}\Big)\Big), \qquad \delta \to 0 \tag{2.34}$$

with $\Psi(\lambda) := \sqrt{\lambda}\varphi(\lambda)$.

Let us sketch the proof of this result. In analogy to the case of linear Tikhonov regularization the main idea how to prove this is to split the error in different types: The error $e_{k+1} := f_{k+1} - f^\dagger$ is decomposed into approximation, noise and nonlinearity errors by

$$\|e_{k+1}\|_{\mathbb{X}} \leq \|e^{\mathrm{app}}_{k+1}\|_{\mathbb{X}} + \|e^{\mathrm{noi}}_{k+1}\|_{\mathbb{X}} + \|e^{\mathrm{nl}}_{k+1}\|_{\mathbb{X}}$$

with

$$\begin{aligned}
e^{\mathrm{app}}_{k+1} &:= -\alpha_k(\alpha_k I + T^* T)^{-1} f^\dagger,\\
e^{\mathrm{noi}}_{k+1} &:= (\alpha_k I + T_k^* T_k)^{-1} T_k^* \left(g^\delta - F(f^\dagger)\right),\\
e^{\mathrm{nl}}_{k+1} &:= (\alpha_k I + T_k^* T_k)^{-1} T_k^* \left(F(f^\dagger) - F(f_k) - T_k e_k\right)\\
&\quad - \alpha_k(\alpha_k I + T_k^* T_k)^{-1} \left(T_k^* (T - T_k) + (T^* - T_k^*) T\right) (\alpha_k I + T^* T)^{-1} f^\dagger,
\end{aligned}$$

where $T_k := \mathrm{D}F(f_k)$, and similarly for the image space errors

$$\|T_{k+1}e_{k+1}\|_{\mathbb{Y}} \leq \|T_{k+1}e_{k+1}^{\mathrm{app}}\|_{\mathbb{Y}} + \|T_{k+1}e_{k+1}^{\mathrm{noi}}\|_{\mathbb{Y}} + \|T_{k+1}e_{k+1}^{\mathrm{nl}}\|_{\mathbb{Y}}.$$

Further estimating these different errors leads to an inequality of the form

$$\|T_{k+1}e_{k+1}\|_{\mathbb{Y}} \leq \tilde{a}\sqrt{\alpha_k}\varphi(\alpha_k) + \tilde{b}\|T_k e_k\|_{\mathbb{Y}} + \tilde{c}\frac{\|T_k e_k\|_{\mathbb{Y}}^2}{\sqrt{\alpha_k}},$$

where the coefficients $\tilde{a}$, $\tilde{b}$ and $\tilde{c}$ depend on the constants $C_S, C_Q, \|v\|_{\mathbb{X}}$ and ϱ which can be chosen small enough if one starts at f_0 close enough to the exact solution $f^\dagger$. From this one can now derive conditions (see for details [37]) to obtain an image space rate

$$\|T_{k+1}e_{k+1}\|_{\mathbb{Y}} \leq C\sqrt{\alpha_k}\varphi(\alpha_k)$$

and therefore get the convergence rate for noiseless case in the theorem. By applying the stopping rule one gets directly the rate with respect the noise level.

2.5 Application to inverse obstacle scattering problems

In this section we describe how one can apply the general framework of regularization theory to inverse obstacle scattering problems described in Problem 1.2.

A common and crucial assumption is that one restricts the set of admissible obstacles to star-shaped domains. For every star-shaped obstacle Ω there is reference point x_0 and a positive periodic radial function $r\colon [0, 2\pi] \to \mathbb{R}_+$ such that

$$\gamma_\Omega(t) = x_0 + r(t)\begin{pmatrix}\cos t\\ \sin t\end{pmatrix} \tag{2.35}$$

for all $t \in [0, 2\pi]$, is a parameterization of $\partial\Omega$. For a sake of simplicity one can always assume that $x_0 = (0, 0)^\top$ by a shift of coordinates if x_0 is known. In this manner we identify the admissible obstacles with the positive periodic radial functions with respect to the origin. One chooses the solution space $\mathbb{X} \subset C^{1,\beta}([0, 2\pi])$ to be subspace of sufficient regularity.

Under this assumption one can define the forward operator

$$F\colon\ \mathrm{dom}(F) \subset \mathbb{X} \to L^2(\mathbb{S}^1), \qquad r \mapsto u_\infty \tag{2.36}$$

mapping a radial function to the solution of the corresponding direct problem 1.1. This approach is intensively discussed in [9]. There it is shown that F is compact, continuous and Fréchet differentiable and even that the derivative $DF(r)$ is injective.

These properties allows to apply the iteratively regularized Gauss-Newton method to solve the nonlinear ill-posed operator equation.

Apart from the applicability of the algorithm the more interesting part is again the question of convergence, convergence rates and source conditions. It is well-known that this problem is exponentially ill-posed. That means that the only source conditions one can expect to be satisfied for a sufficiently general class of objects are of the logarithmic form (2.15). In [25] it is proved that these source conditions

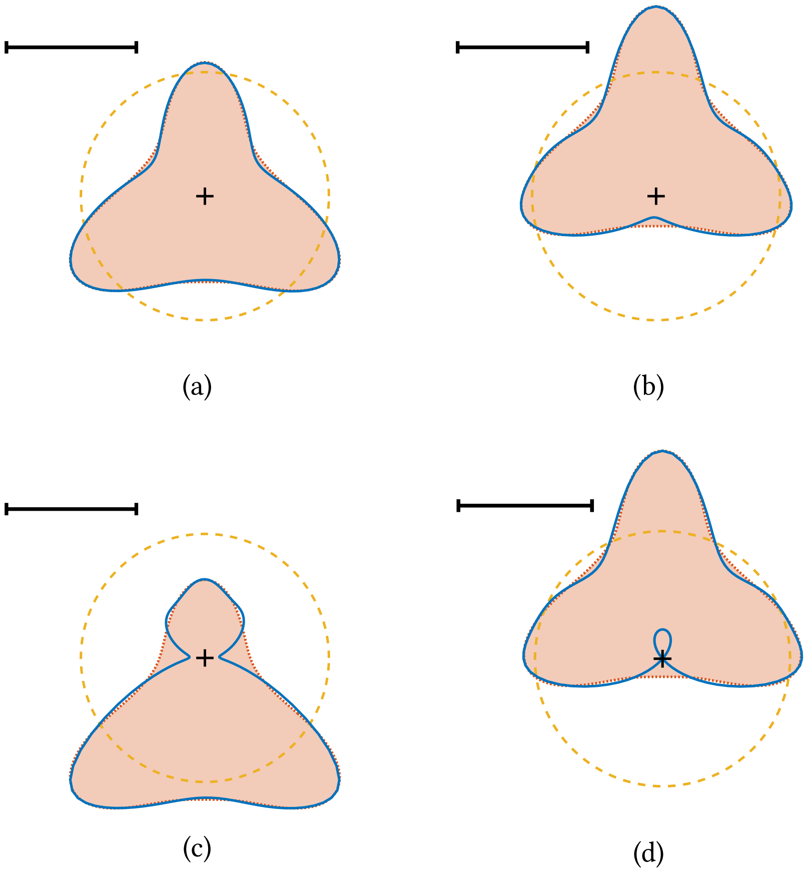


Figure 2.1: Reconstruction of a star-shape domain with respect to different origins with 1% Gaussian white noise. We use in all examples 8 equidistant incident waves, where the half of a wavelength is illustrated by the black plotting scale. Red dotted lines indicate the exact solution, blue solid line the reconstruction and yellow dashed lines initial guesses and the black cross highlights the origin of the radial functions. The radial functions are represented by their first 64 Fourier coefficients. For more details about the simulations itself we refer to the Section 6.2. The reconstruction in (a) is very good and in (b) it becomes worse at the boundary. But the examples in (c) contains unreasonable deformations and (d) is even a complete Failure as the computed curve contains a self-intersection.

imposed on $f^\dagger - f_0$ roughly correspond to a Sobolev smoothness. The index function (2.15) used in the proof of [25] is equivalent to

$$\varphi(\lambda) := \left(\log(3 + \lambda^{-1})\right)^{-p} \tag{2.37}$$

for some $p > 0$.

Nevertheless this whole approach for inverse obstacle scattering problems contains two major disadvantages. The first is the dependence on the parameterization. For one reconstruction process fix the origin or the point x_0 and then compute the unique parameterization given by (2.35). Such an approach necessitates a-priori knowledge: One needs to know an appropriate inner point of the obstacle to be able to state the mathematical model in the above sense. Note that the numerical reconstruction can fail if the origin is chosen badly. For an illustration see Figure 2.1. We can observe unwanted deformations in the reconstruction or even a failure if the center point is chosen too close to the boundary of the exact domain. This is expected since the penalty term corresponding to the exact solution explodes as the origin tends to the boundary.

Second it is obvious that the assumption on the domains to be star-shape is too restrictive. In Figure 2.2 we shown that one cannot compute the shape of a nonstar-shaped domain using this approach. The exclusion of many interesting classes of domains, which are not star-shaped limits the usage for practical purposes.

Therefore it is preferable to establish a new approach to model a space of admiss-

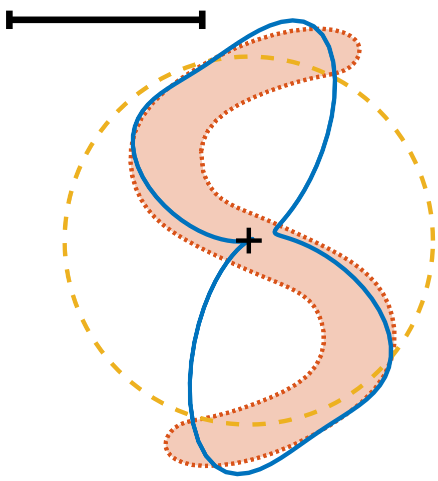


Figure 2.2: Failure of reconstructing a nonstar-shape domain by using radial functions. Parameters, line styles and colors are chosen as in Figure 2.1.

ible curves to overcome these disadvantages.

3 A class of shape manifolds

We become what we behold. We shape our tools and then our tools shape us.
— Marshall McLuhan, *Understanding media*

In this chapter we introduce a new shape space emerging from the geometrical description of curves in $\mathbb{R}^2$. As a preliminary result we introduce the construction of shape manifolds of closed curves. Moreover we give the definition of the attached tangent spaces including their Riemannian metric. Furthermore, the extrinsic curvature of the shape space is given explicitly. This is the second main result of this chapter, enlarging the investigation of shape manifolds in Section 3.1. It is known that from an extrinsic description of the curvature one can deduce the intrinsic curvature tensor. Nevertheless we prove also a bound on the curvature tensor without using a bound on the second fundamental form in Section 3.4.

In the following many notions and constructions from Riemannian geometry play a major role. We refer the reader for details to the Appendix A, where the needed concepts are recalled.

3.1 Construction of shape manifolds

The main parts of this section are published in the article [12].

Let $\Gamma \subset \mathbb{R}^2$ be a regular, closed curve of class H^2 of length L. That is, there is a parameterization $\gamma \in H^2([0,1];\mathbb{R}^2)$ satisfying $\gamma'(t) \neq 0$ for all $t \in [0,1]$ and the closing conditions

$$\gamma(0) = \gamma(1) \quad \text{and} \quad \gamma'(0) = \gamma'(1). \tag{3.1}$$

Without loss of generality, we assume that γ is of constant speed, i.e., $|\gamma'(t)| = L$ for all $t \in (0,1)$. This parameterization is well-known as *arc-length parameterization*. Thus, we represent γ by a triple $m = (\theta, L, p)$ with a base point $p := \gamma(0)$, curve's length L, and angle function $\theta \in H^1([0,1])$ via

$$\gamma(t) = \gamma_m(t) := p + \int_0^t \gamma'(\tau)\,\mathrm{d}\tau = p + L \int_0^t \big(\cos(\theta(\tau)), \sin(\theta(\tau))\big)\,\mathrm{d}\tau. \tag{3.2}$$

In order to meet the closing conditions (3.1), θ needs to satisfy

$$\int_0^1 \cos(\theta(t))\,\mathrm{d}t = 0, \quad \int_0^1 \sin(\theta(t))\,\mathrm{d}t = 0 \quad \text{and} \quad \theta(1) - \theta(0) \in 2\,\pi\,\mathbb{Z}. \tag{3.3}$$

The number $\frac{\theta(1)-\theta(0)}{2\,\pi}$ is called the *turning number* of γ (not to be confused with the winding number). The map γ is an *embedding* of Γ if it is a diffeomorphism onto Γ and a necessary (but not sufficient) condition for Γ to be embedded is that γ has turning number ± 1, see [11, Sec. 5.7, Thm. 2]. Since our application focuses on boundary curves of simply connected domains, we restrict the curves turning number to be +1 and define the space:

$$\Theta := \left\{\theta \in H^1([0,1]) \,\middle|\, \textstyle\int_0^1 \big(\cos(\theta(t)), \sin(\theta(t))\big)\,\mathrm{d}t = 0,\ \theta(1) - \theta(0) = 2\,\pi\right\}. \tag{3.4}$$

The nonlinear equations in the definition of Θ prevent any kind of linear space structure. Nevertheless, Θ is a manifold as we show in the next theorem.

Theorem 3.1. *The space Θ is an embedded submanifold of $H^1([0,1])$.*

Proof. First we define a constraint mapping $\boldsymbol{\Phi}\colon H^1([0,1]) \to \mathbb{R}^2 \times \mathbb{R}$ given by

$$\boldsymbol{\Phi}(\theta) := (\Phi_{\mathrm{cld}}(\theta), \Phi_{\mathrm{per}}(\theta)) \tag{3.5a}$$

where

$$\Phi_{\mathrm{cld}}\colon H^1([0,1]) \to \mathbb{R}^2, \qquad \theta \mapsto \begin{pmatrix} \int_0^1 \cos(\theta(s))\,\mathrm{d}s \\ \int_0^1 \sin(\theta(s))\,\mathrm{d}s \end{pmatrix} \tag{3.5b}$$

$$\Phi_{\mathrm{per}}\colon H^1([0,1]) \to \mathbb{R}, \qquad \theta \mapsto \theta(1) - \theta(0) - 2\pi. \tag{3.5c}$$

Using this notation we rewrite Θ as follows

$$\Theta = \left\{ \theta \in H^1([0,1]) \;\middle|\; \boldsymbol{\Phi}(\theta) = 0 \right\}.$$

By the implicit function theorem we have to show that $\boldsymbol{\Phi}$ is a submersion to show that Θ is a submanifold. That is, $\mathrm{D}\boldsymbol{\Phi}(\theta)$ admits a bounded linear right inverse for each $\theta \in \boldsymbol{\Phi}^{-1}(0)$ (see section A.1 or [36, II.2]).

First, we calculate derivatives of the functions Φ_{cld} and Φ_{per} used in the definition of $\boldsymbol{\Phi}$. Let $\theta \in H^1([0,1])$ be given. Then for $u \in H^1([0,1])$ we deduce

$$\mathrm{D}\Phi_{\mathrm{cld}}(\theta)\, u = \begin{pmatrix} -\int_0^1 \sin(\theta(t))\, u(t)\, \mathrm{d}t \\ \int_0^1 \cos(\theta(t))\, u(t)\, \mathrm{d}t \end{pmatrix} = \begin{pmatrix} -\langle s_\theta, u \rangle_{L^2} \\ \langle c_\theta, u \rangle_{L^2} \end{pmatrix}$$

and

$$\mathrm{D}\Phi_{\mathrm{per}}(\theta)\, u = u(1) - u(0).$$

Here we used the notation

$$s_\theta(t) := \sin(\theta(t)) \qquad \text{and} \qquad c_\theta(t) := \cos(\theta(t)) \qquad \text{for } t \in [0,1]. \tag{3.6}$$

To prove the subjectivity of $\mathrm{D}\boldsymbol{\Phi}(\theta)$ for $\theta \in \Theta$ it suffices to construct a function $u \in H^1([0,1])$ such that it solves $\mathrm{D}\boldsymbol{\Phi}(\theta)\, u = \lambda$ and depends linearly on $\lambda \in \mathbb{R}^3$. For this define

$$u(t) := a s_\theta(t) + b c_\theta(t) + ct.$$

For $\theta \in \boldsymbol{\Phi}^{-1}(0)$ the function u is a pre-image of λ, if the coefficients a, b and c solve the linear system

$$\mathrm{D}\boldsymbol{\Phi}(\theta)\, u = \begin{pmatrix} -\langle s_\theta, s_\theta \rangle_{L^2} & -\langle s_\theta, c_\theta \rangle_{L^2} & -\langle s_\theta, t \rangle_{L^2} \\ \langle c_\theta, s_\theta \rangle_{L^2} & \langle c_\theta, c_\theta \rangle_{L^2} & \langle c_\theta, t \rangle_{L^2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

The determinant D of the matrix given by $D = -\|s_\theta\|_{L^2}^2 \|c_\theta\|_{L^2}^2 + |\langle s_\theta, c_\theta \rangle_{L^2}|^2$ is negative by the Cauchy-Schwarz inequality. It attains zero if and only if s_θ and c_θ are linearly dependent. This is a contradiction since $\theta \in \boldsymbol{\Phi}^{-1}(0)$ is continuous, but never constant. The result follows because this proves surjectivity of $\mathrm{D}\boldsymbol{\Phi}(\theta)$. □

As pointed out in the Appendix A the tangent space of a manifold is essential for differential geometry. The tangent space $\mathcal{T}_\theta \Theta$ of Θ is given by [1]

[1] In this thesis we use calligraphic letters as $\mathcal{M}$, $\mathcal{T}$, $\mathcal{E}$ or $\mathcal{exp}$ to denote sets, functions or constructions arising from differential geometry to incorporate well-established notations from the two fields of Riemannian geometry and regularization theory.

$$\mathcal{T}_\theta \Theta = \left\{ u \in H^1([0,1]) \,\middle|\, \mathrm{D}\boldsymbol{\Phi}(\theta)\, u = 0 \right\}.$$

The family of inner products $(g_\theta)_{\theta\in\Theta}$ defined by

$$g_\theta(u,v) := \int_0^1 \big(u(t)\, v(t) + u'(t)\, v'(t)\big)\, \mathrm{d}t \quad \text{for } u, v \in \mathcal{T}_\theta\Theta$$

turns (Θ, g) into a infinite-dimensional Riemannian manifold (in the sense of [36], see Section A.1 for details).

Remark 3.2. From the proof of Theorem 3.1 and the construction of the constraint mapping $\boldsymbol{\Phi}$ one can see that the structure of Θ is almost only influenced by the closure constraint Φ_{cld}. Indeed we can describe the manifold structure also slightly differently. Note that

$$\Phi_{\mathrm{per}}^{-1}(0) = \left\{ \theta \in H^1([0,1]) \,\middle|\, \theta(1) - \theta(0) = 2\pi \right\}$$

is an affine subspace in $H^1([0,1])$. Thus, it suffices to consider the constraint mapping

$$\Phi\colon H^1_{\mathrm{per}}([0,1]) \to \mathbb{R}^2, \qquad \theta \mapsto \begin{pmatrix} \int_0^1 c_\theta(t)\,\mathrm{d}t \\ \int_0^1 s_\theta(t)\,\mathrm{d}t \end{pmatrix} \tag{3.7}$$

with

$$H^1_{\mathrm{per}}([0,1]) := \left\{ u \in H^1([0,1]) \,\middle|\, u(0) = u(1) \right\}.$$

Especially for the tangent space the equivalent description is useful

$$\mathcal{T}_\theta\Theta = \left\{ u \in H^1_{\mathrm{per}}([0,1]) \,\middle|\, \mathrm{D}\Phi(\theta)\, u = 0 \right\}. \tag{3.8}$$

For a compact, convex set of base points $B \subset \mathbb{R}^2$ and for bounds of acceptable curve lengths L_1, L_2, we define the space of feasible curves by

$$\mathcal{M} := \Theta \times [L_1, L_2] \times B.$$

Then $\mathcal{M}$ is a smooth submanifold with corners in the Hilbert space

$$\mathbb{X} := H^1([0,1]) \times \mathbb{R} \times \mathbb{R}^2$$

and its tangent space at an interior point $m = (\theta, L, p)$ is given by

$$\mathcal{T}_m\mathcal{M} = \mathcal{T}_\theta\Theta \oplus \mathbb{R} \oplus \mathbb{R}^2.$$

3.2 The second fundamental form

The following sections are dedicated to the investigation of the constraint function $\Phi\colon H^1_{\mathrm{per}}([0,1]) \to \mathbb{R}^2$ defined in (3.7). The main result of this section is the proof of an explicit formula for the second fundamental form of the submanifold Θ in $H^1([0,1])$. By the proof of Theorem 3.1 and Remark 3.2 Θ is given by the submersion Φ. It is a well-known fact from Riemannian geometry that in this case the second fundamental form $\mathbb{I}^{\Theta}$ is given by the formula (A.26) in terms of derivatives and a Moore-Penrose inverse.

The next lemma shows the smoothness of the constraint operator Φ.

Lemma 3.3. *The map $\Phi\colon H^1_{\mathrm{per}}([0,1]) \to \mathbb{R}^2$ is a C^∞ operator and for $n \in \mathbb{N}_0$ the derivative of an odd order*

$$\mathrm{D}^{2n+1}\Phi(\theta)\colon \left(H^1_{\mathrm{per}}([0,1])\right)^{2n+1} \to \mathbb{R}^2$$

is given by

$$\mathrm{D}^{2n+1}\Phi(\theta)\,(u_1,\dots,u_{2n+1}) = \begin{pmatrix} (-1)^{n+1}\int_0^1 s_\theta(t)u_1(t)\cdots u_{2n+1}(t)\,\mathrm{d}t \\ (-1)^n \int_0^1 c_\theta(t)u_1(t)\cdots u_{2n+1}(t)\,\mathrm{d}t \end{pmatrix}.$$

In the even case, the derivative

$$\mathrm{D}^{2n}\Phi(\theta)\colon \left(H^1_{\mathrm{per}}([0,1])\right)^{2n} \to \mathbb{R}^2$$

is given by

$$\mathrm{D}^{2n}\Phi(\theta)\,(u_1,\dots,u_{2n}) = (-1)^n \begin{pmatrix} \int_0^1 c_\theta(t)u_1(t)\cdots u_{2n+1}(t)\,\mathrm{d}t \\ \int_0^1 s_\theta(t)u_1(t)\cdots u_{2n+1}(t)\,\mathrm{d}t \end{pmatrix}.$$

Furthermore,

$$|\mathrm{D}^k\Phi(\theta)\,(u_1,\dots,u_k)|_{\mathbb{R}^2} \le \|u_1\|_{H^1}\cdots\|u_k\|_{H^1} \tag{3.9}$$

for all $u_1,\dots,u_k \in H^1_{\mathrm{per}}([0,1])$ and all $k \in \mathbb{N}$. [2]

Proof. The stated formulas are proven by a straight forward calculation. We prove (3.9) for k odd. The even case can be proven analogously. Let $n \in \mathbb{N}_0$, then it follows

$$|\mathrm{D}^{2n+1}\Phi(\theta)\,(u_1,\dots,u_{2n+1})|^2_{\mathbb{R}^2} = \left|\int_0^1 s_\theta(t)u_1(t)\cdots u_{2n+1}(t)\,\mathrm{d}t\right|^2$$

[2]Note that in this thesis norms and inner product in Hilbert spaces $\mathbb{X}$ are denoted by $\|\cdot\|_{\mathbb{X}}$ and $\langle\cdot,\cdot\rangle_{\mathbb{X}}$, where we use the notation $|\cdot|_{\mathbb{R}^n}$ and $(\cdot,\cdot)_{\mathbb{R}^n}$ for norms and inner products in $\mathbb{R}^n$.

$$
\begin{aligned}
&+ \left| \int_0^1 c_\theta(t) u_1(t) \cdots u_{2n+1}(t) \, \mathrm{d}t \right|^2 \\
&\leq \int_0^1 (s_\theta^2(t) + c_\theta^2(t)) \left(u_1(t) \cdots u_{2n+1}(t) \right)^2 \mathrm{d}t \\
&\leq \|u_1\|_{L^2}^2 \cdots \|u_{2n+1}\|_{L^2}^2
\end{aligned}
$$

and by taking the square root on both side one can deduce $\|\mathrm{D}^k \Phi(\theta)\| \leq 1$ for all $k \in \mathbb{N}_0$. □

Next, we prove an expression for the Moore-Penrose inverse of the derivative $\mathrm{D}\Phi$.

Lemma 3.4. *The Moore-Penrose inverse* $\mathrm{D}\Phi(\theta)^\dagger \colon \mathbb{R}^2 \longrightarrow H^1_{\mathrm{per}}([0,1])$ *for* $\theta \in \Theta$ *is given by*

$$
\mathrm{D}\Phi(\theta)^\dagger (x) = \frac{-\left(\|g_\theta\|_{H^1}^2 x_1 + \langle f_\theta, g_\theta \rangle_{H^1} x_2 \right)}{D_\theta} f_\theta + \frac{\langle g_\theta, f_\theta \rangle_{H^1} x_1 + \|f_\theta\|_{H^1} x_2}{D_\theta} g_\theta \tag{3.10}
$$

for $x = (x_1, x_2)^\top \in \mathbb{R}^2$*, where*

$$
D_\theta := \|f_\theta\|_{H^1}^2 \|g_\theta\|_{H^1}^2 - |\langle f_\theta, g_\theta \rangle|^2 \tag{3.11}
$$

with $f_\theta, g_\theta \in H^2_{\mathrm{per}}([0,1]) = \{ f \in H^1([0,1]) \mid f(0) = f(1),\ f'(0) = f'(1) \}$ *solving the ordinary differential equations*

$$
f_\theta - f_\theta'' = s_\theta \tag{3.12a}
$$

and

$$
g_\theta - g_\theta'' = c_\theta. \tag{3.12b}
$$

Proof. In the following we need the existence of the functions $f_\theta, g_\theta \in H^2_{\mathrm{per}}([0,1])$ as solutions of the ordinary differential equations (3.12). At the end of the proof we give the explicit formulas for them, but first we construct the generalized inverse. We compute the adjoint operator $\mathrm{D}\Phi(\theta)^* \colon \mathbb{R}^2 \longrightarrow H^1_{\mathrm{per}}([0,1])$ satisfying the equation

$$
(\mathrm{D}\Phi(\theta)\, u, x)_{\mathbb{R}^2} = \langle u, \mathrm{D}\Phi(\theta)^*\, x \rangle_{H^1}
$$

for all $u \in H^1_{\mathrm{per}}([0,1])$ and $x \in \mathbb{R}^2$. Note that by (3.12) and integration by parts for $u \in H^1_{\mathrm{per}}([0,1])$ it follows that

$$
\begin{aligned}
\langle f_\theta, u \rangle_{H^1} = \langle f_\theta, u \rangle_{L^2} + \langle f_\theta', u' \rangle_{L^2} &= \langle f_\theta - f_\theta'', u \rangle_{L^2} \\
&= \langle s_\theta, u \rangle_{L^2}
\end{aligned}
$$

and analogously $\langle g_\theta, u\rangle_{H^1} = \langle c_\theta, u\rangle_{L^2}$. Therefore with $x = (x_1, x_2)^\top \in \mathbb{R}^2$ we get

$$\begin{aligned}\langle u, \mathrm{D}\Phi(\theta)^* x\rangle_{H^1} = (\mathrm{D}\Phi(\theta)\, u, x)_{\mathbb{R}^2} &= -\langle s_\theta, u\rangle_{L^2}\, x_1 + \langle c_\theta, u\rangle_{L^2}\, x_2 \\ &= -\langle f_\theta, u\rangle_{H^1}\, x_1 + \langle g_\theta, u\rangle_{H^1}\, x_2 \\ &= \langle u, -f_\theta\, x_1 + g_\theta\, x_2\rangle_{H^1}\end{aligned}$$

and thus

$$\mathrm{D}\Phi(\theta)^* \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -f_\theta\, x_1 + g_\theta\, x_2. \tag{3.13}$$

Since $\mathrm{D}\Phi(\theta)$ is surjective as seen in Remark 3.2, we can compute the Moore-Penrose inverse via

$$\mathrm{D}\Phi(\theta)^\dagger = \mathrm{D}\Phi(\theta)^* \left(\mathrm{D}\Phi(\theta)\, \mathrm{D}\Phi(\theta)^*\right)^{-1}. \tag{3.14}$$

This formula is true for all linear surjective operators and can be found for example in [45]. $\mathrm{D}\Phi(\theta)\, \mathrm{D}\Phi(\theta)^* \colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by the matrix expression

$$\begin{aligned}\mathrm{D}\Phi(\theta)\, \mathrm{D}\Phi(\theta)^* \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} \langle s_\theta, f_\theta\rangle_{L^2} & -\langle s_\theta, g_\theta\rangle_{L^2} \\ -\langle c_\theta, f_\theta\rangle_{L^2} & \langle c_\theta, g_\theta\rangle_{L^2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} \langle f_\theta, f_\theta\rangle_{H^1} & -\langle f_\theta, g_\theta\rangle_{H^1} \\ -\langle g_\theta, f_\theta\rangle_{H^1} & \langle g_\theta, g_\theta\rangle_{H^1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\end{aligned}$$

with the identities as above. The determinant D_θ of this matrix is given in (3.11). Note that $D_\theta \geq 0$ by the Cauchy-Schwarz inequality. To prove that it never attains zero assume the contrary. Then f_θ and g_θ have to be linearly dependent, i.e. there is a $\lambda \in \mathbb{R} \setminus \{0\}$ such that $f_\theta = \lambda\, g_\theta$. Plugging this in the equations (3.12), we deduce linear dependence between s_θ and c_θ, which cannot be the case, since $\theta \in \Theta$ is continuous and not constant. Therefore, $D_\theta > 0$ and the inverse matrix is given by

$$\left(\mathrm{D}\Phi(\theta)\, \mathrm{D}\Phi(\theta)^*\right)^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{D_\theta} \begin{pmatrix} \langle g_\theta, g_\theta\rangle_{H^1} & \langle f_\theta, g_\theta\rangle_{H^1} \\ \langle g_\theta, f_\theta\rangle_{H^1} & \langle f_\theta, f_\theta\rangle_{H^1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

By combining this matrix with the adjoint operator we obtain the stated formula for the generalized inverse.

For the sake of completeness we give here explicit formulas for f_θ and g_θ, which solves the equations (3.12). We consider the function

$$f_\theta(t) := a_{\mathrm{s},\theta}(t)\mathrm{e}^t + b_{\mathrm{s},\theta}(t)\mathrm{e}^{-t} \tag{3.15a}$$

with

$$a_{\mathrm{s},\theta}(t) := \tilde{a}_{\mathrm{s},\theta}(t) + \frac{e}{1-e}\tilde{a}_{\mathrm{s},\theta}(1) \qquad \text{and} \qquad b_{\mathrm{s},\theta}(t) := \tilde{b}_{\mathrm{s},\theta}(t) + \frac{e^{-1}}{1-e^{-1}}\tilde{b}_{\mathrm{s},\theta}(1) \tag{3.15b}$$

where

$$\tilde{a}_{s,\theta}(t) := -\frac{1}{2}\int_0^t e^{-\xi} s_\theta(\xi)\,d\xi \qquad \text{and} \qquad \tilde{b}_{s,\theta}(t) := \frac{1}{2}\int_0^t e^{\xi} s_\theta(\xi)\,d\xi. \tag{3.15c}$$

We verify that this function indeed solves the equation (3.12a). Note that $a'_{s,\theta}(t) = \frac{1}{2}s_\theta(t)e^{-t}$ and $b'_{s,\theta}(t) = -\frac{1}{2}s_\theta(t)e^{t}$. Using this we deduce for the first derivative

$$f'_\theta(t) = a'_{s,\theta}(t)e^t + a_{s,\theta}(t)e^t + b'_{s,\theta}(t)e^{-t} - b_{s,\theta}(t)e^{-t} = a_{s,\theta}(t)e^t - b_{s,\theta}(t)e^{-t}$$

and for the second derivative

$$f''_\theta(t) = a'_{s,\theta}(t)e^t + a_{s,\theta}(t)e^t - b'_{s,\theta}(t)e^{-t} + b_{s,\theta}(t)e^{-t} = f_\theta(t) - s_\theta(t),$$

which shows that f_θ is a solution of the ordinary differential equation (3.12a).

To verify that this function also satisfies the periodic boundary condition we compute directly

$$\begin{aligned} f_\theta(1) = a_{s,\theta}(1)e + b_{s,\theta}(0)e^{-1} &= \Big(1 + \frac{e}{1-e}\Big)\tilde{a}_{s,\theta}(1)e + \Big(1 + \frac{e^{-1}}{1-e^{-1}}\Big)\tilde{b}_{s,\theta}(1)e^{-1} \\ &= \frac{e}{1-e}\tilde{a}_{s,\theta}(1) + \frac{e^{-1}}{1-e^{-1}}\tilde{b}_{s,\theta}(1) \\ &= a_{s,\theta}(0) + b_{s,\theta}(0) = f_\theta(0). \end{aligned}$$

Thus f_θ solves equation (3.12a) and satisfies the boundary conditions. Completely analogously it holds that

$$g_\theta(t) := a_{c,\theta}(t)e^t + b_{c,\theta}(t)e^{-t} \tag{3.16a}$$

with

$$a_{c,\theta}(t) := \tilde{a}_{c,\theta}(t) + \frac{e}{1-e}\tilde{a}_{c,\theta}(1) \qquad \text{and} \qquad \tilde{a}_{c,\theta}(t) := -\frac{1}{2}\int_0^t e^{-\xi} c_\theta(\xi)\,d\xi, \tag{3.16b}$$

and

$$b_{c,\theta}(t) := \tilde{b}_{c,\theta}(t) + \frac{e^{-1}}{1-e^{-1}}\tilde{b}_{c,\theta}(1) \qquad \text{and} \qquad \tilde{b}_{c,\theta}(t) := \frac{1}{2}\int_0^t e^{\xi} c_\theta(\xi)\,d\xi. \tag{3.16c}$$

By an analog argument one verifies that g_θ solves (3.12b). □

In order to derive a concrete bound of the Moore-Penrose inverse $D\Phi(\theta)^\dagger$, which will play a major role in later computations, one needs to bound D_θ from below. This is done in the following lemma.

Lemma 3.5. *For $\theta \in \Theta$ and $f_\theta, g_\theta \in H^2_{\text{per}}([0,1])$ defined in Lemma 3.4 it holds the lower bound*

$$D_\theta = \|f_\theta\|^2_{H^1}\|g_\theta\|^2_{H^1} - |\langle f_\theta, g_\theta\rangle_{H^1}|^2 \geq \frac{1}{C_D(\theta)} \tag{3.17}$$

with

$$C_D(\theta) := \frac{16}{\pi^2}\|\theta'\|^2_{L^2}\left(1+\|\theta'\|_{L^2}\right)^4\left(1+\frac{2}{\sqrt{2}}\left(1+\|\theta'\|_{L^2}\right)\right)^2. \tag{3.18}$$

Proof. From (3.12a) we deduce by integration by parts

$$\|s_\theta\|^2_{L^2} = \|f_\theta\|^2_{L^2} + 2\|f'_\theta\|^2_{L^2} + \|f''_\theta\|^2_{L^2} = \|f_\theta\|^2_{H^1} + \|f'_\theta\|^2_{H^1}. \tag{3.19}$$

By leaving out the last term on the right-hand side implies immediately the upper bound

$$\|f_\theta\|_{H^1} \leq \|s_\theta\|_{L^2} \leq 1. \tag{3.20}$$

Furthermore, by using again once more integration by parts and Cauchy-Schwarz inequality we arrive at

$$\begin{aligned}\|f'_\theta\|^2_{H^1} &= \langle f'_\theta, f'_\theta\rangle_{L^2} + \langle f''_\theta, f''_\theta\rangle_{L^2} \\ &= \langle f''_\theta, -f_\theta + f''_\theta\rangle_{L^2} = -\langle f''_\theta, s_\theta\rangle_{L^2} = \langle f'_\theta, \theta' c_\theta\rangle_{L^2} \leq \|f'_\theta\|_{L^2}\|\theta' c_\theta\|_{L^2}.\end{aligned} \tag{3.21}$$

Thus, we deduce

$$\|s_\theta\|^2_{L^2} \leq \|f_\theta\|^2_{H^1} + \|f'_\theta\|_{L^2}\|\theta'\|_{L^2} \leq \left(1+\|\theta'\|_{L^2}\right)\|f_\theta\|_{H^1}.$$

By (3.12) c_θ and g_θ are connected by an analogous equation than s_θ and f_θ. This implies that an analogous result holds true by exchanging s_θ with c_θ and f_θ with g_θ. Because of the identity

$$\|s_\theta\|^2_{L^2} + \|c_\theta\|^2_{L^2} = \int_0^1 \left(\sin(\theta(t))\right)^2 + \left(\cos(\theta(t))\right)^2 \,\mathrm{d}t = 1 \tag{3.22}$$

one of the summands on the left-hand side of (3.22) has to be larger than $\frac{1}{2}$. Thus, we assume without loss of generality that $\|s_\theta\|^2_{L^2} \geq \frac{1}{2}$. Otherwise we change in the following the roles of c_θ and s_θ respectively g_θ and f_θ. Therefore we get the lower estimate

$$\|f_\theta\|_{H^1} \geq \frac{1}{2\left(1+\|\theta'\|_{L^2}\right)}. \tag{3.23}$$

On the other hand by (3.20) and (3.22) it holds that $\|g_\theta\|^2_{H^1} \leq \|c_\theta\|^2_{L^2} \leq \frac{1}{2}$. Next, we define

$$\lambda := \frac{\langle f_\theta, g_\theta\rangle_{H^1}}{\|f_\theta\|^2_{H^1}}. \tag{3.24}$$

Then by straightforward computations and (3.23) we can estimate

$$\begin{aligned}
\|g_\theta - \lambda f_\theta\|_{H^1}^2 &= \|g_\theta\|_{H^1}^2 - 2\lambda\langle g_\theta, f_\theta\rangle_{H^1} + \lambda^2\|f_\theta\|_{H^1}^2 \\
&= \|g_\theta\|_{H^1}^2 - 2\frac{\left|\langle f_\theta, g_\theta\rangle_{H^1}\right|^2}{\|f_\theta\|_{H^1}^2} + \frac{\left|\langle f_\theta, g_\theta\rangle_{H^1}\right|^2}{\|f_\theta\|_{H^1}^2} \\
&= \frac{\|g_\theta\|_{H^1}^2\|f_\theta\|_{H^1}^2 - \left|\langle f_\theta, g_\theta\rangle_{H^1}\right|^2}{\|f_\theta\|_{H^1}^2} = \frac{D_\theta}{\|f_\theta\|_{H^1}^2} \leq 4\big(1 + \|\theta'\|_{L^2}\big)^2 D_\theta.
\end{aligned}$$

To continue with a lower bound of $\|g_\theta - \lambda f_\theta\|_{H^1}$ we use the same strategy as in (3.19) to get

$$\|c_\theta - \lambda s_\theta\|_{L^2}^2 = \|g_\theta - \lambda f_\theta\|_{H^1}^2 + \|g_\theta' - \lambda f_\theta'\|_{H^1}^2.$$

In analogy to (3.21)

$$\|g_\theta' - \lambda f_\theta'\|_{H^1}^2 = \langle g_\theta' - \lambda f_\theta', \theta'\big(- s_\theta - \lambda c_\theta\big)\rangle_{L^2} \leq \|g_\theta - \lambda f_\theta\|_{H^1}\,\|\theta'\|_{L^2}\,\big(1 + \lambda\big).$$

Further, it follows from (3.23) and $\|g\|_{H^1}^2 \leq \frac{1}{2}$ that

$$\lambda = \frac{\langle f_\theta, g_\theta\rangle_{H^1}}{\|f_\theta\|_{H^1}^2} \leq \frac{\|g_\theta\|_{H^1}}{\|f_\theta\|_{H^1}} \leq \frac{2}{\sqrt{2}}\big(1 + \|\theta'\|_{L^2}\big). \tag{3.25}$$

The bound (3.20) implies $\|g_\theta - \lambda f_\theta\|_{H^1} \leq 1 + \lambda$ and therefore

$$\begin{aligned}
\|c_\theta - \lambda s_\theta\|_{L^2}^2 &\leq \big(1 + \|\theta'\|_{L^2}\big)\,\big(1 + \lambda\big)\,\|g_\theta - \lambda f_\theta\|_{H^1} \\
&\leq \big(1 + \|\theta'\|_{L^2}\big)\,\Big(1 + \frac{2}{\sqrt{2}}\,\big(1 + \|\theta'\|_{L^2}\big)\Big)\,\|g_\theta - \lambda f_\theta\|_{H^1}.
\end{aligned}$$

Next, we continue the estimate from below and derive a lower bound on $\|c_\theta - \lambda s_\theta\|_{L^2}$. Thus, define $\varphi \in (0, \pi)$ such that

$$\cot\varphi = \lambda. \tag{3.26}$$

Denote the bound on the right-hand side in (3.25) with C_θ. Since $0 < |\lambda| \leq C_\theta < \infty$ it is $|\varphi| \leq \operatorname{arccot}(C_\theta) < \infty$. Further by $|\sin\varphi| \leq 1$ it is $\frac{1}{|\sin\varphi|} \geq 1$. By applying (3.26) and an angle addition theorem for the sine one obtains the estimate

$$\begin{aligned}
\|c_\theta - \lambda s_\theta\|_{L^2}^2 &= \int_0^1 \big|\cos\theta(t) - \cot\varphi\,\sin\theta(t)\big|^2\,\mathrm{d}t \\
&= \frac{1}{|\sin\varphi|^2}\int_0^1 \big|\cos\theta(t)\,\sin\varphi - \cos\varphi\,\sin\theta(t)\big|^2\,\mathrm{d}t
\end{aligned}$$

$$\geq \int_0^1 \left|\cos\theta(t)\ \sin\varphi - \cos\varphi\ \sin\theta(t)\right|^2 \mathrm{d}t$$
$$= \int_0^1 \left|\sin\left(\theta(t) - \varphi\right)\right|^2 \mathrm{d}t$$

Next we introduce a shifted version $\tilde{\theta}\colon [0,1] \to [\varphi, \varphi + 2\pi)$ of θ such that $\tilde{\theta}(t) - \theta(t) \in 2\pi\mathbb{Z}$ for all $t \in [0,1]$. Since $\theta \in \Theta$, i.e. $\theta \in H^1([0,1])$ and $\theta(1) - \theta(0) = 2\pi$ there exist $t_0 < t_1$ such that

$$\tilde{\theta}(t) \in \left[\varphi + \frac{\pi}{4}, \varphi + \frac{3\pi}{4}\right] \qquad \text{for all } t \in [t_0, t_1]$$

and $\tilde{\theta}(t_0) = \varphi + \frac{\pi}{4}$ and $\tilde{\theta}(t_1) = \varphi + \frac{3\pi}{4}$. It holds that $\tilde{\theta}'(t) = \theta'(t)$ for all $t \in [0,1]$ and by fundamental theorem of calculus and Cauchy-Schwarz

$$\frac{\pi}{2} = \left|\tilde{\theta}(t_1) - \tilde{\theta}(t_0)\right| = \left|\int_{t_0}^{t_1} 1 \cdot \tilde{\theta}'(t)\,\mathrm{d}t\right| \leq \sqrt{t_1 - t_0}\left(\int_{t_0}^{t_1} |\tilde{\theta}'(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}}$$
$$\leq \sqrt{t_1 - t_0}\|\theta'\|_{L^2}.$$

By the choice of the interval one can bound from below $|\sin(\varphi - \tilde{\theta}(t))| \geq \frac{1}{\sqrt{2}}$ for all $t \in [t_0, t_1]$ and we can use the above to estimate

$$\int_{t_0}^{t_1} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t \geq \frac{t_1 - t_0}{2} \geq \frac{\pi^2}{8\|\theta'\|_{L^2}^2}.$$

Further there exists also $t_2 < t_3$ such that

$$\tilde{\theta}(t) \in \left[\varphi + \frac{5\pi}{4}, \varphi + \frac{7\pi}{4}\right] \qquad \text{for all } t \in [t_2, t_3]$$

and now analogously as above it holds that $t_3 - t_2 \geq \pi^2/(4\|\theta'\|_{L^2}^2)$ and also the corresponding lower estimate for the integral over the interval $[t_2, t_3]$. Now the combination of the last arguments yields

$$\|c_\theta - \lambda s_\theta\|_{L^2}^2 \geq \int_{t_0}^{t_1} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t + \int_{t_2}^{t_3} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t \geq \frac{\pi^2}{4\|\theta'\|_{L^2}^2}.$$

By putting the above estimates together we obtain the assertion. □

Using the last lemma one proves a bound of the operator norm of the Moore-Penrose inverse.

Lemma 3.6. *The map* $\mathrm{D}\Phi(\theta)^\dagger\colon \mathbb{R}^2 \to H^1([0,1])$ *is bounded in terms of* θ*, i.e. it holds*

$$\|\mathrm{D}\Phi(\theta)^\dagger\,(x)\|_{H^1} \leq 2\sqrt{2}\,C_D(\theta)\,|x|_{\mathbb{R}^2} \tag{3.27}$$

for all $x = (x_1, x_2)^\top \in \mathbb{R}^2$*, where* $C_D(\theta)$ *is defined in (3.18).*

Proof. Plugging in the formula (3.10), we obtain by straightforward manipulations and $\|f_\theta\|_{H^1}, \|g_\theta\|_{H^1} \leq 1$ that

$$\begin{aligned}\|\mathrm{D}\Phi(\theta)^\dagger(x)\|_{H^1} &\leq \frac{1}{D_\theta}\Big(\big|\|g_\theta\|_{H^1}^2 x_1 + \langle f_\theta, g_\theta\rangle_{H^1} x_2\big| \|f_\theta\|_{H^1} \\ &\qquad + \big|\langle g_\theta, f_\theta\rangle_{H^1} x_1 + \|f_\theta\|_{H^1}^2 x_2\big| \|g_\theta\|_{H^1}\Big) \\ &\leq \frac{2}{D_\theta}\big(|x_1| + |x_2|\big).\end{aligned}$$

Using that $|x_1| + |x_2| \leq \sqrt{2}\,|x|_{\mathbb{R}^2}$ together with Lemma 3.5, the statement follows. □

For the description of the extrinsic curvature of a submanifold the notion of the second fundamental form is essential. We apply the formulas from Section A.5 and and results above to get the following theorem.

Theorem 3.7. *The second fundamental form* $\mathrm{I\!I}_\theta^\Theta$ *at* θ *of the submanifold* Θ *in* $H^1([0,1])$ *is given by*

$$\begin{aligned}\mathrm{I\!I}_\theta^\Theta(u,v) = &-\frac{\|g_\theta\|_{H^1}^2 \langle c_\theta, u\,v\rangle_{L^2} + \langle f_\theta, g_\theta\rangle_{H^1}\langle s_\theta, u\,v\rangle_{L^2}}{D_\theta} f_\theta \\ &+ \frac{\langle g_\theta, f_\theta\rangle_{H^1}\langle c_\theta, u\,v\rangle_{L^2} + \|f_\theta\|_{H^1}^2\langle s_\theta, u\,v\rangle_{L^2}}{D_\theta} g_\theta.\end{aligned} \tag{3.28}$$

and it holds that

$$\|\mathrm{I\!I}_\theta^\Theta(u,v)\|_{H^1} \leq C_{\mathrm{I\!I}}(\theta)\,\|u\|_{H^1}\|v\|_{H^1} \tag{3.29}$$

with

$$C_{\mathrm{I\!I}}(\theta) := 2\sqrt{2}\,C_D(\theta), \tag{3.30}$$

for all $u, v \in \mathcal{T}_\theta\Theta$*, where* $C_D(\theta)$ *is defined in (3.18).*

Proof. Since Θ is an embedded submanifold of $H^1([0,1])$ we can apply the characterization (A.26) of the second fundamental form by the submersion Φ. This gives

$$\mathrm{I\!I}_\theta^\Theta(u,v) = -\mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}^2\Phi(\theta)(u,v). \tag{3.31}$$

From Lemma 3.3 we know that

$$\mathrm{D}^2\Phi(\theta)(u,v) = -\begin{pmatrix}\langle c_\theta, uv\rangle_{L^2} \\ \langle s_\theta, uv\rangle_{L^2}\end{pmatrix}.$$

Substituting this into the formula (3.10) yields the statement (3.28).

The inequality (3.29) follows by applying the bounds in the Lemmas 3.3 and 3.6 to (3.31), which finishes the proof. □

As an immediately consequence the next corollary provides a local bound on the curvature tensor.

Corollary 3.8. *Let $\theta \in \Theta$ and $u, v \in \mathcal{T}_\theta\Theta$. Then it holds that*

$$|\langle \mathcal{R}_\theta(u,v)u, v\rangle_{H^1}| \leq 2\left(C_{\mathbb{I}}(\theta)\right)^2 \|u\|_{H^1}^2 \|v\|_{H^1}^2. \tag{3.32}$$

Proof. Recall the Gauss equation A.24

$$\langle \mathcal{R}_\theta(u,v)\, u, v\rangle_{H^1} = \langle \mathbb{I}_\theta^\Theta(u,v), \mathbb{I}_\theta^\Theta(v,u)\rangle_{H^1} - \langle \mathbb{I}_\theta^\Theta(u,u), \mathbb{I}_\theta^\Theta(v,v)\rangle_{H^1}.$$

By using Cauchy-Schwarz and Theorem 3.7 the statement follows. □

3.3 Lipschitz-type properties

In this part we prove some properties and estimates as preparations for later usage. Mostly they are local Lipschitz continuity properties. For this we use *geodesic balls* in Θ. That is, for $\theta \in \Theta$ and $\varrho \geq 0$ the geodesic ball $B_\varrho^\Theta(\theta)$ at θ of radius ϱ is given by

$$B_\varrho^\Theta(\theta) := \left\{\gamma(1) \in \Theta \mid \gamma \text{ is a geodesic, } \gamma(0) = \theta,\ \|\gamma'(0)\|_{\gamma(0)} \leq \varrho\right\}. \tag{3.33}$$

For a sake of simplicity we always assume that ϱ is sufficiently small to guarantee the existence of minimizing geodesics between θ and $\widehat{\theta}$ for all $\widehat{\theta} \in B_\varrho^\Theta(\theta)$.

Lemma 3.9. *Let $\theta \in \Theta$, $\varrho \geq 0$ and $\widehat{\theta} \in B_\varrho^\Theta(\theta)$. We denote the parallel transport along the minimizing geodesic between θ to $\widehat{\theta}$ by $\mathcal{P}_\theta^{\widehat{\theta}} \colon \mathcal{T}_\theta\Theta \to \mathcal{T}_{\widehat{\theta}}\Theta$. Then there is a bound $C_{\mathcal{P}}$ depending on θ and ϱ (explicit formula given in the proof), such that*

$$\left\| u - \mathcal{P}_\theta^{\widehat{\theta}} u \right\|_{H^1} \leq C_{\mathcal{P}}(\theta,\varrho)\, dist(\theta,\widehat{\theta})\, \|u\|_{H^1} \qquad \textit{for all } u \in \mathcal{T}_\theta\Theta. \tag{3.34}$$

Before we prove the statement it is to emphasizes that the left-hand side in (3.34) is a difference of elements in the surrounding space $H^1([0,1])$.

Proof. Let γ be the minimizing geodesic connecting θ and $\widehat{\theta}$ with $\gamma(a) = \theta$ and $\gamma(b) = \widehat{\theta}$, where its length L is given by $L = \int_a^b \|\gamma'(t)\|_{H^1}\, \mathrm{d}t$. It follows immediately that $\gamma(t) \in B_\varrho^\Theta(\theta)$ for all t. Furthermore we can assume that it is parameterized by the arc-length, i.e. $\|\gamma'(t)\|_{H^1} = 1$ for all $t \in [a,b]$. Therefore, it holds that

$$L = \int_a^b \|\gamma'(t)\|_{H^1}\, \mathrm{d}t = |b-a|. \tag{3.35}$$

Let $u \in \mathcal{T}_\theta \Theta$ and define the tangential vector field V

$$V(\gamma(t)) := \mathcal{P}_{\gamma(0)}^{\gamma(t)} u \in \mathcal{T}_{\gamma(t)}\Theta$$

along γ. For a normal field N along γ, defined by $N(\gamma(t)) \in \mathcal{N}_{\gamma(t)}\Theta = \mathcal{T}_{\gamma(t)}\Theta^\perp$, it holds therefore that

$$\langle V(\gamma(t)), N(\gamma(t))\rangle_{H^1} = 0$$

for all $t \in [a, b]$ and by taking the derivative with respect to t, one gets

$$\Big\langle \frac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t)), N(\gamma(t))\Big\rangle_{H^1} = -\Big\langle V(\gamma(t)), \frac{\mathrm{d}}{\mathrm{d}t} N(\gamma(t))\Big\rangle_{H^1}. \tag{3.36}$$

The expression $\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))$ is a short version of the derivative $\mathrm{D}V(\gamma(t))\,(\gamma'(t))$ of V along the direction γ'. Since $V(\gamma(t))$ is γ-parallel it holds that

$$\Big\langle \frac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t)), W(\gamma(t))\Big\rangle_{H^1} = 0 \tag{3.37}$$

for all vector fields $W(\gamma(t)) \in \mathcal{T}_{\gamma(t)}\Theta$, which is equivalent to $\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))$ being a normal field. Define for $\theta \in \Theta$ the orthogonal projection $P_\theta \colon H^1([0,1]) \to \mathcal{T}_\theta\Theta$ and note that

$$V(\gamma(t)) = P_{\gamma(t)}\big(V(\gamma(t))\big). \tag{3.38}$$

Combining these facts with (3.36) and using that the projection is idempotent and self-adjoint, we can conclude

$$\begin{aligned} \Big\|\frac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t))\Big\|_{H^1}^2 &= -\Big\langle P_{\gamma(t)}\big(V(\gamma(t))\big), P_{\gamma(t)}\Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2} V(\gamma(t))\Big)\Big\rangle_{H^1} \\ &\leq \|V(\gamma(t))\|_{H^1} \Big\|P_{\gamma(t)}\Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2} V\big(\gamma(t)\big)\Big)\Big\|_{H^1}. \end{aligned} \tag{3.39}$$

Recall that the derivative of a normal field corresponds to the Weingarten map given by (A.22). Furthermore the equation (A.23) connects it with the second fundamental form. Using those, Cauchy-Schwarz and Theorem 3.7 it follows that

$$\begin{aligned} \Big\|P_{\gamma(t)}\Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2} V\big(\gamma(t)\big)\Big)\Big\|_{H^1}^2 &= \big\|\mathcal{S}_{\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))}(\gamma'(t))\big\|_{H^1}^2 \\ &= \Big\langle \frac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t)), \mathrm{I\!I}\big(\gamma'(t), \mathcal{S}_{\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))}(\gamma'(t))\big)\Big\rangle_{H^1} \\ &\leq C_{\mathrm{I\!I}}(\gamma(t)) \Big\|\frac{\mathrm{d}}{\mathrm{d}t} V(\gamma(t))\Big\|_{H^1} \big\|\mathcal{S}_{\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))}(\gamma'(t))\big\|_{H^1} \end{aligned}$$

where $\|\gamma'(t)\|_{H^1} = 1$ holds by construction as the arc length derivative. Dividing by the norm of the Weingarten map gives the bound for the left-hand side. The parallel transport maps isometrically through the tangent spaces and thus

$$\left\|V(\gamma(t))\right\|_{H^1} = \|\mathcal{P}_{\gamma(a)}^{\gamma(t)} u\|_{H^1} = \|u\|_{H^1}.$$

By combining this with the last two inequalities gives

$$\begin{aligned}\left\|\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))\right\|_{H^1}^2 &\le C_{\mathbb{I}}(\gamma(t)) \left\|\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))\right\|_{H^1} \|V(\gamma(t))\|_{H^1} \\ &= C_{\mathbb{I}}(\gamma(t)) \left\|\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))\right\|_{H^1} \|u\|_{H^1} \qquad (3.40)\end{aligned}$$

which can divided by the norm of $\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))$ to obtain a bound for the left-hand side. Now by using the above constructions we deduce

$$\begin{aligned}\left\| u - \mathcal{P}_{\theta}^{\widehat{\theta}} u \right\|_{H^1} &= \left\|\mathcal{P}_{\gamma(a)}^{\gamma(a)} u - \mathcal{P}_{\gamma(a)}^{\gamma(b)} u\right\|_{H^1} \\ &= \left\|V(\gamma(b)) - V(\gamma(a))\right\|_{H^1} \\ &\le \int_a^b \left\|\frac{\mathrm{d}}{\mathrm{d}t}V(\gamma(t))\right\|_{H^1} \mathrm{d}t \le \sup_{t\in[a,b]} \left\{C_{\mathbb{I}}(\gamma(t))\right\} |b-a| \, \|u\|_{H^1}. \qquad (3.41)\end{aligned}$$

Since γ connects θ and $\widehat{\theta}$ as a geodesic path of length L it is by (3.35)

$$|b-a| = L = \mathit{dist}(\theta, \widehat{\theta}). \qquad (3.42)$$

Only the consideration of the term $C_{\mathbb{I}}(\gamma(t))$ is left. Recall the definition of this bound $C_{\mathbb{I}}(\gamma(t)) = 2\sqrt{2}\, C_D(\gamma(t))$ in (3.30). The term C_D given in (3.18) can be written as a polynomial p_D in $\|\theta'\|_{L^2}$ by $p_D(\|\theta'\|_{L^2}) := C_D(\theta)$. The derivative of p_D

$$p_D'(t) = \frac{16}{\pi^2}\, 2t \left(\sqrt{2}\, t + \sqrt{2} + 1\right) \left(4\sqrt{2}\, t^2 + \left(3 + 5\sqrt{2}\right) t + \sqrt{2} + 1\right) (1+t)^3$$

in monotonous increasing in the positive reals and thus we can estimate

$$\begin{aligned}C_D(\gamma(t)) &= C_D(\theta) + p_D(\|\gamma'(t)\|_{L^2}) - p_D(\|\theta'\|_{L^2}) \\ &= C_D(\theta) + \int_{\|\theta'\|_{L^2}}^{\|\gamma'(t)\|_{L^2}} p_D'(\xi)\, \mathrm{d}\xi \\ &\le C_D(\theta) + \Big(\sup_{\xi\in[\|\theta'\|_{L^2}, \|\gamma'(t)\|_{L^2}]} p_D'(\xi)\Big) \left|\|\gamma'(t)\|_{L^2} - \|\theta'\|_{L^2}\right|\end{aligned}$$

for $\|\theta'\|_{L^2} \le \|\gamma'(t)\|_{L^2}$. In the other case we obtain the analog inequality by exchanging the roles of $\|\theta'\|_{L^2}$ and $\|\gamma'(t)\|_{L^2}$. Since γ is parameterized by the arc-length and thus

$\|\gamma'(t)\|_{L^2} \leq \|\gamma'(t)\|_{H^1} \leq 1$, it follows that

$$\max\Big\{ \sup_{\xi\in[\|\theta'\|_{L^2},\|\gamma'(t)\|_{L^2}]} p_D'(\xi),\ \sup_{\xi\in[\|\gamma'(t)\|_{L^2},\|\theta'\|_{L^2}]} p_D'(\xi)\Big\} \leq \max\left\{ p_D'(\|\theta'\|_{L^2}),\, 1\right\}.$$

By using that the norm difference in the surrounding Hilbert space is always a lower bound on the Riemannian distance and that $\gamma(t) \in B_\varrho^\Theta(\theta)$ we deduce

$$\left|\|\gamma'(t)\|_{L^2} - \|\theta'\|_{L^2}\right| \leq \|\gamma'(t) - \theta'\|_{L^2} \leq \|\gamma(t) - \theta\|_{H^1} \leq dist(\gamma(t), \theta) \leq \varrho.$$

By combining (3.30) with the estimates (3.41), (3.42) and

$$C_{\mathcal{P}}(\theta, \varrho) := 2\sqrt{2}\left(C_D(\theta) + \max\left\{ p_D'(\|\theta'\|_{L^2}),\, 1\right\}\varrho\right) \tag{3.43}$$

the statement follows. □

We use now the last lemma to prove in the following local Lipschitz continuity of $\mathrm{D}^2\Phi$.

Lemma 3.10. *The operator* $\mathrm{D}^2\Phi\colon \Theta \to L(\mathcal{T}\Theta\times\mathcal{T}\Theta, \mathbb{R}^2)$ *is locally Lipschitz continuous, i.e. for all* $\theta \in \Theta$ *and sufficiently small* $\varrho \geq 0$ *there is a bound* $C_{\mathrm{D}2\Phi}(\theta, \varrho) > 0$ *depending on* θ *and* ϱ *(explicit formula in the proof) such that*

$$|\mathrm{D}^2\Phi(\theta)(u, u) - \mathrm{D}^2\Phi(\widehat{\theta})(\mathcal{P}_\theta^{\widehat{\theta}}u, \mathcal{P}_\theta^{\widehat{\theta}}u)|_{\mathbb{R}^2} \leq C_{\mathrm{D}2\Phi}(\theta, \varrho)\, dist(\theta, \widehat{\theta})\, \|u\|_{H^1}^2 \tag{3.44}$$

for all $u \in \mathcal{T}_\theta\Theta$ *and all* $\widehat{\theta} \in B_\varrho^\Theta(\theta)$.

Proof. Let $\theta \in \Theta$, $\widehat{\theta} \in B_\varrho^\Theta(\theta)$ for a $\varrho \geq 0$ and $u \in \mathcal{T}_\theta\Theta$ be given. Then it is

$$\begin{aligned}
&|\mathrm{D}^2\Phi(\theta)(u, u) - \mathrm{D}^2\Phi(\widehat{\theta})(\mathcal{P}_\theta^{\widehat{\theta}}u, \mathcal{P}_\theta^{\widehat{\theta}}u)|_{\mathbb{R}^2}^2 \\
=&\Big(\int_0^1 c_\theta(s)\,(u(s))^2 - c_{\widehat{\theta}}(s)\left(\mathcal{P}_\theta^{\widehat{\theta}}u(s)\right)^2 \mathrm{d}s\Big)^2 \\
&\qquad + \Big(\int_0^1 s_\theta(s)\,(u(s))^2 - s_{\widehat{\theta}}(s)\left(\mathcal{P}_\theta^{\widehat{\theta}}u(s)\right)^2 \mathrm{d}s\Big)^2.
\end{aligned} \tag{3.45}$$

Here the first term can be bounded as follows:

$$\begin{aligned}
&\Big(\int_0^1 c_\theta(s)\,(u(s))^2 - c_{\widehat{\theta}}(s)\left(\mathcal{P}_\theta^{\widehat{\theta}}u(s)\right)^2 \mathrm{d}s\Big)^2 \\
\leq&\int_0^1 \left(c_\theta(s)\,(u(s))^2 - c_{\widehat{\theta}}(s)\left(\mathcal{P}_\theta^{\widehat{\theta}}u(s)\right)^2\right)^2 \mathrm{d}s \\
\leq&2\int_0^1 \left(c_\theta(s) - c_{\widehat{\theta}}(s)\right)^2 \left(u(s)\right)^4 \mathrm{d}s + 2\int_0^1 \left(c_{\widehat{\theta}}(s)\right)^2 \left(u^2(s) - \left(\mathcal{P}_\theta^{\widehat{\theta}}u(s)\right)^2\right)^2 \mathrm{d}s.
\end{aligned} \tag{3.46}$$

Recall the Morrey embedding $H^1_{\text{per}}([0,1]) \hookrightarrow C^0([0,1])$, see for example [33, Thm. 8.4]. Therefore, there is a constant C_∞ such that

$$\|v\|_\infty \leq C_\infty \|v\|_{H^1} \tag{3.47}$$

for $v \in H^1_{\text{per}}([0,1])$. The first term on the right-hand side of (3.46) can be bounded using standard estimates, $s_{\text{id}}(t) = \sin(t)$ and (3.47) by

$$\begin{aligned}
\int_0^1 \left(c_\theta(s) - c_{\widehat{\theta}}(s)\right)^2 \left(u(s)\right)^4 \mathrm{d}s &\leq \|u\|_\infty^4 \int_0^1 \left(\cos(\theta(s)) - \cos(\widehat{\theta}(s))\right)^2 \mathrm{d}s \\
&\leq C_\infty^4 \|u\|_{H^1}^4 \int_0^1 \left(\int_{\widehat{\theta}(s)}^{\theta(s)} \sin(\xi)\, \mathrm{d}\xi \right)^2 \mathrm{d}s \\
&\leq C_\infty^4 \|u\|_{H^1}^4 \int_0^1 \|s_{\text{id}}\|_\infty^2 \left(\theta(s) - \widehat{\theta}(s)\right)^2 \mathrm{d}s \\
&= C_\infty^4 \|u\|_{H^1}^4 \|\theta - \widehat{\theta}\|_{L^2}^2 \\
&\leq C_\infty^4 \, \mathit{dist}(\theta, \widehat{\theta})^2 \|u\|_{H^1}^4 .
\end{aligned}$$

In the last step we used that the norm distance in a submanifold is always a lower bound to the Riemannian distance, i.e. $\|\theta - \widehat{\theta}\|_{H^1} \leq \mathit{dist}(\theta, \widehat{\theta})$. Further we use (3.34) from Lemma 3.9 to obtain

$$\begin{aligned}
\int_0^1 \left(c_{\widehat{\theta}}(s)\right)^2 \left(u^2(s) - (\mathcal{P}_\theta^{\widehat{\theta}} u(s))^2\right)^2 \mathrm{d}s &\leq \|c_{\widehat{\theta}}\|_\infty^2 \left\|(u - \mathcal{P}_\theta^{\widehat{\theta}} u)u + (\mathcal{P}_\theta^{\widehat{\theta}} u)\left(u - \mathcal{P}_\theta^{\widehat{\theta}} u\right)\right\|_{L^2}^2 \\
&\leq 2\left(\|u\|_\infty^2 + \left\|\mathcal{P}_\theta^{\widehat{\theta}} u\right\|_\infty^2\right) \left\| u - \mathcal{P}_\theta^{\widehat{\theta}} u \right\|_{L^2}^2 \\
&\leq 4 C_\infty^2 (C_{\mathcal{P}}(\theta, \varrho))^2 \mathit{dist}(\theta, \widehat{\theta})^2 \|u\|_{H^1}^4 .
\end{aligned}$$

The sum of the last two estimates provides a bound of the first term on the right-hand side in (3.45). For the second term one proves the analogous estimates by replacing cosine with sine. Therefore, by using

$$C_{\mathrm{D}2\Phi}(\theta, \varrho) := \sqrt{C_\infty^4 + 8 C_\infty^2 (C_{\mathcal{P}}(\theta, \varrho))^2}.$$

we obtain the statement. □

Before we can prove a Lipschitz property of $\mathrm{D}\Phi(\theta)^\dagger$ we need the following proposition as preparation.

Proposition 3.11. *The maps* $\boldsymbol{g}\colon \Theta \to H^1_{\text{per}}([0,1])$, $\theta \mapsto g_\theta$ *and* $\boldsymbol{f}\colon \Theta \to H^1_{\text{per}}([0,1])$, $\theta \mapsto f_\theta$ *are differentiable and the derivatives can be bounded independently of* θ*, i.e.*

there is a constant $C_{\mathrm{e}} > 0$ *such that for all* $\theta \in \Theta$

$$\|\mathrm{D}\boldsymbol{g}(\theta)\,(u)\|_{H^1} \le C_{\mathrm{e}}\,\|u\|_{H^1} \tag{3.48}$$

and

$$\|\mathrm{D}\boldsymbol{f}(\theta)\,(u)\|_{H^1} \le C_{\mathrm{e}}\,\|u\|_{H^1} \tag{3.49}$$

for all $u \in \mathcal{T}_\theta\Theta$.

Proof. Recall that

$$\begin{aligned} g_\theta(t) = &-\frac{1}{2}\Big(\int_0^t \mathrm{e}^{-\xi} c_\theta(\xi)\,\mathrm{d}\xi + \frac{e}{1-e}\int_0^1 \mathrm{e}^{-\xi} c_\theta(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^t \\ &+ \frac{1}{2}\Big(\int_0^t \mathrm{e}^{\xi} c_\theta(\xi)\,\mathrm{d}\xi + \frac{e^{-1}}{1-e^{-1}}\int_0^1 \mathrm{e}^{\xi} c_\theta(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^{-t}. \end{aligned}$$

This is clearly differentiable in θ with derivative

$$\begin{aligned} \mathrm{D}\boldsymbol{g}(\theta)\,(u)(t) = &\frac{1}{2}\Big(\int_0^t \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e}{1-e}\int_0^1 \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^t \\ &- \frac{1}{2}\Big(\int_0^t \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e^{-1}}{1-e^{-1}}\int_0^1 \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^{-t} \end{aligned}$$

and its derivative with respect to t is given by

$$\begin{aligned} \big(\mathrm{D}\boldsymbol{g}(\theta)\,(u)\big)'(t) = &\frac{1}{2}\Big(\int_0^t \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e}{1-e}\int_0^1 \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^t \\ &+ \frac{1}{2}\Big(\int_0^t \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e^{-1}}{1-e^{-1}}\int_0^1 \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)\mathrm{e}^{-t}. \end{aligned}$$

Using $(a+b)^2 \le 2a^2 + 2b^2$, we can derive bounds for these expressions:

$$\begin{aligned} \|\mathrm{D}\boldsymbol{g}(\theta)\,(u)\|_{H^1}^2 \le\ &2\int_0^1 \Big(\int_0^t \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e}{1-e}\int_0^1 \mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)^2 \mathrm{e}^{2t}\,\mathrm{d}t \\ &+ 2\int_0^1 \Big(\int_0^t \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi + \frac{e^{-1}}{1-e^{-1}}\int_0^1 \mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\,\mathrm{d}\xi\Big)^2 \mathrm{e}^{-2t}\,\mathrm{d}t \\ \le\ &2\Big(1 + \Big(\frac{e}{1-e}\Big)^2\Big)\int_0^1 \big(\mathrm{e}^{-\xi} s_\theta(\xi)u(\xi)\big)^2\,\mathrm{d}\xi \int_0^1 \mathrm{e}^{2t}\,\mathrm{d}t \\ &+ 2\Big(1 + \Big(\frac{e^{-1}}{1-e^{-1}}\Big)^2\Big)\int_0^1 \big(\mathrm{e}^{\xi} s_\theta(\xi)u(\xi)\big)^2\,\mathrm{d}\xi \int_0^1 \mathrm{e}^{-2t}\,\mathrm{d}t \\ \le\ &\Big(2 + \Big(\frac{e}{1-e}\Big)^2 + \Big(\frac{e^{-1}}{1-e^{-1}}\Big)^2\Big)\mathrm{e}^2\|s_\theta u\|_{L^2}^2. \end{aligned}$$

By using $\|s_\theta u\|_{L^2} \le \|s_\theta\|_\infty\,\|u\|_{L^2} \le \|u\|_{H^1}$ bound in the statement follows with the constant

$$C_{\mathrm{e}} := \mathrm{e}\sqrt{2 + \left(\frac{e}{1-e}\right)^2 + \left(\frac{e^{-1}}{1-e^{-1}}\right)^2}.$$

By complete analogous computations for f_θ one obtains the same bound, which finishes the proof. □

In the last lemma of this section we prove that the Moore-Penrose inverse of $\mathrm{D}\Phi$ is also locally Lipschitz continuous.

Lemma 3.12. *The operator $\theta \mapsto \mathrm{D}\Phi(\theta)^\dagger$ is locally Lipschitz continuous, i.e. for all $\theta \in \Theta$ and sufficiently small $\varrho \geq 0$ there exists a positive bound $C_{\mathrm{D}\Phi^\dagger}(\theta, \varrho) > 0$ such that*

$$\left\| \mathrm{D}\Phi(\theta)^\dagger(x) - \mathrm{D}\Phi(\widehat{\theta})^\dagger(x) \right\|_{H^1} \leq C_{\mathrm{D}\Phi^\dagger}(\theta, \varrho)\, dist(\theta, \widehat{\theta})\, |x|_{\mathbb{R}^2} \tag{3.50}$$

for all $x \in \mathbb{R}^2$ and $\widehat{\theta} \in B^\Theta_\varrho(\theta)$

Proof. Let $\theta \in \Theta$, $\varrho \geq 0$, $\widehat{\theta} \in B^\Theta_\varrho(\theta)$ and $x = (x_1, x_2)^\top \in \mathbb{R}^2$ be given. From the expression (3.10) we add suitable zero terms and obtain a decomposition for the estimate as follows:

$$\begin{aligned}
&\left\|\mathrm{D}\Phi(\theta)^\dagger(x) - \mathrm{D}\Phi(\widehat{\theta})^\dagger(x)\right\|_{H^1} \\
=&\left\| \frac{-\left(\|g_\theta\|^2_{H^1} x_1 + \langle g_\theta, f_\theta\rangle_{H^1} x_2\right)}{D_\theta} f_\theta + \frac{\langle f_\theta, g_\theta\rangle_{H^1} x_1 + \|f_\theta\|^2_{H^1} x_2}{D_\theta} g_\theta \right. \\
&\left. \quad - \frac{-\left(\|g_{\widehat{\theta}}\|^2_{H^1} x_1 + \langle g_{\widehat{\theta}}, f_{\widehat{\theta}}\rangle_{H^1} x_2\right)}{D_{\widehat{\theta}}} f_{\widehat{\theta}} - \frac{\langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} x_1 + \|f_{\widehat{\theta}}\|^2_{H^1} x_2}{D_{\widehat{\theta}}} g_{\widehat{\theta}} \right\|_{H^1} \\
\leq& \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5,
\end{aligned}$$

where we define

$$\begin{aligned}
\Sigma_1 &:= \left|\frac{1}{D_\theta} - \frac{1}{D_{\widehat{\theta}}}\right| \left\|\left(\|g_\theta\|^2_{H^1} x_1 + \langle g_\theta, f_\theta\rangle_{H^1} x_2\right) f_\theta - \left(\langle f_\theta, g_\theta\rangle_{H^1} x_1 + \|f_\theta\|^2_{H^1} x_2\right) g_\theta\right\|_{H^1}, \\
\Sigma_2 &:= \frac{1}{D_{\widehat{\theta}}} \left\|\left(\left(\|g_\theta\|^2_{H^1} x_1 + \langle g_\theta, f_\theta\rangle_{H^1} x_2\right) - \left(\|g_{\widehat{\theta}}\|^2_{H^1} x_1 + \langle g_{\widehat{\theta}}, f_{\widehat{\theta}}\rangle_{H^1} x_2\right)\right) f_\theta\right\|_{H^1}, \\
\Sigma_3 &:= \frac{1}{D_{\widehat{\theta}}} \left\|\left(\left(\langle f_\theta, g_\theta\rangle_{H^1} x_1 + \|f_\theta\|^2_{H^1} x_2\right) - \left(\langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} x_1 + \|f_{\widehat{\theta}}\|^2_{H^1} x_2\right)\right) g_\theta\right\|_{H^1}, \\
\Sigma_4 &:= \frac{1}{D_{\widehat{\theta}}} \left\|\left(\|g_{\widehat{\theta}}\|^2_{H^1} x_1 + \langle g_{\widehat{\theta}}, f_{\widehat{\theta}}\rangle_{H^1} x_2\right)\left(f_\theta - f_{\widehat{\theta}}\right)\right\|_{H^1}, \\
\Sigma_5 &:= \frac{1}{D_{\widehat{\theta}}} \left\|\left(\langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} x_1 + \|f_{\widehat{\theta}}\|^2_{H^1} x_2\right)\left(g_\theta - g_{\widehat{\theta}}\right)\right\|_{H^1}.
\end{aligned}$$

In the following we estimate all five summands Σ_j such that

$$\Sigma_j \leq C_{\Sigma_j}(\theta, \widehat{\theta})\, dist(\theta, \widehat{\theta})\, |x|_{\mathbb{R}^2}.$$

Starting with the first it is

$$\begin{aligned}
\left|\frac{1}{D_\theta} - \frac{1}{D_{\widehat{\theta}}}\right| &= \frac{|D_{\widehat{\theta}} - D_\theta|}{D_\theta D_{\widehat{\theta}}} \\
&= \frac{1}{D_\theta D_{\widehat{\theta}}} \Big| \|f_{\widehat{\theta}}\|_{H^1}^2 \|g_{\widehat{\theta}}\|_{H^1}^2 - |\langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1}|^2 \\
&\qquad\qquad - \|f_\theta\|_{H^1}^2 \|g_\theta\|_{H^1}^2 + |\langle f_\theta, g_\theta\rangle_{H^1}|^2 \Big| \\
&\leq \frac{1}{D_\theta D_{\widehat{\theta}}} \Big(\big| \|f_{\widehat{\theta}}\|_{H^1}^2 - \|f_\theta\|_{H^1}^2 \big| \, \|g_{\widehat{\theta}}\|_{H^1}^2 + \|f_\theta\|_{H^1}^2 \big| \|g_{\widehat{\theta}}\|_{H^1}^2 - \|g_\theta\|_{H^1}^2 \big| \\
&\qquad + \big| \langle f_\theta, g_\theta\rangle_{H^1} + \langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} \big| \, \big| \langle f_\theta, g_\theta\rangle_{H^1} - \langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} \big| \Big).
\end{aligned}$$

Here we can use $\|f_\theta\|_{H^1} \leq 1$ to obtain that

$$\begin{aligned}
\big| \|f_{\widehat{\theta}}\|_{H^1}^2 - \|f_\theta\|_{H^1}^2 \big| &= \Big| \int_0^1 (f_{\widehat{\theta}}(t))^2 - (f_\theta(t))^2 + (f'_{\widehat{\theta}}(t))^2 - (f'_\theta(t))^2 \,\mathrm{d}t \Big| \\
&= \Big| \int_0^1 (f_{\widehat{\theta}}(t) - f_\theta(t)) f_{\widehat{\theta}}(t) + f_\theta(t)(f_{\widehat{\theta}}(t) - f_\theta(t)) \\
&\qquad + (f'_{\widehat{\theta}}(t) - f'_\theta(t)) f'_{\widehat{\theta}}(t) + f'_\theta(t)(f'_{\widehat{\theta}}(t) - f'_\theta(t)) \,\mathrm{d}t \Big| \\
&\leq | \|f_{\widehat{\theta}} - f_\theta\|_{L^2} \left(\|f_{\widehat{\theta}}\|_{L^2} + \|f_\theta\|_{L^2} \right) + \|f'_{\widehat{\theta}} - f'_\theta\|_{L^2} \left(\|f'_{\widehat{\theta}}\|_{L^2} + \|f'_\theta\|_{L^2} \right) | \\
&\leq 2 \|f_{\widehat{\theta}} - f_\theta\|_{H^1}.
\end{aligned}$$

Now we can apply Proposition 3.11 as follows: using $\gamma(t) := exp_\theta(t\, exp_\theta^{-1}(\widehat{\theta}))$, we get

$$\begin{aligned}
2\|f_{\widehat{\theta}} - f_\theta\|_{H^1} \leq 2 \int_0^1 \|\mathrm{D}\boldsymbol{f}(\gamma(t))\,(exp_\theta^{-1}(\widehat{\theta}))\|_{H^1} \,\mathrm{d}t &\leq 2C_{\mathrm{e}}\, \|exp_\theta^{-1}(\widehat{\theta})\|_{H^1} \\
&= 2C_{\mathrm{e}}\, dist(\theta, \widehat{\theta}),
\end{aligned}$$

where the inequality follows by Proposition 3.11 and the equality is due to (A.9). The analog estimate holds true for the terms with g_θ and $g_{\widehat{\theta}}$: Similarly to above, it holds that

$$\big| \|g_{\widehat{\theta}}\|_{H^1}^2 - \|g_\theta\|_{H^1}^2 \big| \leq 2C_{\mathrm{e}}\, dist(\theta, \widehat{\theta}).$$

Similar to above it is

$$\big| \langle f_\theta, g_\theta\rangle_{H^1} - \langle f_{\widehat{\theta}}, g_{\widehat{\theta}}\rangle_{H^1} \big| \leq \big| \langle f_\theta - f_{\widehat{\theta}}, g_\theta\rangle_{H^1} \big| + \big| \langle f_{\widehat{\theta}}, g_\theta - g_{\widehat{\theta}}\rangle_{H^1} \big|$$

$$\leq \|f_\theta - f_{\widehat{\theta}}\|_{H^1} + \|g_\theta - g_{\widehat{\theta}}\|_{H^1} \leq 2C_e\, dist(\theta, \widehat{\theta}).$$

The other part of Σ_1 can be estimated using $\|f_\theta\|_{H^1} \leq 1$ and $\|g_\theta\|_{H^1} \leq 1$ by

$$\Big\|\big(\|g_\theta\|_{H^1}^2\, x_1 + \langle g_\theta, f_\theta\rangle_{H^1}\, x_2\big) f_\theta - \big(\langle f_\theta, g_\theta\rangle_{H^1}\, x_1 + \|f_\theta\|_{H^1}^2\, x_2\big) g_\theta\Big\|_{H^1} \leq 2\,\big(|x_1| + |x_2|\big) \leq 2\sqrt{2}\,|x|_{\mathbb{R}^2}$$

Thus we define

$$C_{\Sigma_1}(\theta, \widehat{\theta}) := \frac{16\sqrt{2}\,C_e}{D_\theta\, D_{\widehat{\theta}}}.$$

The side computations for Σ_1 can also be used in the following estimates. We obtain

$$\begin{aligned}
\Sigma_2 &= \frac{1}{D_{\widehat{\theta}}} \Big\|\Big(\big(\|g_\theta\|_{H^1}^2\, x_1 + \langle g_\theta, f_\theta\rangle_{H^1}\, x_2\big) - \big(\|g_{\widehat{\theta}}\|_{H^1}^2\, x_1 + \langle g_{\widehat{\theta}}, f_{\widehat{\theta}}\rangle_{H^1}\, x_2\big)\Big) f_\theta\Big\|_{H^1} \\
&\leq \frac{1}{D_{\widehat{\theta}}} \Big(\big|\|g_{\widehat{\theta}}\|_{H^1}^2 - \|g_\theta\|_{H^1}^2\big|\,|x_1| + \big|\langle g_\theta, f_\theta\rangle_{H^1} - \langle g_{\widehat{\theta}}, f_{\widehat{\theta}}\rangle_{H^1}\big|\,|x_2|\Big) \\
&\leq \frac{1}{D_{\widehat{\theta}}}\, 4C_e\, dist(\theta, \widehat{\theta})\big(|x_1| + |x_2|\big)
\end{aligned}$$

and thus we define

$$C_{\Sigma_2}(\theta, \widehat{\theta}) := \frac{4\sqrt{2}\,C_e}{D_{\widehat{\theta}}}.$$

The computations for Σ_3 work completely analogously and we define the same constant

$$C_{\Sigma_3}(\theta, \widehat{\theta}) := \frac{4\sqrt{2}\,C_e}{D_{\widehat{\theta}}}.$$

Also for the last two terms Σ_4 and Σ_5 the estimates are straightforward and by applying the above estimates one obtains

$$C_{\Sigma_4}(\theta, \widehat{\theta}) := \frac{4\sqrt{2}\,C_e}{D_{\widehat{\theta}}} \qquad \text{and} \qquad C_{\Sigma_5}(\theta, \widehat{\theta}) := \frac{4\sqrt{2}\,C_e}{D_{\widehat{\theta}}}.$$

Now we can apply the lower bound from Lemma 3.5 to obtain

$$C(\theta, \widehat{\theta}) := 16\sqrt{2}\,C_e\,\big(C_D(\theta) + 1\big)\,C_D(\widehat{\theta}). \tag{3.51}$$

As last step we estimate the term $C_D(\widehat{\theta})$. Recall the notation at the end of the proof of Lemma 3.9 and write $p_D(\|\widehat{\theta}'\|_{L^2}) = C_D(\widehat{\theta})$. Analogously to the proof of Lemma 3.9

we deduce

$$\begin{aligned} C_D(\widehat{\theta}) &\le C_D(\theta) + \max\left\{ \int_{\|\theta'\|_{L^2}}^{\|\widehat{\theta}'\|_{L^2}} p_D'(\xi)\,\mathrm{d}\xi\,,\ \int_{\|\widehat{\theta}'\|_{L^2}}^{\|\theta'\|_{L^2}} p_D'(\xi)\,\mathrm{d}\xi \right\} \\ &\le C_D(\theta) + \max\left\{ p_D'(\|\theta'\|_{L^2})\,,\ p_D'(\|\widehat{\theta}'\|_{L^2}) \right\}. \end{aligned}$$

From

$$\|\widehat{\theta}'\|_{L^2} - \|\theta'\|_{L^2} \le \|\widehat{\theta}' - \theta'\|_{L^2} \le \|\widehat{\theta} - \theta\|_{H^1} \le \mathit{dist}(\widehat{\theta}, \theta) \le \varrho \tag{3.52}$$

and the monotonicity of p_D' in the positive real numbers we follow that

$$\begin{aligned} C_D(\widehat{\theta}) &\le C_D(\theta) + \max\left\{ p_D'(\|\theta'\|_{L^2})\,,\ p_D'(\|\theta'\|_{L^2} + \varrho) \right\} \\ &\le C_D(\theta) + p_D'(\|\theta'\|_{L^2} + \varrho). \end{aligned} \tag{3.53}$$

Therefore, the combination of the inequality (3.51) with (3.53) and

$$C_{\mathrm{D}\Phi^\dagger}(\theta, \varrho) := 16\sqrt{2}\, C_{\mathrm{e}} \left(C_D(\theta) + 1 \right) \left(C_D(\theta) + p_D'(\|\theta'\|_{L^2} + \varrho) \right) \tag{3.54}$$

proves the statement. □

3.4 Intrinsic curvature

In the last section of the chapter we derived a formula for the second fundamental form, which measures the extrinsic curvature of a submanifold. The well-known Gauss equation (A.24) is a formula for the intrinsic curvature tensor $\mathcal{R}$ on the manifold in terms of the extrinsic curvature. Therefore by Theorem 3.7 one can derive an explicit formula for $\mathcal{R}$. Nevertheless in this section we derive a bound on the curvature tensor without using the second fundamental form. Recall first the definition (A.10) for $\theta \in \Theta$

$$R_\theta(u, v)w = \nabla_U \nabla_V W(\theta) - \nabla_V \nabla_U W(\theta) - \nabla_{[U,V]} W(\theta)$$

and $u, v, w \in \mathcal{T}_\theta \Theta$ and their corresponding U, V, W vector fields with $U(\theta) = u$, $V(\theta) = v$ and $W(\theta) = w$, also shortly written as $U_\theta = u$. Furthermore from (A.19)

$$\nabla_V W(\theta) = P_\theta\big(\mathrm{D}W(\theta)\,(V(\theta))\big)$$

where $P_\theta\colon H^1([0,1]) \to \mathcal{T}_\theta\Theta$ is the orthogonal projection. In the article [17] it was already shown that in this case the curvature tensor can be written as

$$R_\theta(u, v)w = \big(\mathrm{D}\boldsymbol{P}(\theta)\,(V_\theta)\ \mathrm{D}\boldsymbol{P}(\theta)\,(U_\theta) - \mathrm{D}\boldsymbol{P}(\theta)\,(U_\theta)\ \mathrm{D}\boldsymbol{P}(\theta)\,(V_\theta)\big)\, W_\theta \tag{3.55}$$

where we defined the map $\boldsymbol{P}(\theta) := P_\theta$. The idea to handle this expression is to find a suitable way to express and estimate the derivative of $\boldsymbol{P}$.

Recall that the tangent space can be written as

$$\mathcal{T}_\theta\Theta = \left\{u \in H^1_{\mathrm{per}}([0,1]) \,\middle|\, \langle s_\theta, u\rangle_{L^2} = \langle c_\theta, u\rangle_{L^2} = 0\right\} \tag{3.56}$$

where $s_\theta(t) := \sin(\theta(t))$ and $c_\theta(t) := \cos(\theta(t))$. The expression above is written using the orthogonal projection $P^1_\theta = P_\theta \colon H^1([0,1]) \to \mathcal{T}_\theta\Theta$ with respect to the H^1 norm. Instead of using this we define a projection operator with respect to the L^2 norm and show afterwards how one can use it to estimate the derivative of $\boldsymbol{P}$.

First we derive a formula for the orthogonal projection $P^0_\theta \colon H^1_{\mathrm{per}}([0,1]) \to \mathcal{T}_\theta\Theta$ with respect to the L^2 norm. Remark that intuitively such kind of an operator does not need to exists, since $\mathcal{T}_\theta\Theta$ is not closed with respect to L^2. Nevertheless we can define it as follows.

Lemma 3.13. *The orthogonal projection $P^0_\theta \colon H^1_{\mathrm{per}}([0,1]) \to \mathcal{T}_\theta\Theta$ with respect to the L^2 norm exists and is uniquely given by the following formula*

$$P^0_\theta v = v + \alpha_\theta s_\theta + \beta_\theta c_\theta \tag{3.57a}$$

where

$$\alpha_\theta := \frac{\langle c_\theta, s_\theta\rangle_{L^2}\langle c_\theta, v\rangle_{L^2} - \|c_\theta\|^2_{L^2}\langle s_\theta, v\rangle_{L^2}}{D^0_\theta} \tag{3.57b}$$

and

$$\beta_\theta := \frac{\langle s_\theta, c_\theta\rangle_{L^2}\langle s_\theta, v\rangle_{L^2} - \|s_\theta\|^2_{L^2}\langle c_\theta, v\rangle_{L^2}}{D_0(\theta)} \tag{3.57c}$$

with

$$D^0_\theta := \|s_\theta\|^2_{L^2}\|c_\theta\|^2_{L^2} - |\langle s_\theta, c_\theta\rangle_{L^2}|^2. \tag{3.57d}$$

Proof. Let

$$U_\theta := \{u \in L^2([0,1]) \mid \langle s_\theta, u\rangle_{L^2} = \langle c_\theta, u\rangle_{L^2} = 0\}.$$

First we check that this is a closed subspace. From the linear condition the subspace property is clear. For the closeness let $(u_n)_{n\in\mathbb{N}} \subset U_\theta$ a convergent sequence with respect to L^2 with limit $u \in L^2([0,1])$. We verify that $\langle s_\theta, u\rangle_{L^2} = 0$. Remark first that $\langle s_\theta, u\rangle_{L^2} = \langle s_\theta, u - u_n\rangle_{L^2}$ for all $n \in \mathbb{N}$, since $u_n \in U_\theta$. Now by Cauchy-Schwarz

$$|\langle s_\theta, u\rangle_{L^2}| = |\langle s_\theta, u - u_n\rangle_{L^2}| \le \|s_\theta\|_{L^2}\|u - u_n\|_{L^2}.$$

The right-hand side is always bounded since $s_\theta \in H^1([0,1])$ and $\theta \in H^1([0,1])$. Since u_n

converges to u with respect to L^2 it follows that the right-hand side converges to 0 and therefore the left-hand side has to be zero independently of n. Thus $\langle s_\theta, u\rangle_{L^2} = 0$. Analogously, $\langle c_\theta, u\rangle_{L^2} = 0$ and we conclude that $u \in U_\theta$ and hence U_θ is closed. For every closed subspace the orthogonal projection exists and we define $P_{U_\theta} : L^2([0,1]) \longrightarrow U_\theta$ the orthogonal projection with respect to the L^2 norm.

Let $v \in L^2([0,1])$ be given. We define

$$w \coloneqq v + \alpha s_\theta + \beta c_\theta \tag{3.58}$$

with some constants $\alpha, \beta \in \mathbb{R}$. Obviously w satisfies the orthogonality property of a projection onto a subspace:

$$\langle w - v, u\rangle_{L^2} = \alpha\langle s_\theta, u\rangle_{L^2} + \beta\langle c_\theta, u\rangle_{L^2} = 0$$

for all $u \in U_\theta$. Now we choose the constants α and β such that $w \in U_\theta$ and then w is the orthogonal projection of v onto U_θ, i.e. $P_{U_\theta} v \coloneqq w$. Thus w has to satisfy $\langle s_\theta, w\rangle_{L^2} = \langle c_\theta, w\rangle_{L^2} = 0$, which can be written as the linear system

$$\begin{pmatrix} \langle s_\theta, s_\theta\rangle_{L^2} & \langle s_\theta, c_\theta\rangle_{L^2} \\ \langle c_\theta, s_\theta\rangle_{L^2} & \langle c_\theta, c_\theta\rangle_{L^2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = -\begin{pmatrix} \langle s_\theta, v\rangle_{L^2} \\ \langle c_\theta, v\rangle_{L^2} \end{pmatrix}. \tag{3.59}$$

The matrix is invertible since the determinate D_θ^0 given by

$$D_\theta^0 = \|s_\theta\|_{L^2}^2 \|c_\theta\|_{L^2}^2 - |\langle s_\theta, c_\theta\rangle_{L^2}|^2, \tag{3.60}$$

is never equal to zero. We have already seen this in the proof of Theorem 3.1. Then the solution of the system (3.59) is given by

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = -\frac{1}{D_\theta^0} \begin{pmatrix} \langle c_\theta, c_\theta\rangle_{L^2} & -\langle s_\theta, c_\theta\rangle_{L^2} \\ -\langle c_\theta, s_\theta\rangle_{L^2} & \langle s_\theta, s_\theta\rangle_{L^2} \end{pmatrix} \begin{pmatrix} \langle s_\theta, v\rangle_{L^2} \\ \langle c_\theta, v\rangle_{L^2} \end{pmatrix}. \tag{3.61}$$

This proves $P_{U_\theta} v = w$ with w given in (3.58) together with (3.61).

From $\theta \in \Theta$ it follows that

$$s_\theta(1) = \sin(\theta(1)) = \sin(\theta(0) + 2\pi) = s_\theta(0)$$

and thus $s_\theta \in H^1_{\text{per}}([0,1])$. Analogously $c_\theta \in H^1_{\text{per}}([0,1])$. Therefore if $v \in H^1_{\text{per}}([0,1])$ one can conclude from the expression (3.58) that even $P_{U_\theta} v \in H^1_{\text{per}}([0,1])$. Using this we define

$$P_\theta^0 \coloneqq P_{U_\theta}|_{H^1_{\text{per}}} \tag{3.62}$$

the projection in H^1 just as a restriction of the L^2 projection. □

To shorten the notation we write in the rest of this section $H := H^1_{\text{per}}([0,1])$. Now we consider the operators

$$\boldsymbol{P}^j\colon \Theta \to L(H,H), \qquad \theta \mapsto P^j_\theta, \qquad j = 0,1, \tag{3.63}$$

which maps θ to the projections with respect to L^2 respectively H^1 norm on the corresponding tangent spaces. From the expression (3.57) it is clear that $\boldsymbol{P}^0$ is differentiable. Concerning $\boldsymbol{P}^1$ recall form the last section that on the one hand $\mathcal{T}_\theta\Theta = \ker(\mathrm{D}\Phi(\theta))$ and the projection onto it is given by $\boldsymbol{P}^1(\theta) = I - \mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}\Phi(\theta)$, which is differentiable as one can directly see from the Lemmas 3.3 and 3.4.

Knowing the differentiability of the two families of projection operators the next lemma provides a formula for combining their derivatives.

Lemma 3.14. *The derivative of $\boldsymbol{P}^1$ at $\theta \in \Theta$ into the direction $u \in H$ evaluated at $v \in H$ is given by*

$$\mathrm{D}\boldsymbol{P}^1(\theta)\,(P^0_\theta u) = (1 - P^1_\theta)\,\mathrm{D}\boldsymbol{P}^0(\theta)\,(P^0_\theta u). \tag{3.64}$$

The explicit formula for the derivative $\boldsymbol{P}^0$ on the whole space is given in the proof. For $u, v \in \mathcal{T}_\theta\Theta$ it furthermore holds that

$$\begin{aligned}(\mathrm{D}\boldsymbol{P}^0(\theta)\,u)\,v = \frac{1}{D^0_\theta}\Big(&\big(\|s_\theta\|^2_{L^2}\langle s_\theta u, v\rangle_{L^2} + \langle s_\theta, c_\theta\rangle_{L^2}\langle c_\theta u, v\rangle_{L^2}\big)c_\theta \\ &- \big(\|c_\theta\|^2_{L^2}\langle c_\theta u, v\rangle_{L^2} + \langle c_\theta, s_\theta\rangle_{L^2}\langle s_\theta u, v\rangle_{L^2}\big)s_\theta\Big)\end{aligned} \tag{3.65}$$

Proof. Recall that both projections map onto the same space $\mathcal{T}_\theta\Theta$ but with respect to different norms. From this analogously to the facts already used in [17] it is

$$P^1_\theta = P^0_\theta\,P^1_\theta \qquad \text{and} \qquad P^0_\theta = P^1_\theta\,P^0_\theta \tag{3.66}$$

and thus

$$\mathrm{D}\boldsymbol{P}^0(\theta) = \mathrm{D}(\boldsymbol{P}^1\,\boldsymbol{P}^0)(\theta) = \mathrm{D}\boldsymbol{P}^1(\theta)\,P^0_\theta + P^1_\theta\,\mathrm{D}\boldsymbol{P}^0(\theta).$$

After rearrangement and multiplying both sides from the right with P^0_θ we obtain the formula (3.64).

The derivative can be computed straight forward from the expression (3.57):

$$(\mathrm{D}\boldsymbol{P}^0(\theta)\,u)\,v = \big((\mathrm{D}\alpha_\theta\,u)\,v\big)\,s_\theta + \alpha_\theta\,c_\theta\,u + \big((\mathrm{D}\beta_\theta\,u)\,v\big)\,c_\theta - \beta_\theta\,s_\theta\,u.$$

Here we used that the derivative of s_θ in the direction u is given by $(\mathrm{D}s_\theta\,u)(t) = \cos(\theta(t))\,u(t)$ and analogously $(\mathrm{D}c_\theta\,u)(t) = -\sin(\theta(t))\,u(t)$. Further by writing $\alpha_\theta :=$

$N_\alpha(\theta)/D^0_\theta$, where $N_\alpha(\theta)$ is the numerator in (3.57b) and (3.57c), we get

$$(\mathrm{D}\alpha_\theta\, u)\, v = \frac{1}{(D^0_\theta)^2}\Big(\big((\mathrm{D}N_\alpha(\theta)\, u)\, v\big)\, D^0_\theta - N_\alpha(\theta)\,\big((\mathrm{D}D^0_\theta\, u)\, v\big)\Big)$$

where

$$\begin{aligned}(\mathrm{D}N_\alpha(\theta)\, u)\, v = {} & \big(\langle c_\theta u, c_\theta\rangle_{L^2} - \langle s_\theta u, s_\theta\rangle_{L^2}\big)\langle c_\theta, v\rangle_{L^2} - \langle c_\theta, s_\theta\rangle_{L^2}\langle s_\theta u, v\rangle_{L^2}\\ & + 2\langle s_\theta u, c_\theta\rangle_{L^2}\langle s_\theta, v\rangle_{L^2} - \|c_\theta\|^2_{L^2}\langle c_\theta u, v\rangle_{L^2}\end{aligned}$$

and

$$(\mathrm{D}D^0_\theta\, u)\, v = 2\big(\|c_\theta\|^2_{L^2} - \|s_\theta\|^2_{L^2}\big)\langle s_\theta u, c_\theta\rangle_{L^2} - 2\langle s_\theta, c_\theta\rangle_{L^2}\big(\langle c_\theta u, c_\theta\rangle_{L^2} - \langle s_\theta u, s_\theta\rangle_{L^2}\big).$$

Now we write $\beta_\theta \coloneqq N_\beta(\theta)/D^0_\theta$ analogously as in (3.57b) and (3.57c) and compute

$$(\mathrm{D}\beta_\theta\, u)\, v = \frac{1}{(D^0_\theta)^2}\Big(\big((\mathrm{D}N_\beta(\theta)\, u)\, v\big)\, D^0_\theta - N_\beta(\theta)\,\big((\mathrm{D}D^0_\theta\, u)\, v\big)\Big)$$

with

$$\begin{aligned}(\mathrm{D}N_\beta(\theta)\, u)\, v = {} & \big(\langle c_\theta u, c_\theta\rangle_{L^2} - \langle s_\theta u, s_\theta\rangle_{L^2}\big)\langle s_\theta, v\rangle_{L^2} + \langle s_\theta, c_\theta\rangle_{L^2}\langle c_\theta u, v\rangle_{L^2}\\ & - 2\langle c_\theta u, s_\theta\rangle_{L^2}\langle c_\theta, v\rangle_{L^2} + \|s_\theta\|^2_{L^2}\langle s_\theta u, v\rangle_{L^2}\end{aligned}$$

Remark that this function is indeed an element of H. Since $u, v \in H = H^1_{\mathrm{per}}([0,1])$ it is $(s_\theta\, u)(0) = \sin(\theta(0))\, u(0) = \sin(\theta(1) - 2\pi)\, u(1) = (s_\theta\, u)(1)$ and analogously $c_\theta\, u \in H$ and thus $(\mathrm{D}\boldsymbol{P}^0(\theta)\, u)\, v \in H^1_{\mathrm{per}}([0,1])$.

Now let $u, v \in \mathcal{T}_\theta\Theta$. From this we know $\langle c_\theta, v\rangle_{L^2} = \langle s_\theta, v\rangle_{L^2} = 0$. Then $N_\alpha(\theta) = N_\beta(\theta) = 0$ and hence $\alpha_\theta = \beta_\theta = 0$. By leaving out all zero terms we get the formula in (3.65). □

By applying the Lemma 3.14 to (3.55) the next lemma follows.

Lemma 3.15. *Let $\theta \in \Theta$ and $u, v \in \mathcal{T}_\theta\Theta$. Then it holds that*

$$\begin{aligned}\langle \mathcal{R}_\theta(u,v)u, v\rangle_{H^1} = {} & \big\langle (1 - P^1_\theta)\,(\mathrm{D}\boldsymbol{P}^0(\theta)\, u)\, u, (1 - P^1_\theta)\,(\mathrm{D}\boldsymbol{P}^0(\theta)\, v)\, v\big\rangle_{H^1}\\ & - \big\langle (1 - P^1_\theta)\,(\mathrm{D}\boldsymbol{P}^0(\theta)\, v)\, u, (1 - P^1_\theta)\,(\mathrm{D}\boldsymbol{P}^0(\theta)\, u)\, v\big\rangle_{H^1}.\end{aligned} \tag{3.67}$$

Proof. The derivative $\mathrm{D}\boldsymbol{P}^1(\theta)\, u$ inherits the self-adjointness of the projection operators: by taking the derivative of the equation $\langle P^1_\theta u, w\rangle_{H^1} = \langle u, P^1_\theta w\rangle_{H^1}$ one gets

$$\langle (\mathrm{D}\boldsymbol{P}^1(\theta)\, u)\, v, w\rangle_{H^1} = \langle v, (\mathrm{D}\boldsymbol{P}^1(\theta)\, u)\, w\rangle_{H^1}.$$

The formula (3.67) follows by combining this self-adjointness with (3.64) in (3.55). □

Before we can prove a bound on the curvature tensor we need the following lower bound.

Lemma 3.16. *Let $\theta \in \Theta$. Then the lower estimate*

$$D_\theta^0 = \|s_\theta\|_{L^2}^2 \|c_\theta\|_{L^2}^2 - |\langle s_\theta, c_\theta \rangle_{L^2}|^2 \geq \frac{\pi^2}{8\|\theta'\|_{L^2}^2} \tag{3.68}$$

holds true.

Proof. Note that the main idea of this proof is similar to the proof of Lemma 3.5. By the identity

$$\|s_\theta\|_{L^2}^2 + \|c_\theta\|_{L^2}^2 = \int_0^1 \big(\sin\theta(t)\big)^2 + \big(\cos\theta(t)\big)^2 \,\mathrm{d}t = 1 \tag{3.69}$$

we can assume without loss of generality that $\|s_\theta\|_{L^2}^2 \geq \frac{1}{2}$. By setting

$$\lambda := \frac{\langle s_\theta, c_\theta \rangle_{L^2}}{\|s_\theta\|_{L^2}^2}$$

we obtain

$$\begin{aligned} \|c_\theta - \lambda s_\theta\|_{L^2}^2 &= \|s_\theta\|_{L^2}^2 - 2\lambda \langle c_\theta, s_\theta \rangle_{L^2} + \lambda^2 \|s_\theta\|_{L^2}^2 \\ &= \|s_\theta\|_{L^2}^2 - \frac{|\langle c_\theta, s_\theta \rangle_{L^2}|^2}{\|s_\theta\|_{L^2}^2} = \frac{D_\theta^0}{\|s_\theta\|_{L^2}^2} \leq 2D_\theta^0. \end{aligned} \tag{3.70}$$

Next we introduce $\varphi \in (0, \pi)$ such that

$$\cot\varphi = \lambda.$$

By Cauchy-Schwarz and (3.69) it follows that $|\lambda| \leq 1$ and hence $\varphi \in [\frac{\pi}{4}, \frac{3\pi}{4}]$. Thus on the one hand $\sin\varphi \geq \frac{1}{\sqrt{2}}$ and on the other $\frac{1}{|\sin(\varphi)|} \geq 1$. Therefore by an angle addition theorem for the sinus

$$\begin{aligned} \|c_\theta - \lambda s_\theta\|_{L^2}^2 &= \int_0^1 |\cos\theta(t) - \cot\varphi \sin\theta(t)|^2 \,\mathrm{d}t \\ &= \frac{1}{|\sin\varphi|^2} \int_0^1 |\cos\theta(t)\sin\varphi - \cos\varphi \sin\theta(t)|^2 \,\mathrm{d}t \\ &\geq \int_0^1 |\sin(\varphi - \theta(t))|^2 \,\mathrm{d}t. \end{aligned}$$

Next we introduce a shifted version $\tilde{\theta}\colon [0,1] \to [\varphi, \varphi + 2\pi)$ of θ such that $\tilde{\theta}(t) - \theta(t) \in 2\pi\mathbb{Z}$ for all $t \in [0,1]$. Since $\theta \in \Theta$, i.e. $\theta \in H^1([0,1])$ and $\theta(1) - \theta(0) = 2\pi$ there exist

$t_0 < t_1$ such that

$$\tilde{\theta}(t) \in \left[\varphi + \frac{\pi}{4}, \varphi + \frac{3\pi}{4}\right] \qquad \text{for all } t \in [t_0, t_1]$$

and $\tilde{\theta}(t_0) = \varphi + \frac{\pi}{4}$ and $\tilde{\theta}(t_1) = \varphi + \frac{3\pi}{4}$. It holds that $\tilde{\theta}'(t) = \theta'(t)$ for all $t \in [0, 1]$ and by the fundamental theorem of calculus and Cauchy-Schwarz, it follows that

$$\begin{aligned}\frac{\pi}{2} = \left|\tilde{\theta}(t_1) - \tilde{\theta}(t_0)\right| = \left|\int_{t_0}^{t_1} 1 \cdot \tilde{\theta}'(t)\,\mathrm{d}t\right| &\leq \sqrt{t_1 - t_0}\,\Big(\int_{t_0}^{t_1} |\tilde{\theta}'(t)|^2\,\mathrm{d}t\Big)^{\frac{1}{2}} \\ &\leq \sqrt{t_1 - t_0}\,\|\theta'\|_{L^2}.\end{aligned}$$

By the choice of the interval it holds that $|\sin(\varphi - \tilde{\theta}(t))| \geq 1/\sqrt{2}$ for all $t \in [t_0, t_1]$ and we can use the above to estimate

$$\int_{t_0}^{t_1} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t \geq \frac{t_1 - t_0}{2} \geq \frac{\pi^2}{8\|\theta'\|_{L^2}^2}.$$

Further there exists also $t_2 < t_3$ such that

$$\tilde{\theta}(t) \in \left[\varphi + \frac{5\pi}{4}, \varphi + \frac{7\pi}{4}\right] \qquad \text{for all } t \in [t_2, t_3]$$

and now analogously as above it holds that $t_3 - t_2 \geq \pi^2/(4\|\theta'\|_{L^2}^2)$ and also the corresponding lower estimate for the integral over the interval $[t_2, t_3]$. Now the combination of the last arguments yields

$$\|c_\theta - \lambda s_\theta\|_{L^2}^2 \geq \frac{1}{2}\int_{t_0}^{t_1} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t + \frac{1}{2}\int_{t_2}^{t_3} |\sin(\varphi - \theta(t))|^2\,\mathrm{d}t \geq \frac{\pi^2}{8\|\theta'\|_{L^2}^2}.$$

By putting the above estimates together we obtain the assertion. □

The main result of this section is the following theorem. We prove an alternative local bound on the intrinsic curvature in Θ without using the bound on the second fundamental form in Theorem 3.7.

Theorem 3.17. *Let $\theta \in \Theta$ and $u, v \in \mathcal{T}_\theta\Theta$. Then there is bound $C_\mathcal{R}(\theta)$ depending only on θ (explicit formula in the proof) such that*

$$|\langle \mathcal{R}_\theta(u, v)u, v\rangle_{H^1}| \leq C_\mathcal{R}(\theta)\,\|u\|_{H^1}^2\,\|v\|_{H^1}^2. \tag{3.71}$$

Proof. First we denote $(\mathrm{D}\boldsymbol{P}^0(\theta)\,u)\,v = c_1\,s_\theta + c_2\,c_\theta$ with $c_1, c_2 \in \mathbb{R}$ the values in the formula (3.65). Then the following holds:

$$\|(\mathrm{D}\boldsymbol{P}^0(\theta)\,u)\,v\|_{H^1}^2 = \|c_1\,s_\theta + c_2\,c_\theta\|_{L^2}^2 + \|c_1\,\theta'\,c_\theta - c_2\,\theta'\,s_\theta\|_{L^2}^2$$

$$
\begin{aligned}
&\leq 2\left(|c_1|^2\left(\|s_\theta\|_{L^2}^2+\|\theta' c_\theta\|_{L^2}^2\right)+|c_2|^2\left(\|c_\theta\|_{L^2}^2+\|\theta' s_\theta\|_{L^2}^2\right)\right)\\
&\leq 2\left(1+\frac{1}{(2\pi)^2}\right)\|\theta'\|_{L^2}^2\left(|c_1|^2+|c_2|^2\right).
\end{aligned}
$$

Here we used the Almansi inequality (see [42]),

$$
\|s_\theta\|_{L^2}=\int_0^1|\sin\theta(t)|^2\,\mathrm{d}t\leq\frac{1}{(2\pi)^2}\int_0^1|\theta'(t)\,\cos\theta(t)|^2\,\mathrm{d}t,
$$

which holds since $s_\theta\in H^1_{\mathrm{per}}([0,1])$ and $\int_0^1\sin\theta(t)\,\mathrm{d}t=\langle s_\theta,1\rangle_{L^2}=0$. Clearly it is $\|c_\theta\|_\infty\leq 1$ and so $\|\theta' c_\theta\|_{L^2}^2\leq\|\theta'\|_{L^2}^2$. The analog estimate for c_θ holds true.

We estimate

$$
\begin{aligned}
|\langle c_\theta u,v\rangle_{L^2}|^2=\left(\int_0^1\cos(\theta(t))u(t)v(t)\,\mathrm{d}t\right)^2&\leq\|c_\theta\|_\infty^2\,|\langle u,v\rangle_{L^2}|^2\\
&\leq\|u\|_{L^2}^2\|v\|_{L^2}^2\leq\|u\|_{H^1}^2\|v\|_{H^1}^2
\end{aligned}
$$

and analogously $|\langle s_\theta u,v\rangle_{L^2}|^2\leq\|u\|_{H^1}^2\|v\|_{H^1}^2$. Next we apply this together with $\|c_\theta\|_{L^2}^2\leq\|c_\theta\|_\infty^2\leq 1$ and $\langle c_\theta,s_\theta\rangle_{L^2}\leq\|c_\theta\|_{L^2}\|s_\theta\|_{L^2}\leq 1$ and combine it with (3.68) to get

$$
|c_1|^2=\frac{1}{(D_\theta^0)^2}\left(\|c_\theta\|_{L^2}^2\langle c_\theta u,v\rangle_{L^2}+\langle c_\theta,s_\theta\rangle_{L^2}\langle s_\theta u,v\rangle_{L^2}\right)^2\leq 2\left(\frac{8}{\pi^2}\|\theta'\|_{L^2}^2\right)^2\|u\|_{H^1}^2\,\|v\|_{H^1}^2.
$$

By the same approach, we obtain the estimate

$$
|c_2|^2=\frac{1}{(D_\theta^0)^2}\left(\|s_\theta\|_{L^2}^2\langle s_\theta u,v\rangle_{L^2}+\langle s_\theta,c_\theta\rangle_{L^2}\langle c_\theta u,v\rangle_{L^2}\right)^2\leq 2\left(\frac{8}{\pi^2}\|\theta'\|_{L^2}^2\right)^2\|u\|_{H^1}^2\,\|v\|_{H^1}^2.
$$

Finally, by substituting all of the above estimates into (3.67) and defining

$$
C_{\mathcal{R}}(\theta):=8\left(1+\frac{1}{(2\pi)^2}\right)\left(\frac{8}{\pi^2}\|\theta'\|_{L^2}^2\right)^2\|\theta'\|_{L^2}^2.
$$

the assertion in (3.71) is proved. □

4 Bending energy regularization

You get tragedy where the tree, instead of bending, breaks.
— Ludwig Wittgenstein, *Culture and Value*

This chapter is dedicated to energy functionals for measuring geometric properties. We focus on the bending energy or elasticae, but we take also the Möbius energy into account. Some properties are investigated, especially on shape manifolds as constructed in Chapter 3. Furthermore, as a preparation for the next chapter, we take a closer look at the intrinsic Hessian of bending energy.

Subsequently to this we consider a Tikhonov regularization approach to inverse obstacle problems on shape manifolds penalizing with energy functionals. The fundamental regularizing property is shown and afterwards we prove applicability of the concept to inverse obstacle scattering problems.

4.1 Elastic and Möbius energy

Most parts of this section are published in the article [12].

Recall that the Euler-Bernoulli bending energy (see [14]) of a planar curve $\mathit{\Gamma}$ is given by

$$\int_{\mathit{\Gamma}} \kappa^2 \, \mathrm{d}s,$$

where ds is the line element and κ denotes the (signed) curvature of Γ. The bending energy, or more precisely the curvature, is a geometrical invariant of the curve Γ and thus we gain *independence under reparameterizations*, which is the main benefit of our approach. The bending energy models the stored deformation energy of Γ under the assumption of an undeformed *straight* rest state of the same length as Γ.

Let Γ be parameterized by γ that is represented by $m = (\theta, L, p) \in \mathcal{M}$ as in (3.2). Then we have $\kappa(t) = \frac{\theta'(t)}{L}$ and $\mathrm{d}s = L\,\mathrm{d}t$. This shows that bending energy scales with $1/\lambda$ when Γ is re-scaled by a factor $\lambda > 0$. Thus, without any additional constraints, minimizers of this energy do not exist (the energy of γ_m converges to 0 for $L \to \infty$). We therefore consider the following scale-invariant version $\mathcal{E}_\mathrm{b}\colon \mathcal{M} \to [0,\infty)$ of bending energy which is simply the H^1-seminorm:

$$\mathcal{E}_\mathrm{b}(m) := \int_0^1 \left(\theta'(t)\right)^2 \mathrm{d}t. \tag{4.1}$$

As mentioned above, $\mathcal{E}_\mathrm{b}(m)$ describes the energy required to deform a *straight* elastic rod of length L into Γ. More generally, consider an undeformed rest state Γ_* of non-vanishing curvature (i.e., if Γ_* is pre-curved). Assuming that Γ_* is deformed into Γ by a diffeomorphism $\varphi\colon \Gamma_* \to \Gamma$ that does not change the line element[1], the bending energy is given by

$$\int_{\Gamma_*} \left(\kappa_*(s) - \kappa(\varphi(s))\right)^2 \mathrm{d}s.$$

Representing Γ_* by $m_* = (\theta_*, L, p_*)$ as above, the scale-invariant version of this energy is given by

$$\mathcal{E}_\mathrm{b}(m, m_*) = \int_0^1 \left(\theta'(t) - \theta_*'(t)\right)^2 \mathrm{d}t. \tag{4.2}$$

This formulation is useful when Γ_* represents a reasonable initial guess that is further optimized in order to obtain the desired solution.

While reconstructing a domain, one requires a boundary curve that is free of self-intersections. In this context, the following lemma is useful.

Lemma 4.1. *The set of non-self-intersecting curves is open in the $\mathbb{X}$-topology.*

Proof. First notice that curves of finite bending energy correspond to elements of the Sobolev space $H^2([0,1];\mathbb{R}^2)$. Furthermore, by construction, each point of $\mathcal{M}$ represents a C^1-immersion $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^2$; indeed, due to periodic boundary conditions we can take $\mathbb{S}^1$ as the domain for γ. Since injective immersions of compact domains

[1]Notice that for any two (sufficiently regular) planar curves of the *same* total length L, there exists a diffeomorphism between them that preserves infinitesimal length at every point. In particular, such a mapping is not necessarily a Euclidean motion.

are embeddings, we may employ Theorem 3.10 from [44], stating that the set of C^1-embeddings is open in $C^1(\mathbb{S}^1;\mathbb{R}^2)$. Now, the fact that $H^2(\mathbb{S}^1;\mathbb{R}^2)$ embeds continuously into $C^1(\mathbb{S}^1;\mathbb{R}^2)$ implies the result. □

Remark 4.2. If a sufficiently good initial guess $m_* \in \mathcal{M}$ of the true solution is available and if m_* is free of self-intersections, then Lemma 4.1 ensures that we can choose

$$\mathcal{M}_0 := \{\, m \in \mathcal{M} \mid \|m - m_*\|_{\mathbb{X}} \leq \delta \,\} \tag{4.3}$$

containing only non-self-intersecting curves.

Although we have not encountered the problem of self-intersections in practice for our method, we briefly outline how to avoid this issue whenever needed. A popular and widely studied energy that is *self-avoiding* (i.e., finite energy guarantees that the curve is free of self-intersections) is the so-called *Möbius energy* defined as

$$\mathcal{E}_{\mathrm{M}}(\Gamma) := \int_\Gamma \int_\Gamma \left(\frac{1}{|x-y|^2} - \frac{1}{\mathit{dist}^\Gamma(x,y)^2} \right) \mathrm{d}s(x)\, \mathrm{d}s(y), \tag{4.4}$$

where $\mathit{dist}^\Gamma(x,y)$ denotes the geodesic distance between x and y along the curve Γ as a one dimensional manifold in $\mathbb{R}^2$ and integration is performed with respect to the line elements. This parameterization-invariant energy was introduced by O'Hara [47] and its analytical properties have been studied by several authors [6, 7, 16, 20, 34, 35]. The self-avoiding property is ensured by the first summand of the integrand, while the second summand is introduced in order to remove the singularity along the diagonal $x = y$. The Möbius energy is invariant under Möbius transformations (i.e., under conformal transformations of $\mathbb{C} \cong \mathbb{R}^2$) and thus in particular scale-invariant. We will show in Section 4.4 that using the Möbius energy as an additional penalty term ensures that minimizers of the regularized problem are indeed free of self-intersections.

4.2 Properties of the energy functionals

Most parts of this section are published in the article [12].

The analysis of well-posedness and convergence properties of Tikhonov regularization in Section 4.4 requires some properties of the energy functionals $\mathcal{E}_{\mathrm{b}}$ and $\mathcal{E}_{\mathrm{M}}$ on the Riemannian manifold $\mathcal{M}$. For showing existence of solutions via the direct method of the calculus of variations, weakly sequential lower semi-continuity of the objective functional is a desirable property. Weak convergence, however, is a concept that is *not* invariant under nonlinear changes of coordinates. Since we parameterized

$\mathcal{M}$ as in (3.2), the bending energy becomes a *convex quadratic functional*, enabling us to derive the following result.

Proposition 4.3. *Let* $\mathcal{E} \in \{ \mathcal{E}_\mathrm{b}, \mathcal{E}_\mathrm{b}(\cdot, m_*) \}$. *With respect to the* $\mathbb{X}$*-topology, we have:*

(i) $\mathcal{M} \subset \mathbb{X}$ *is weakly sequentially closed.*

(ii) $\mathcal{E}$ *is weakly sequentially lower semi-continuous.*

(iii) Modulo shifts by elements of $2\,\pi\,\mathbb{Z}$ *, the sublevel sets* $\mathcal{E}^{-1}([0, a]) \subset \mathcal{M}$ *are weakly sequentially compact.*

Proof. We proceed in the usual manner of the direct method of calculus of variations. In order to show (i), consider a sequence $(m_n = (\theta_n, L_n, p_n))_{n\in\mathbb{N}}$ in $\mathcal{M}$ that converges weakly to some $m = (\theta, L, p) \in \mathcal{M}$. By the Rellich compactness theorem, $H^1([0, 1])$ is compactly embedded in $C([0, 1])$ equipped with the supremum norm. Thus, weak convergence of $\theta_n \rightharpoonup \theta$ in $H^1([0, 1])$ implies strong convergence in $C([0, 1])$. Since the closing conditions (3.3) are continuous on $C([0, 1])$, this implies that $\theta \in \Theta$ and thus $m \in \mathcal{M}$.
In order to show (ii), notice that $\mathcal{E}$ is defined in terms of a squared seminorm on $\mathbb{X}$, which is a continuous and convex functional, whose sublevel sets are therefore sequentially closed and convex. The fact that sequentially closed convex sets are weakly sequentially closed implies (ii).
For showing (iii), we first observe that for each $z \in 2\,\pi\,\mathbb{Z}$, the curve represented by $(\theta + z, L, p)$ is the same as the one represented by (θ, L, p). Now let $m_n = (\theta_n, L_n, p_n)$ be a sequence in a sublevel set $\mathcal{E}^{-1}([0, a])$. Modulo shifting by $z_n \in 2\,\pi\,\mathbb{Z}$, we may assume that $\theta_n(0) \in [0, 2\pi]$. We may define an equivalent norm on $H^1([0, 1])$ by $\|\theta\|_* := |\theta(0)| + \|\theta'\|_{L^2}$. We then either have $\|\theta_n\|_{H^1} \leq 2\,\pi + \sqrt{\mathcal{E}(\theta_n)}$ (for the case of $\mathcal{E} = \mathcal{E}_\mathrm{b}$) or $\|\theta_n - \theta_*\|_{H^1} \leq 2\,\pi + |\theta_*(0)| + \sqrt{\mathcal{E}(\theta_n)}$ (for the case of $\mathcal{E} = \mathcal{E}_\mathrm{b}(\cdot, m_*)$). In either case, the sequence $(m_n)_{n\in\mathbb{N}}$ is bounded in H^1, and hence it has a subsequence (θ_{n_k}) converging weakly to some $\theta \in H^1([0, 1])$. Moreover, $[L_1, L_2] \times B$ is compact so that we may find a further subsequence so that m_{n_k} converges weakly to some $m = (\theta, L, p) \in H^1([0, 1]) \times [L_1, L_2] \times B$. Because of (ii), we have $\mathcal{E}(m) \leq a$ and because of (i), m is indeed an element of $\mathcal{M}$. □

Lemma 4.4. *The Möbius energy* $\mathcal{E}_\mathrm{M} \colon \mathcal{M} \to [0, \infty]$ *defined by* (4.4) *is weakly sequentially lower semi-continuous with respect to the weak topology of* $\mathbb{X}$.

Proof. Recall that (3.2) constitutes a smooth mapping from $\mathcal{M}$ to $H^2(\mathbb{S}^1; \mathbb{R}^2)$. As shown in [7], the Möbius energy is continuously differentiable (and thus continuous) on the space of embeddings of class $C^{0,1}(\mathbb{S}^1; \mathbb{R}^2) \cap H^{3/2}(\mathbb{S}^1; \mathbb{R}^2)$. Now the statement follows from the compactness of the embedding of $H^2(\mathbb{S}^1; \mathbb{R}^2)$ into this space. More precisely, let m_n, $m \in \mathcal{M}$ with $m_n \rightharpoonup m$. We have to show that $\mathcal{E}_\mathrm{M}(m) \leq c := \liminf_{n\to\infty} \mathcal{E}_\mathrm{M}(m_n)$. In the case of $c = \infty$, there is nothing to show, so assume that

$c < \infty$. Since $\mathcal{E}_{\mathrm{M}}$ is invariant under scaling and translation, we may assume that $L_n = L = 1$ and $p_n = p = 0$. Denote by γ_n, $\gamma \in H^2(\mathbb{S}^1; \mathbb{R}^2)$ the corresponding parameterizations. Due to the Rellich embedding, we may pick a subsequence such that $c = \lim_{k\to\infty} \mathcal{E}_{\mathrm{M}}(m_{n_k})$ and such that $\gamma_{n_k} \to \gamma$ strongly in $C^{0,1} \cap H^{3/2}$. The latter shows that

$$\mathcal{E}_{\mathrm{M}}(m) = \lim_{k\to\infty} \mathcal{E}_{\mathrm{M}}(m_{n_k}) = c,$$

which proves the claim. □

4.3 Hessian of the bending energy

In this section we investigate the Hessian of the bending energy. Recall that the Hessian of a functional is the covariant derivative of the gradient field, see Chapter A for details. Furthermore we prove that this Hessian is locally Lipschitz continuous.

In the following lemma we compute the first and second derivative of the bending energy (4.2). Recall that the energy for a curve $m = (\theta, L, p)$ is given by

$$\mathcal{E}_{\mathrm{b}}(m, m_*) = \int_0^1 (\theta'(s) - \theta_*'(s))^2 \,\mathrm{d}s.$$

Lemma 4.5. *The first and second derivative of $\mathcal{E}_{\mathrm{b}}(\cdot, m_*)$ at m are given by*

$$\mathrm{D}\mathcal{E}_{\mathrm{b}}(m, m_*)\,(\boldsymbol{v}) = 2\int_0^1 \big(\theta'(s) - \theta_*'(s)\big)\, v'(s)\,\mathrm{d}s, \tag{4.5}$$

$$\mathrm{D}^2\mathcal{E}_{\mathrm{b}}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) = 2\int_0^1 \big(v'(s)\big)^2\,\mathrm{d}s \tag{4.6}$$

with $\boldsymbol{v} = (v, v_L, v_0) \in \mathcal{T}_m\mathcal{M}$ and they are bounded operators on $\mathcal{T}_m\mathcal{M}$ and the bound only depends on θ, i.e. it holds that

$$|\mathrm{D}\mathcal{E}_{\mathrm{b}}(m, m_*)\,(\boldsymbol{v})| \le 2\|\theta' - \theta_*'\|_{L^2}\|\boldsymbol{v}\|_m \tag{4.7}$$

and

$$|\mathrm{D}^2\mathcal{E}_{\mathrm{b}}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v})| \le 2\|\boldsymbol{v}\|_m^2 \tag{4.8}$$

for all $\boldsymbol{v} \in \mathcal{T}_m\mathcal{M}$.

Proof. The formulas for the derivatives as well as the estimates are straightforward to compute. □

The next theorem states an explicit formula for the Hessian of the bending energy.

Theorem 4.6. *Let $m = (\theta, L, p) \in \mathcal{M}$ and $\boldsymbol{v} = (v, v_L, v_0) \in \mathcal{T}_m\mathcal{M}$ be given. Then the intrinsic Hessian of the bending energy at m is given by*

$$\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) = \mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) - \mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*)\,\mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}^2\Phi(\theta)\,(v, v). \tag{4.9}$$

Here the second fundamental form $\mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}^2\Phi(\theta)\,(v, v) = \mathrm{II}^\Theta_\theta\,(v, v)$ as an element of $H^1([0,1])$ (see Theorem 3.7) is canonically embedded into $\mathbb{X}$.

Proof. From basic Riemannian geometry it is known that the intrinsic Hessian of a functional on a submanifold can be expressed using the second fundamental form. The general equation (A.25) is in this case given by

$$\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) = \mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) + \mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*)\;\mathrm{II}^\mathcal{M}_m(\boldsymbol{v}, \boldsymbol{v}).$$

Recall that $\mathcal{M} = \Theta \times [L_1, L_2] \times B$. Here $[L_1, L_2]$ is a trivial submanifold (with boundary) in $\mathbb{R}$. Since we always assume $L \in [L_1, L_2]$ to be an inner point, the curvature vanishes there. One can see this as follows: Since the tangent space at any inner point is isomorphic to $\mathbb{R}$ itself, the orthogonal space is just zero and therefore the second fundamental form as an element in the normal bundle is always zero. For B with nonempty interior with any inner point the argumentation works analogously. Thus we have

$$\mathrm{II}^\mathcal{M}_m(\boldsymbol{v}, \boldsymbol{v}) = (\mathrm{II}^\Theta_\theta(v, v), 0, 0).$$

By applying the results of Theorem 3.7 we obtain the statement. □

By combining Lemma 4.5 and the Theorems 3.7 and 4.6 one can derive upper and lower bounds as follows.

Lemma 4.7. *The bilinear form $\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)$ on $\mathcal{T}_m\mathcal{M}$ is bounded and the bound only depends on θ:*

$$\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) \le 2\big(1 + \|\theta' - \theta'_*\|_{L^2}\,C_\mathrm{II}(\theta)\big)\,\|\boldsymbol{v}\|^2_m \tag{4.10}$$

for all $\boldsymbol{v} \in \mathcal{T}_m\mathcal{M}$. Furthermore we can also bound the operator from below by

$$\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) \ge 2\big(1 - 2\,\|\theta' - \theta'_*\|_{L^2}\,C_\mathrm{II}(\theta)\big)\,\|v'\|^2_\theta \tag{4.11}$$

for all $\boldsymbol{v} \in \mathcal{T}_m\mathcal{M}$.

Proof. Both statements follows directly by applying Cauchy-Schwarz and then using Theorem 3.7. For the second one we also used $\|v\|^2_{L^2} \le \|v'\|^2_{L^2}$. □

In the assumptions of the convergence analysis in the next chapter we need local Lipschitz continuity of the Hessian. As preparation for it the corresponding local

Lipschitz property is proven for the first and second derivatives of the bending energy in the next lemma.

Lemma 4.8. *Let $m = (\theta, L, p)$, $\varrho \geq 0$ and $B_\varrho^{\mathcal{M}}(m)$ be a geodesic ball (in the sense of (3.33)). For $\widehat{m} = (\widehat{\theta}, \widehat{L}, \widehat{p}) \in B_\varrho^{\mathcal{M}}(m)$ denote the parallel transport along the minimizing geodesic connecting m and $\widehat{m}$ by $\mathcal{P}_m^{\widehat{m}}$. Then the first and second derivative of $\mathcal{E}_\mathrm{b}$ are locally Lipschitz continuous with respect to m. That is, for all $m = (\theta, L, p) \in \mathcal{M}$ and sufficiently small $\varrho \geq 0$ it holds that*

$$\left|\mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}) - \mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \leq 2\big(1 + C_{\mathcal{P}}(\theta, \varrho)\big)\; dist(m, \widehat{m})\; \|\boldsymbol{v}\|_m \tag{4.12}$$

and

$$\left|\mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) - \mathrm{D}^2\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \leq 4C_{\mathcal{P}}(\theta, \varrho)\; dist(m, \widehat{m})\; \|\boldsymbol{v}\|_m^2 \tag{4.13}$$

for all $\boldsymbol{v} \in \mathcal{T}_m\mathcal{M}$ and all $\widehat{m} \in B_\varrho^{\mathcal{M}}(m)$, where $C_{\mathcal{P}}(\theta, \varrho)$ is defined in (3.43) in the proof of Lemma 3.9.

Proof. By plugging in the explicit formulas and using standard estimates we get

$$\begin{aligned}
&\left|\mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}) - \mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \\
=&2\left|\langle \theta' - \theta_*', v'\rangle_{L^2} - \langle \widehat{\theta}' - \theta_*', (\mathcal{P}_\theta^{\widehat{\theta}} v)'\rangle_{L^2}\right| \\
\leq&2\left|\langle \theta' - \theta_*' - (\widehat{\theta}' - \theta_*'), v'\rangle_{L^2}\right| + 2\left|\langle \widehat{\theta}' - \theta_*', v' - (\mathcal{P}_\theta^{\widehat{\theta}} v)'\rangle_{L^2}\right| \\
\leq&2\big(\|\theta' - \widehat{\theta}'\|_{L^2}\,\|v'\|_{L^2} + \|\widehat{\theta}' - \theta_*'\|_{L^2}\,\|v' - (\mathcal{P}_\theta^{\widehat{\theta}} v)'\|_{L^2}\big).
\end{aligned}$$

Note that $\|\theta' - \widehat{\theta}'\|_{L^2} \leq dist^{\Theta}(\theta, \widehat{\theta}) \leq dist^{\mathcal{M}}(m, \widehat{m})$, where $dist^{\Theta}$ indicates the Riemannian distance only on the submanifold Θ in $H^1([0,1])$ and $\|v'\|_{L^2} \leq \|v\|_\theta \leq \|\boldsymbol{v}\|_m$. By using the result of Lemma 3.9 we obtain the statement.

Concerning the second derivative we can estimate similarly to the above:

$$\begin{aligned}
&\left|\mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) - \mathrm{D}^2\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \\
=&2\left|\|v'\|_{L^2}^2 - \|(\mathcal{P}_\theta^{\widehat{\theta}} v)'\|_{L^2}^2\right| \\
\leq&2\left|\langle v' - (\mathcal{P}_\theta^{\widehat{\theta}} v)', v'\rangle_{L^2}\right| + 2\left|\langle (\mathcal{P}_\theta^{\widehat{\theta}} v)', v' - (\mathcal{P}_\theta^{\widehat{\theta}} v)'\rangle_{L^2}\right| \\
\leq&2\left(\|v\|_{H^1} + \|\mathcal{P}_\theta^{\widehat{\theta}} v\|_{H^1}\right)\|v - (\mathcal{P}_\theta^{\widehat{\theta}} v)\|_{H^1}.
\end{aligned}$$

Using that $\mathcal{P}_\theta^{\widehat{\theta}}$ is an isometry (see Section A.3) and Lemma 3.9 the bound (4.13) follows as well. ◻

Theorem 4.9. *Let $m = (p, L, \theta) \in \mathcal{M}$, $\varrho \geq 0$ and $\widehat{m} = (\widehat{p}, \widehat{L}, \widehat{\theta}) \in B_\varrho^{\mathcal{M}}(m)$ and $\mathcal{P}_m^{\widehat{m}}$ be the parallel transport along the minimizing geodesic connecting m and $\widehat{m}$. Then the*

Hessian on $\mathcal{M}$ of $\mathcal{E}_\mathrm{b}$ is locally Lipschitz continuous with respect to m. That is, for all $m \in \mathcal{M}$ and sufficiently small $\varrho \geq 0$ it holds that that

$$\left|\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)(\boldsymbol{v}, \boldsymbol{v}) - \mathcal{H}ess\,\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \leq C_{\mathcal{H}ess}(m, \varrho)\, dist(m, \widehat{m})\, \|\boldsymbol{v}\|_m^2 \tag{4.14}$$

with

$$\begin{aligned} C_{\mathcal{H}ess}(m, \varrho) := {} & 2\big(1 + C_{\mathcal{P}}(\theta, \varrho)\big) + 4C_{\mathcal{P}}(\theta, \varrho)\, C_{\mathrm{II}}(\theta) + 2\left(\|\theta' - \theta'_*\|_{L^2} + \varrho\right) \\ & \times \left(C_{\mathrm{D}\Phi^\dagger}(\theta, \varrho) + 2\sqrt{2}\left(C_D(\theta) + p'_D(\|\theta'\|_{L^2} + \varrho)\right) C_{\mathrm{D}2\Phi}(\theta, \varrho)\right) \end{aligned} \tag{4.15}$$

for all $\boldsymbol{v} \in \mathcal{T}_m\mathcal{M}$ and $\widehat{m} \in B_\varrho^{\mathcal{M}}(m)$.

Proof. It holds that

$$\begin{aligned} & \left|\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_*)(\boldsymbol{v}, \boldsymbol{v}) - \mathcal{H}ess\,\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \\ = & \Big|\mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) - \mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*)\,\mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}^2\Phi(\theta)\,(v, v) \\ & \quad - \mathrm{D}^2\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v}) + \mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,\mathrm{D}\Phi(\widehat{\theta})^\dagger\,\mathrm{D}^2\Phi(\widehat{\theta})\,(\mathcal{P}_\theta^{\widehat{\theta}}v, \mathcal{P}_\theta^{\widehat{\theta}}v)\Big| \\ \leq & \left|\mathrm{D}^2\mathcal{E}_\mathrm{b}(m, m_*)\,(\boldsymbol{v}, \boldsymbol{v}) - \mathrm{D}^2\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\,(\mathcal{P}_m^{\widehat{m}}\boldsymbol{v}, \mathcal{P}_m^{\widehat{m}}\boldsymbol{v})\right| \\ & \quad + \left|\left(\mathrm{D}\mathcal{E}_\mathrm{b}(m, m_*) - \mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\mathcal{P}_m^{\widehat{m}}\right)\mathrm{D}\Phi(\theta)^\dagger\,\mathrm{D}^2\Phi(\theta)\,(v, v)\right| \\ & \quad + \left|\mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\mathcal{P}_m^{\widehat{m}}\left(\mathrm{D}\Phi(\theta)^\dagger - \mathrm{D}\Phi(\widehat{\theta})^\dagger\right)\mathrm{D}^2\Phi(\theta)\,(v, v)\right| \\ & \quad + \left|\mathrm{D}\mathcal{E}_\mathrm{b}(\widehat{m}, m_*)\mathcal{P}_m^{\widehat{m}}\,\mathrm{D}\Phi(\widehat{\theta})^\dagger\left(\mathrm{D}^2\Phi(\theta)\,(v, v) - \mathrm{D}^2\Phi(\widehat{\theta})\,(\mathcal{P}_\theta^{\widehat{\theta}}v, \mathcal{P}_\theta^{\widehat{\theta}}v)\right)\right|. \end{aligned}$$

The first term on the right-hand side can directly be estimated by Lemma 4.8 via (4.13). The second term is controlled by the same lemma using (4.12) together with the upper bound for the second fundamental form (3.29). For the third term we use the Lemmas 3.3,3.12 and 4.5 and the fact that $\mathcal{P}_\theta^{\widehat{\theta}}$ is an isometry. To estimate the last expression we apply the Lemmas 3.6, 3.10 and 4.5.

Therefore one obtains the bound

$$\begin{aligned} C := {} & 2\big(1 + C_{\mathcal{P}}(\theta, \varrho)\big) + 4C_{\mathcal{P}}(\theta, \varrho)\, C_{\mathrm{II}}(\theta) \\ & + 2\,\|\widehat{\theta}' - \theta'_*\|_{L^2}\left(C_{\mathrm{D}\Phi^\dagger}(\theta, \varrho) + 2\sqrt{2}\, C_D(\widehat{\theta})\, C_{\mathrm{D}2\Phi}(\theta, \varrho)\right). \end{aligned} \tag{4.16}$$

By (3.52) It holds that it

$$\|\widehat{\theta}' - \theta'_*\|_{L^2} \leq \|\theta' - \theta'_*\|_{L^2} + \|\widehat{\theta}' - \theta'\|_{L^2} \leq \|\theta' - \theta'_*\|_{L^2} + \varrho.$$

By using the notation $p_D(\|\widehat{\theta}'\|_{L^2}) = C_D(\widehat{\theta})$ from the proof of Lemma 3.12 we can use the estimate (3.53) given by

$$C_D(\widehat{\theta}) \le C_D(\theta) + p_D'(\|\theta'\|_{L^2} + \varrho)$$

to obtain the statement. □

4.4 Tikhonov regularization

The main parts of this section are published in the article [12].

In this section we consider a general injective operator

$$F\colon \mathcal{M}_0 \subset \mathcal{M} \to \mathbb{Y}$$

mapping a set of embedded curves $\mathcal{M}_0$ into a Hilbert space $\mathbb{Y}$. The unknown exact solution will be denoted by $m^\dagger \in \mathcal{M}_0$. Noisy data is described by a vector $y^\delta \in \mathbb{Y}$ satisfying

$$\|y^\delta - F(m^\dagger)\|_{\mathbb{Y}} \le \delta.$$

In order to approximately recover $m^\dagger$ from the data y^δ, we use Tikhonov regularization with some regularization parameter $\alpha > 0$:

$$m_\alpha^\delta \in \operatorname*{argmin}_{m\in\mathcal{M}_0} \left[\|F(m) - y^\delta\|_{\mathbb{Y}}^2 + \alpha\,\mathcal{E}_{\mathrm{b}}(m, m_*)\right]. \tag{4.17}$$

Here m_* denotes an initial guess of $m^\dagger$ and may be set to 0 if no such initial guess is available. If no appropriate submanifold of embedded curves $\mathcal{M}_0$ containing the true solution is known, we may choose $\mathcal{M}_0$ by setting $\mathcal{M}_0 := \{\, m \in \mathcal{M} \mid \mathcal{E}_{\mathrm{M}}(m) \le c \,\}$ for sufficiently large $c > 0$ or alternatively consider Tikhonov regularization of the form

$$m_\alpha^\delta \in \operatorname*{argmin}_{m\in\mathcal{M}} \left[\|F(m) - y^\delta\|_{\mathbb{Y}}^2 + \alpha\,\mathcal{E}_{\mathrm{b}}(m, m_*) + \alpha\,\mathcal{E}_{\mathrm{M}}(m)\right]. \tag{4.18}$$

Since $\mathcal{E}_{\mathrm{M}}(m) = \infty$ if m is self-intersecting, $\mathcal{E}_{\mathrm{M}}$ acts as a barrier function: Only the values of F on the set of embedded curves are relevant and each curve $\gamma_{m_\alpha^\delta}$ is guaranteed to be embedded.

With the properties of the energy functionals established in the previous section, the following convergence properties follow from the general theory of nonlinear Tikhonov regularization.

Theorem 4.10. *Assume that $\mathcal{M}_0 \subset \mathcal{M}$ contains only non-self-intersecting elements and let $m^\dagger \in \mathcal{M}_0$. Suppose that $F\colon \mathcal{M}_0 \to \mathbb{Y}$ is weakly sequentially continuous (with respect to the topologies of $\mathbb{X}$ and $\mathbb{Y}$) and injective and $\mathcal{M}_0$ is weakly closed in the case of* (4.17).

1. (existence) *The minimum of the Tikhonov functionals in* (4.17) *and* (4.18) *is attained for any* $\alpha > 0$.

2. (regularizing property) *Suppose that* F *is injective. Moreover, consider a sequence of data* (y^{δ_n}) *with* $\|y^{\delta_n} - F(m^\dagger)\| \leq \delta_n \to 0$ *as* $n \to \infty$*. Assume that the regularization parameters are chosen such that*

$$\alpha_n \to 0 \qquad \text{and} \qquad \frac{\delta_n}{\sqrt{\alpha_n}} \to 0\,.$$

Then for any sequence of minimizers of the Tikhonov functionals we have

$$\lim_{n\to\infty} \left\| m_{\alpha_n}^{\delta_n} - m^\dagger \right\|_{\mathbb{X}} = \lim_{n\to\infty} \left\| \gamma_{m_{\alpha_n}^{\delta_n}} - \gamma_{m^\dagger} \right\|_\infty = 0, \tag{4.19}$$

$$\lim_{n\to\infty} \left\| F\left(m_{\alpha_n}^{\delta_n}\right) - F\left(m^\dagger\right) \right\|_{\mathbb{Y}} = 0. \tag{4.20}$$

3. (convergence rates) *Suppose in the case of* (4.17) *that there exists a loss function* $l \colon \mathcal{M} \times \mathcal{M} \to [0,\infty)$ *and a concave, increasing function* $\varphi \colon [0,\infty) \to [0,\infty)$ *with* $\varphi(0) = 0$ *such that* $m^\dagger$ *satisfies the variational source condition*

$$l(m, m^\dagger) \leq \mathcal{E}_{\mathrm{b}}(m, m_*) - \mathcal{E}_{\mathrm{b}}(m^\dagger, m_*) + \varphi\left(\|F(m) - F(m^\dagger)\|_{\mathbb{Y}}^2\right) \tag{4.21}$$

for all $m \in \mathcal{M}_0$*. Then the reconstruction error for an optimal choice* $\overline{\alpha}$ *of* α *is bounded by*

$$l\left(m_{\overline{\alpha}}^{\delta}, m^\dagger\right) \leq 2\,\varphi(\delta^2). \tag{4.22}$$

Proof. We define a functional $\mathcal{E} \colon \mathbb{X} \to [0,\infty)$ by

$$\mathcal{E}(m) := \begin{cases} \mathcal{E}_{\mathrm{b}}(m, m_*), & \text{if } m \in \mathcal{M}_0, \\ \infty, & \text{else} \end{cases} \qquad \text{or}$$

$$\mathcal{E}(m) := \begin{cases} \mathcal{E}_{\mathrm{b}}(m, m_*) + \mathcal{E}_{\mathrm{M}}(m), & \text{if } m \in \mathcal{M}, \\ \infty, & \text{else} \end{cases}$$

in the case of (4.17) or (4.18), respectively. We show that in both cases $\mathcal{E}$ is weakly sequentially lower semi-compact, i.e. sublevel-sets of $\mathcal{E}$ are weakly sequentially compact. In the first case this follows from Proposition 4.3, part ((iii)) and the assumption that $\mathcal{M}_0$ is weakly sequentially closed. In the second case this is a straightforward consequence of Proposition 4.3, part ((iii)) and Lemma 4.4.

Extending F to an operator $\widetilde{F} \colon \mathbb{X} \to \mathbb{Y}$ in an arbitrary fashion, we can formally write the Tikhonov regularizations (4.17) and (4.18) as a minimization problem over $\mathbb{X}$,

$$m_\alpha^\delta \in \operatorname*{argmin}_{m\in\mathbb{X}} \left[\|\widetilde{F}(m) - y^\delta\|_{\mathbb{Y}}^2 + \alpha\,\mathcal{E}(m) \right].$$

and apply standard convergence results for generalized Tikhonov regularization. The first statement now follows from [53, Thm. 3.22] or [15, Thm. 3.2].

To prove the second statement, let $m^\dagger = (\theta^\dagger, L^\dagger, p^\dagger)$ and $m_{\alpha_n}^{\delta_n} = (\theta_n, L_n, p_n)$ and recall from [53, Thm. 3.26] or [15, Thm. 3.4] that (4.20) holds true, and for an injective operator we have weak convergence of $m_{\alpha_n}^{\delta_n}$ to $m^\dagger$ as well as $\lim_{n\to\infty} \mathcal{E}(m_{\alpha_n}^{\delta_n}) = \mathcal{E}(m^\dagger)$. Since $\mathcal{E}_\mathrm{b}$ and $\mathcal{E}_\mathrm{M}$ are both weakly sequentially lower semicontinuous it follows that $\lim_{n\to\infty} \|\theta_n' - \theta_*'\|_{L^2}^2 = \|(\theta^\dagger - \theta_*)'\|_{L^2}^2$. This implies

$$\begin{aligned}\|(\theta_n - \theta^\dagger)'\|_{L^2}^2 &= \|(\theta_n - \theta_*)'\|_{L^2}^2 - \|(\theta^\dagger - \theta_*)'\|_{L^2}^2 + \langle(\theta^\dagger - \theta_*)', (\theta_n - \theta^\dagger)'\rangle_{L^2} \\ &\to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

Modulo shifts in $2\pi\mathbb{Z}$, we may assume that $\theta_n(0) \in [-\pi, \pi]$. By passing to a subsequence, we may assume that $\theta_{n_k}(0) \to \theta^\dagger(0)$. Using the equivalent norm $\|\theta\|_* := |\theta(0)| + \|\theta'\|_{L^2}$ on $H^1([0,1])$ this yields strong convergence of (θ_{n_k}) to $\theta^\dagger$ in $H^1([0,1])$. As weak convergence in $\mathbb{R}^2$ is equivalent to strong convergence, (L_{n_k}, p_{n_k}) also converges strongly to $(L^\dagger, p^\dagger)$. As this holds true for any subsequence, the whole sequence $(m_{\alpha_n}^{\delta_n})$ converges strongly to $m^\dagger$ in $\mathbb{X}$. This implies strong convergence of the corresponding curves in the supremum norm.

The third statement follows from [18] or [15, Thm. 4.11]. □

We point out that the variational source condition (4.21) is related to stability results as worked out for inverse medium scattering problems in [28] where such conditions with logarithmic functions φ hold true under Sobolev smoothness conditions on the solution. However, for inverse obstacle scattering problems no such verifications of variational source conditions are known so far.

Remark 4.11. It can be seen from the references cited in the proof of Theorem 4.10 that the results can be extended to the case where $\mathbb{Y}$ is a Banach space and $\|F(m) - y^\delta\|_{\mathbb{Y}}^2$ is replaced by more general data fidelity terms $\mathcal{S}(F(m), y^\delta)$.

4.5 Application to inverse obstacle scattering problems

In this section we model the Problem 1.2 on the shape manifold introduced in Chapter 3 and apply the regularization approach using the bending energy discussed in the last section. One may describe the inverse problems as operator equation: We introduce the operator $F\colon \mathcal{M} \to L^2(\mathbb{M})$ mapping $m \in \mathcal{M}$ to the far field pattern u_∞ of the scattered field in Problem 1.1 for the domain Ω corresponding to m. More precisely, the boundary Γ is given by the image of the curve parameterization $\gamma_m(\mathbb{S}^1)$

and Ω is the unbounded component of $\mathbb{R}^2 \setminus \gamma_m(\mathbb{S}^1)$. The inverse problem is described by the operator equation

$$F(m) = u_\infty. \tag{4.23}$$

By Schiffer's uniqueness result ([9, Thm. 5.1]) F is injective if $\mathbb{M} = \mathbb{S}^1 \times \mathbb{S}^1$, and by the uniqueness result of Colton and Sleeman ([9, Thm. 5.1]) it is also injective if $\mathbb{M}$ is the product of $\mathbb{S}^1$ with some finite set and if all curves γ_m for $m \in \mathcal{M}_0$ are contained in a ball of a certain size. (Both results are stated in [9] for $\mathbb{R}^3$, but also hold true in $\mathbb{R}^2$.)

Let us show that the operator F also satisfies the remaining assumptions of Theorem 4.10:

Proposition 4.12. *The operator F maps weakly convergent sequences in $\mathcal{M}_0$ (with respect to the topology of $\mathbb{X}$) to strongly convergent sequences in $L^2(\mathbb{S}^1)$ and is continuously Fréchet differentiable.*
In particular, F is strongly and weakly continuous.

Proof. Notice that the linear mapping $\mathbb{X} \to C^1(\mathbb{S}^1; \mathbb{R}^2)$, $m \mapsto \gamma_m$ defined by (3.2) is compact by embedding theorems for Sobolev spaces, and hence it maps weakly convergent sequences to strongly convergent sequences. Moreover, the forward scattering operator $C^1(\mathbb{S}^1; \mathbb{R}^2) \to L^2(\mathbb{S}^1)$, $\gamma_m \mapsto u_\infty$ is continuously Fréchet differentiable, and in particular continuous by [25, Thm. 1.9]. Therefore, the composition of these two mappings is continuously Fréchet differentiable and maps weakly convergent to strongly convergent sequences. □

Notice that by the last proposition the operator equation (4.23) on an infinite-dimensional manifold $\mathcal{M}_0$ is ill-posed in the sense that there cannot exist a strongly continuous inverse of F. Otherwise every weakly convergent sequence in $\mathcal{M}_0$ would be strongly convergent.) This implies the need for regularization to solve this equation.

It has to be remarked that on the one hand this proves a new geometrically conform regularization method for solving inverse obstacle scattering problems, but on the other hand solving the minimization problem (4.17) is a challenging task. The functional is smooth, but nonconvex and therefore an algorithm to compute (4.17) can easily get stuck in a local minima. This motivates to develop and investigate an algorithm, which can overcome such difficulties.

5 Iteratively regularized Gauss-Newton methods on manifolds

Few ideas work on the first try. Iteration is key to innovation.

— Sebastian Thrun

In the last chapter we introduced a variational regularization approach on shape manifolds based on the bending energy for penalizing. Unfortunately, using this approach one has to solve a smooth, but highly nonconvex optimization problem on a manifold. The numerical minimization is challenging because of local minima. This motivates to investigate an iteratively regularized algorithm in analogy to the iteratively regularized Gauss-Newton algorithm introduced in Section 2.4 for nonlinear operators on Hilbert spaces.

A general framework for such algorithms is introduced in the following. Furthermore, the assumptions on the space, the operator and the regularization term are presented and discussed. Moreover, we prove convergence rates of the algorithm for exact and perturbed data.

Finally the general framework is applied to inverse obstacle scattering problems, using shape manifolds and the bending energy introduced in Chapters 3 and 4. The assumptions made in the general case are discussed and partly verified. Unfortunately, the assumption to control the nonlinearity of the forward operator could not be verified. The analogous assumption in the Hilbert space setting described in Section 2.5 could not yet be proven, either. Nevertheless, this drawback in the theory is discussed in Section 5.5. We emphasize that all assumptions can be verified, which arise newly from the generalization of the algorithm from Hilbert spaces to Riemannian manifolds.

5.1 A general Newton-type algorithm

Let $F\colon \operatorname{dom}(F) \subset \mathcal{M} \to \mathbb{Y}$ be a general Fréchet differentiable operator with injective derivative $\mathrm{D}F(m)$ for all elements $m \in \operatorname{dom}(F)$. Denote an unknown exact solution by $m^\dagger \in \operatorname{dom}(F)$ and let the noisy data $y^\delta \in \mathbb{Y}$ satisfy

$$\|F(m^\dagger) - y^\delta\|_{\mathbb{Y}} \le \delta \tag{5.1}$$

with an error bound $\delta \ge 0$.

Let $\mathcal{E}\colon \mathcal{M} \to \mathbb{R}$ be a C^2 functional incorporating a-priori information. In the following, we will examine an iteratively regularized method given by a sequence $(m_k)_{k\in\mathcal{M}}$. In each Newton-type iteration we solve the minimization problem

$$v_k := \operatorname*{argmin}_{v\in\mathcal{T}_{m_k}\mathcal{M}} \mathcal{J}^{m_k}_{y^\delta,\alpha_k}(v) \tag{5.2a}$$

with

$$\mathcal{J}^{m_k}_{y^\delta,\alpha_k}(v) := \left\|F(m_k)+\mathrm{D}F(m_k)v-y^\delta\right\|^2_{\mathbb{Y}}+\alpha_k\left\langle \mathit{Hess}\,\mathcal{E}(m_k)v, v\right\rangle_{m_k}+\alpha_k\left\langle \mathit{grad}\,\mathcal{E}(m_k), v\right\rangle_{m_k} \tag{5.2b}$$

in order to obtain the update direction $v_k \in \mathcal{T}_{m_k}\mathcal{M}$ and compute the next iterate

$$m_{k+1} := \mathit{exp}_{m_k}(v_k), \tag{5.2c}$$

by the Riemannian exponential map (see Section A.3 for details). Here, $\mathcal{T}_m\mathcal{M}$ denotes the tangent space of $\mathcal{M}$ at the point $m \in \mathcal{M}$, the Riemannian metric is denoted by $\langle\cdot,\cdot\rangle_m$ and the induced norm by $\|\cdot\|_m$.

In addition, we choose a sequence of regularization parameters $(\alpha_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ such that

$$\alpha_k > 0, \qquad 1 \le \frac{\alpha_k}{\alpha_{k+1}} \le C_\alpha, \qquad \lim_{k\to\infty} \alpha_k = 0. \tag{5.3}$$

This is a generalization to infinite-dimensional manifolds of the iteratively regularized Gauss-Newton method (IRGNM), which was first introduced by Bakushinskii [4] in the case for Hilbert spaces. Assuming that the Hessian of $\mathcal{E}$ is positive definite (see Assumption 5.8) the minimizer of (5.2a) is given by

$$\begin{aligned} v_k = \Big(\alpha_k\,\mathit{Hess}\,\mathcal{E}(m_k) + \mathrm{D}F(m_k)^*\,\mathrm{D}F(m_k)\Big)^{-1} \\ \times \Big(\mathrm{D}F(m_k)^*\left(y^\delta - F(m_k)\right) - \alpha_k\,\mathit{grad}\,\mathcal{E}(m_k)\Big). \end{aligned} \tag{5.4}$$

Remark 5.1. In comparison to the iteratively regularized Gauss-Newton method

(2.26) on Hilbert spaces, in the algorithm (5.2) both the data fidelity term and the regularization term is replaced by a Taylor approximation. This arises naturally by the generalization to the manifold setting and will play a role in the following convergence analysis in the Sections 5.3 and 5.4.

5.2 Spectral source conditions

As in Section 2.4 we will prove convergence rates using a-priori knowledge about the solution decoded by source conditions. In this section we generalize the notion of source conditions to manifolds.

We will use the following notation

$$g := \mathit{grad}\,\mathcal{E}(m^\dagger) \in \mathcal{T}_{m^\dagger}\mathcal{M}, \quad \mathcal{H} := \mathcal{H}\mathit{ess}\,\mathcal{E}(m^\dagger)\colon \mathcal{T}_{m^\dagger}\mathcal{M} \to \mathcal{T}_{m^\dagger}\mathcal{M}, \quad \mathcal{W} := \mathcal{H}^{-\frac{1}{2}}$$

and

$$T := \mathrm{D}F(m^\dagger)\colon \mathcal{T}_{m^\dagger}\mathcal{M} \to \mathbb{Y}, \qquad L := T\mathcal{W}.$$

We generalize the notion of source conditions to the context of Riemannian manifolds as follows:

Assumption 5.2. *There is some $v \in \mathcal{T}_{m^\dagger}\mathcal{M}$ such that*

$$g = \mathcal{W}^{-1}\,\varphi(L^*L)\,v \tag{5.5}$$

for an index function φ (in the sense of Definition 2.7).

In [38] it was shown that an index function of the form (5.5) always exists. Recall that in the case of mildly ill-posed problems the index function can be expected to be of Hölder-type

$$\varphi_\nu(\lambda) = \lambda^\nu \tag{5.6}$$

for $\nu > 0$, and in the case of severely ill-posed problems of logarithmic form

$$\varphi_p(\lambda) = \begin{cases} (-\ln\lambda)^{-p}, & 0 < \lambda \le \exp(-1) \\ 0, & \lambda = 0 \end{cases} \tag{5.7}$$

with $p > 0$, which was already discussed in Section 2.3. Here we will assume that $\|T^*T\| \le \exp(-1)$, which can always be achieved by rescaling the norm $\|\cdot\|_\mathbb{Y}$. In general we can assume without loss of generality that

$$\varphi(\lambda) \le 1 \qquad \text{for all} \quad \lambda \in (0, \|L\|^2]. \tag{5.8}$$

Remark 5.3. One can see that (5.5) is a natural choice for generalization into a manifold setting. Consider the special case that the manifold is itself a Hilbert space, i.e. $\mathcal{M} = \mathbb{X}$. Then every tangent space can be identified with the model space. If we use the variable transformation $x \mapsto \widetilde{x} = \mathcal{H}^{-\frac{1}{2}}x$ or the weighted inner product $\langle x, y\rangle_{\mathcal{M}} := \langle x, \mathcal{H}y\rangle_{\mathbb{X}}$ then the source condition (5.5) is equivalent to

$$\widetilde{g} = \varphi(\mathcal{H}^{-\frac{1}{2}}\, T^*\; T\mathcal{H}^{-\frac{1}{2}})\, v,$$

which coincides with the known source conditions for weighted operators over Hilbert spaces (see (2.20) in Theorem 2.11).

In order to prove convergence of the algorithm 5.2 using general source conditions, we apply additional assumptions used in the literature for the index function φ, see [39, 38]. For this, denote the filter functions corresponding to Tikhonov regularization by

$$q_k(\lambda) := \frac{1}{\alpha_k + \lambda} \qquad \text{and} \qquad r_k(\lambda) := 1 - \lambda q_k(\lambda). \tag{5.9}$$

In the following, we only consider filter functions which are covered by the identity, i.e. $\varphi \succcurlyeq \mathrm{id}$ with some constant $C > 0$ as used in the Definition 2.8. In [39] it is shown that in this case there is a constant $c_\varphi > 0$ with $c_\varphi < \max\{1, \frac{1}{C}\}$ such that

$$\sup_{0<\lambda\leq\|L\|^2} |r_k(\lambda)|\, \varphi(\lambda) \leq c_\varphi\, \varphi(\alpha_k), \qquad k \in \mathbb{N}. \tag{5.10a}$$

If, in addition, $\varphi \succcurlyeq \sqrt{\mathrm{id}}$ is satisfied, then it holds that

$$\sup_{0<\lambda\leq\|L\|^2} \sqrt{\lambda}\, |r_k(\lambda)|\, \varphi(\lambda) \leq c_\varphi\, \sqrt{\alpha_k}\, \varphi(\alpha_k), \qquad k \in \mathbb{N}. \tag{5.10b}$$

Provided the index function is concave (which is the case for Hölder index functions with $\nu < 1$ and logarithmic index functions), we get by

$$\varphi(\lambda) = \varphi\!\left(\frac{1}{C_\alpha}\left(C_\alpha\, \lambda\right) + \left(1 - \frac{1}{C_\alpha}\right) 0\right) \geq \frac{1}{C_\alpha}\, \varphi\!\left(C_\alpha\, \lambda\right) + \left(1 - \frac{1}{C_\alpha}\right) \varphi(0)$$

and the fact that $\varphi(0) = 0$ in addition the estimate

$$\frac{\varphi\left(C_\alpha\, \lambda\right)}{\varphi(\lambda)} \leq C_\alpha \qquad \text{for all} \quad \lambda \in (0, \|L\|^2/C_\alpha]. \tag{5.11}$$

As in Chapter 2 we use the a-priori stopping rule (see (2.32): the iteration stops at the first index $\boldsymbol{k} = \boldsymbol{k}(\alpha, \delta)$ such that $\tau\sqrt{\alpha_{\boldsymbol{k}}}\varphi(\alpha_{\boldsymbol{k}})$ is less or equal than δ, for some

$\tau > 0$, i.e.

$$\tau\sqrt{\alpha_{\boldsymbol{k}}}\varphi(\alpha_{\boldsymbol{k}}) < \delta \leq \tau\sqrt{\alpha_k}\varphi(\alpha_k), \qquad 0 \leq k < \boldsymbol{k}. \tag{5.12}$$

5.3 Spatial and nonlinearity assumptions

Being able to change between different tangent spaces is necessary to prove convergence rates. We will use two slightly different maps which encode this behavior: the parallel transport and the derivative of the exponential map (see Section A.3 for details). These maps are closely related and in parts they map identically. From now on, we denote by $\mathcal{P}_m^{\widehat{m}} \colon \mathcal{T}_m\mathcal{M} \to \mathcal{T}_{\widehat{m}}\mathcal{M}$ the parallel transport along the unique geodesic from m to $\widehat{m}$. Uniqueness is guaranteed by restricting to a sufficiently small set, as discussed below. In our setting $\mathcal{P}_m^{\widehat{m}}$ becomes unitary, i.e. the inverses and adjoint mappings satisfy

$$\mathcal{P}_{\widehat{m}}^m = (\mathcal{P}_m^{\widehat{m}})^* = (\mathcal{P}_m^{\widehat{m}})^{-1}.$$

Here, $\mathcal{P}_{\widehat{m}}^m$ transports vectors along the geodesic starting from $\widehat{m}$ to m, the reversed geodesic ([51, Sec. 2]). Furthermore, we use the derivative of the exponential map $\mathrm{D}exp_m(v) \colon \mathcal{T}_m\mathcal{M} \to \mathcal{T}_{\widehat{m}}\mathcal{M}$ with $v := exp_m^{-1}(\widehat{m}) \in \mathcal{T}_m\mathcal{M}$. Note that the derivative of the exponential map and the parallel transport coincide along the direction v:

$$\mathrm{D}exp_m(v)\,(v) = \mathcal{P}_m^{\widehat{m}} v.$$

In section 5.4 we prove convergence of the algorithm (5.2) under the source condition (5.5) and the stopping rule introduced in (5.12). This is done locally in a neighborhood of the exact solution $m^\dagger$. In fact, we consider a ball $B_\varrho(m^\dagger) = \{m \in \mathcal{M} \mid dist(m, m^\dagger) \leq \varrho\}$ of radius $\varrho > 0$ contained in $\mathrm{dom}(F)$. In the following we need the crucial assumption that the curvature of the manifold is bounded in the ball $B_\varrho(m^\dagger)$.

Assumption 5.4. *There is constant $C_{\mathfrak{D}} > 0$ depending on ϱ and $m^\dagger$ such that*

$$\|\mathcal{R}_m(u, v)w\|_m \leq C_{\mathfrak{D}}\,\|u\|_m\,\|v\|_m\,\|w\|_m \tag{5.13}$$

for all $m \in B_\varrho(m^\dagger)$ and $u, v, w \in \mathcal{T}_m\mathcal{M}$. For all $m \in B_\varrho(m^\dagger)$ the map exp_m is a diffeomorphism.

On finite-dimensional manifolds the first part of this assumptions is always satisfied. To see this, note that the ball is compact on finite dimensional manifolds. Therefore the curvature tensor can be bounded uniformly in this ball. In infinite dimension one cannot expect to get a uniform bound and has to verify the assumption locally. The second part of the assumption guarantees that for every two points in the ball $B_\varrho(m^\dagger)$ there exists a unique connecting geodesic. By choosing ϱ small enough we can always guarantee the existence of a minimizing geodesic between two points in the ball.

From literature (see e.g. [30]) we know that the replacement of the data fidelity term by a Taylor approximation implies the need for a condition to treat the nonlinearity of the operator F.

Assumption 5.5. *There are mappings $S(\widehat{m}, m) \in L(\mathbb{Y}, \mathbb{Y})$ and $Q(\widehat{m}, m) \in L(\mathcal{T}_m \mathcal{M}, \mathbb{Y})$ for $\widehat{m}, m \in B_\varrho(m^\dagger)$ and constants $C_S, C_Q > 0$ such that*

$$\mathrm{D}F(\widehat{m})\,\mathrm{D}exp_m(v) = S(\widehat{m}, m)\,\mathrm{D}F(m) + Q(\widehat{m}, m) \tag{5.14a}$$

$$\|I - S(\widehat{m}, m)\|_{\mathbb{Y}} \leq C_S \tag{5.14b}$$

$$\|Q(\widehat{m}, m)\| \leq C_Q \|\mathrm{D}F(m^\dagger)\,\mathrm{D}exp_m(exp_m^{-1}(m^\dagger))\,(v)\|_{\mathbb{Y}} \tag{5.14c}$$

for all $\widehat{m}, m \in B_\varrho(m^\dagger)$ where $v := exp_m^{-1}(\widehat{m})$ or equivalently $exp_m(v) = \widehat{m}$.

Remark 5.6. Note that if $\mathcal{M}$ is a Hilbert space $\mathbb{X}$, then (5.14) reduces to a widely used nonlinearity Assumption 2.12, see also [29, 30]. In the flat space the exponential map is given by the translation operator whose derivative is the identity mapping.

Remark 5.7. The operator $\mathrm{D}exp_m(v)$ is a map from one tangent space to another, which is called a vector transport. The choice of $\mathrm{D}exp$ is not arbitrary and arises from using exp in the algorithm (5.2) as update step. Although we know from differential geometry that the parallel transport map behaves identically to the exponential map in one direction, see (A.13). By applying Assumption 5.4 and Lemma 3.9, the difference of the operators is controlled by a bound on the curvature. Whereas we cannot exchange the choice of vector transport when it is composed with an ill-posed operator.

We illustrate this fact in the following example. Consider a linear, injective, compact and exponentially ill-posed operator $T\colon \mathbb{X} \to \mathbb{Y}$ between Hilbert spaces. Assume there is a linear operator $J\colon \mathbb{X} \to \mathbb{X}$ satisfying $\|J - \mathrm{id}_{\mathbb{X}}\| \leq c$ for c small. In general, there does not exist an operator $A\colon \mathbb{Y} \to \mathbb{Y}$ such that

$$A\,T = T\,J$$

with $\|A - \mathrm{id}_{\mathbb{Y}}\| \leq C$ for a constant $C > 0$. One can see this as follows: Denote with (f_n, g_n, σ_n) the singular system of T (see [13]) and assume that T is exponentially ill-posed in the sense that $\sigma_n = \exp(-n^p)$ for some $p > 0$. Consider the operator

$$J\colon \begin{cases} f_{2n} \mapsto f_{2n} + \frac{1}{2} f_n, \\ f_{2n+1} \mapsto f_{2n+1} + \frac{1}{2} f_n \end{cases}$$

which satisfies $\|J - \mathrm{id}_X\| = \frac{1}{2}$. Assume that a operator A satisfying the above identity exists, then it follows

$$A\colon \begin{cases} g_{2n} \mapsto g_{2n} + \frac{1}{2}\frac{\sigma_n}{\sigma_{2n}} g_n \\ g_{2n+1} \mapsto g_{2n+1} + \frac{1}{2}\frac{\sigma_n}{\sigma_{2n+1}} g_n. \end{cases}$$

As one gets

$$\frac{\sigma_n}{\sigma_{2n}} = \frac{\exp(-n^p)}{\exp(-2^p n^p)} = \exp\left((2^p - 1)\, n^p\right)$$

the right-hand side tends to infinity for $n \to \infty$. Therefore, the operator A is unbounded, which contradicts the assumption.

This shows that one cannot simply replace $\mathrm{D}\mathcal{exp}_m$ with the parallel transport in the decomposition (5.14), just because their difference is small in a small neighborhood.

To ensure the existence of the direction v_k given by (5.2a), we assume that the Hessian of the bending energy is bounded, boundedly invertible and strictly positive definite:

Assumption 5.8. *There are constants $C_{\mathcal{H}}^U, C_{\mathcal{H}}^L > 0$, such that the family of operators $\mathcal{Hess}\,\mathcal{E}(m)\colon \mathcal{T}_m\mathcal{M} \to \mathcal{T}_m\mathcal{M}$ satisfies*

$$\|\mathcal{Hess}\,\mathcal{E}(m)\, v\|_m \le C_{\mathcal{H}}^U \|v\|_m \qquad \text{and} \qquad \langle(\mathcal{Hess}\,\mathcal{E}(m))v, v\rangle_m \ge \frac{1}{C_{\mathcal{H}}^L}\|v\|_m^2 \tag{5.15}$$

for all $m \in B_\varrho(m^\dagger)$ and $v \in \mathcal{T}_m\mathcal{M}$ and define the constant

$$C_{\mathcal{H}} := \max\left\{\sqrt{C_{\mathcal{H}}^U}, \sqrt{C_{\mathcal{H}}^L}\right\}.$$

We will use the notation

$$g_k := \mathcal{grad}\,\mathcal{E}(m_k) \in \mathcal{T}_{m_k}\mathcal{M}, \qquad \mathcal{H}_k := \mathcal{Hess}\,\mathcal{E}(m_k)\colon \mathcal{T}_{m_k}\mathcal{M} \to \mathcal{T}_{m_k}\mathcal{M}$$

and

$$\mathcal{W}_k := \mathcal{H}_k^{-\frac{1}{2}}, \qquad T_k := \mathrm{D}F(m_k)\colon \mathcal{T}_{m_k}\mathcal{M} \to \mathrm{Y}, \qquad L_k := T_k\mathcal{W}_k.$$

Note that the Assumption 5.8 is necessary for defining $\mathcal{W}$ and $\mathcal{W}_k$ and in this case we can rewrite (5.4) as

$$v_k = \mathcal{W}_k \left(\alpha_k I + L_k^* L_k\right)^{-1} \Big(L_k^* \left(y^\delta - F(m_k)\right) - \alpha_k \mathcal{W}_k g_k\Big). \tag{5.16}$$

Likewise, note that the regularization functional $\mathcal{E}$ is nonlinear. Similar to the assumption for the forward operator, we need to impose a condition $\mathcal{E}$ to control the nonlinearity.

Assumption 5.9. *The Hessian of $\mathcal{E}$ is locally Lipschitz continuous, i.e. there is a constant $C_{\mathcal{E}} > 0$ such that*

$$\left\|\mathcal{P}_{\widehat{m}}^{m}\, \mathcal{H}ess\, \mathcal{E}(\widehat{m})\, \mathcal{P}_{m}^{\widehat{m}}\, v - \mathcal{H}ess\, \mathcal{E}(m)\, v\right\|_{m} \leq C_{\mathcal{E}}\, dist(m, \widehat{m})\, \|v\|_{m} \tag{5.17}$$

for all $v \in \mathcal{T}_m \mathcal{M}$ and $m, \widehat{m} \in B_{\varrho}(m^{\dagger})$.

5.4 Convergence analysis on infinite-dimensional manifolds

In the following we will use the notation:

$$e_k^{\dagger} := exp_{m_k}^{-1}(m^{\dagger}) \in \mathcal{T}_{m_k}\mathcal{M}, \qquad e_{\dagger}^{k} := exp_{m^{\dagger}}^{-1}(m_k) \in \mathcal{T}_{m^{\dagger}}\mathcal{M}.$$

Note these two tangent vectors are elements of different tangent spaces and connected by the identity

$$e_{\dagger}^{k} = -\mathcal{P}_{m_k}^{m^{\dagger}} e_k^{\dagger}. \tag{5.18}$$

The main strategy for the proof of local convergence is based on the proof of local convergence in the Hilbert space setting, see Theorem 2.13. In the first four lemmas we establish bounds for the distance between the exact solution and the $(k+1)$-th iterate in terms of the k-th iterate. Using general source conditions this is not sufficient to prove convergence rates, which is discussed for the Hilbert space setting in [30]. We examine the corresponding errors in the observation space as well. From these recursive formulas, we prove convergence rates first for exact data and afterwards for noisy data.

Lemma 5.10. *Suppose the Assumptions 5.4 and 5.8 hold true and $m_k \in B_{\varrho}(m^{\dagger})$. Then*

$$\|e_{\dagger}^{k+1}\|_{m^{\dagger}} = \|e_{k+1}^{\dagger}\|_{m_{k+1}} \leq \left(1 + C_{\mathcal{D}}\, \|e_k^{\dagger}\|_{m_k}^2\right) \|v_k - e_k^{\dagger}\|_{m_k} + \frac{1}{3}\, C_{\mathcal{D}}\, \|v_k - e_k^{\dagger}\|_{m_k}^3. \tag{5.19}$$

Proof. The tangent vectors above measure the distance between the k-th iterate and the exact solution in different tangent spaces, i.e.

$$\|e_k^{\dagger}\|_{m_k} = dist(m_k, exp_{m_k}(e_k^{\dagger})) = dist(m_k, m^{\dagger}) = dist(exp_{m^{\dagger}}(e_{\dagger}^{k}), m^{\dagger}) = \|e_{\dagger}^{k}\|_{m^{\dagger}},$$

due to (A.9). On the other hand, the distance between $m_{k+1} = exp_{m_k}(v_k)$ and $m^\dagger = exp_{m_k}(e_k^\dagger)$ is given by the infimum of the length of all path connecting these two points. Therefore, we choose the path $\gamma(t) := exp_{m_k}(e_k^\dagger + t(v_k - e_k^\dagger))$ and estimate

$$\begin{aligned} dist(m_{k+1}, m^\dagger) &= dist(exp_{m_k}(v_k), exp_{m_k}(e_k^\dagger)) \\ &\leq \int_0^1 \|\mathrm{D}exp_{m_k}(e_k^\dagger + t(v_k - e_k^\dagger))\,(v_k - e_k^\dagger)\|_{\gamma(t)}\,\mathrm{d}t. \end{aligned}$$

Using $w_k := e_k^\dagger + t(v_k - e_k^\dagger)$ we apply (A.18), Assumption 5.4 and the isometry property of parallel transport (see Section A.3) to get

$$\begin{aligned} \|\mathrm{D}exp_{m_k}(w_k)\,(v_k - e_k^\dagger)\|_{\gamma(t)} &\leq \|\mathcal{P}_{m_k}^{\gamma(t)}(v_k - e_k^\dagger)\|_{\gamma(t)} + \frac{1}{2}\|\mathcal{R}_{m_k}(w_k, v_k - e_k^\dagger)\,w_k\|_{m_k} \\ &\leq \|v_k - e_k^\dagger\|_{m_k} + \frac{1}{2}C_{\mathcal{D}}\|v_k - e_k^\dagger\|_{m_k}\,\|w_k\|_{m_k}^2. \end{aligned}$$

Since $(a+b)^2 \leq 2a^2 + 2b^2$, it therefore follows that

$$\begin{aligned} dist(m_{k+1}, m^\dagger) &\leq \|v_k - e_k^\dagger\|_{m_k} + \frac{1}{2}C_{\mathcal{D}}\,\|v_k - e_k^\dagger\|_{m_k}\int_0^1 2\big(\|e_k^\dagger\|_{m_k}^2 + t^2\,\|v_k - e_k^\dagger\|_{m_k}^2\big)\,\mathrm{d}t \\ &\leq \big(1 + C_{\mathcal{D}}\,\|e_k^\dagger\|_{m_k}^2\big)\,\|v_k - e_k^\dagger\|_{m_k} + \frac{1}{3}C_{\mathcal{D}}\,\|v_k - e_k^\dagger\|_{m_k}^3 \end{aligned}$$

and together with the above identity we deduce (5.19). □

Lemma 5.11. *Suppose the Assumptions 5.4 and 5.8 hold true and $m_k \in B_\varrho(m^\dagger)$. Then*

$$\|Te_\dagger^{k+1}\|_{\Upsilon} \leq \frac{1 + \frac{1}{2}C_Q\|v_k - e_k^\dagger\|_{m_k}}{1 - C_S - \frac{1}{2}C_Q\|e_\dagger^{k+1}\|_{m^\dagger}}\Big((1 + C_S)\|T_k(v_k - e_k^\dagger)\|_{\Upsilon} + C_Q\|v_k - e_k^\dagger\|_{m_k}\|Te_\dagger^k\|_{\Upsilon}\Big) \tag{5.20}$$

if $C_S + \frac{1}{2}C_Q\|e_\dagger^{k+1}\|_{m^\dagger} < 1$.

Proof. The proof of (5.20) in analogous to the proof of (5.19). We start by estimating the distance in Υ by two different integrals. Firstly, for the lower bound we use the nonlinearity decomposition (5.14) of F using the path $t \mapsto F(exp_{m^\dagger}(te_\dagger^{k+1}))$:

$$\begin{aligned} &\big\|F(m_{k+1}) - F(m^\dagger)\big\|_{\Upsilon} \\ &= \Big\|\int_0^1 \mathrm{D}F(exp_{m^\dagger}(te_\dagger^{k+1}))\,\mathrm{D}exp_{m^\dagger}(te_\dagger^{k+1})\,(e_\dagger^{k+1})\,\mathrm{d}t\Big\|_{\Upsilon} \\ &= \Big\|\int_0^1 S(exp_{m^\dagger}(te_\dagger^{k+1}), m^\dagger)\,Te_\dagger^{k+1} + Q(exp_{m^\dagger}(te_\dagger^{k+1}), m^\dagger)\,e_\dagger^{k+1}\,\mathrm{d}t\Big\|_{\Upsilon} \\ &\geq \Big(1 - C_S - \frac{1}{2}C_Q\|e_\dagger^{k+1}\|_{m^\dagger}\Big)\,\|Te_\dagger^{k+1}\|_{\Upsilon}. \end{aligned}$$

The positivity of the right-hand side is guaranteed if $C_S + \frac{1}{2}C_Q\|e_\dagger^{k+1}\|_{m^\dagger} < 1$ holds true.

We deduce the upper bound using the path $\gamma(t) \coloneqq exp_{m_k}(w_k(t))$ with $w_k(t) \coloneqq e_k^\dagger + t(v_k - e_k^\dagger)$ by

$$\begin{aligned}
\left\|F(m_{k+1}) - F(m^\dagger)\right\|_Y &\leq \int_0^1 \left\|DF(\gamma(t))\, Dexp_{m_k}(w_k(t))\,(v_k - e_k^\dagger)\right\|_Y \,dt \\
&= \int_0^1 \left\|\left(S(\gamma(t), m_k)\, T_k + Q(\gamma(t), m_k)\right)(v_k - e_k^\dagger)\right\|_Y \,dt \\
&\leq \int_0^1 \left\|S(\gamma(t), m_k)\, T_k(v_k - e_k^\dagger)\right\|_Y + \|Q(\gamma(t), m_k)\,(v_k - e_k^\dagger)\|_Y \,dt \\
&\leq (1 + C_S)\left\|T_k(v_k - e_k^\dagger)\right\|_Y \\
&\qquad + C_Q\|v_k - e_k^\dagger\|_{m_k} \int_0^1 \left\|T\,Dexp_{m_k}(e_k^\dagger)\,(exp_{m_k}^{-1}(\gamma(t)))\right\|_Y \,dt.
\end{aligned}$$

Note that $exp_{m_k}^{-1}(\gamma(t)) = w_k(t)$ and by (A.13), (A.15) and (5.18) we conclude that

$$Dexp_{m_k}(e_k^\dagger)\,(e_k^\dagger) = \mathcal{P}_{m_k}^{m^\dagger} e_k^\dagger = -e_\dagger^k.$$

Using this together with (5.14) we obtain for the last integral

$$\begin{aligned}
&\int_0^1 \|T\,Dexp_{m_k}(e_k^\dagger)\,(w_k(t))\|_Y \,dt \\
&\leq \|Te_\dagger^k\|_Y + \int_0^1 t\,\|T\,Dexp_{m_k}(e_k^\dagger)\,(v_k - e_k^\dagger)\|_Y \,dt \\
&\leq \|Te_\dagger^k\|_Y + \frac{1}{2}\left(\left\|\left(I - S(m^\dagger, m_k)\right) T_k\,(v_k - e_k^\dagger)\right\|_Y + \left\|S(m^\dagger, m_k)\, T_k\,(v_k - e_k^\dagger)\right\|_Y \right.\\
&\qquad\qquad \left. + \left\|Q(m^\dagger, m_k)\,(v_k - e_k^\dagger)\right\|_Y\right) \\
&\leq \|Te_\dagger^k\|_Y + \frac{1}{2}\left((1 + C_S)\,\|T_k(v_k - e_k^\dagger)\|_Y + C_Q\,\|Te_\dagger^k\|_Y\,\|v_k - e_k^\dagger\|_{m_k}\right) \\
&= \frac{1 + C_S}{2}\,\|T_k(v_k - e_k^\dagger)\|_Y + \left(1 + \frac{1}{2}C_Q\|v_k - e_k^\dagger\|_{m_k}\right)\|Te_\dagger^k\|_Y.
\end{aligned}$$

Combining the lower and the upper bound yields the stated inequality (5.20). □

We denote

$$\mathcal{D}_k^\dagger \coloneqq Dexp_{m_k}(e_k^\dagger) \qquad \text{and} \qquad \mathcal{D}_\dagger^k \coloneqq Dexp_{m^\dagger}(e_\dagger^k).$$

Then due to (A.16) and (5.18) the identity $\left(\mathcal{D}_\dagger^k\right)^* = \mathcal{D}_k^\dagger$ holds true. By (A.18) and (5.13), we can derive a bound for these operators:

$$\left\| \mathcal{D}_k^\dagger w \right\|_{m^\dagger} \leq \left(1 + \frac{1}{2} C_{\mathcal{D}} \|e_k^\dagger\|_{m_k}^2\right) \|w\|_{m_k}. \tag{5.21}$$

Concerning the invertibility of these operators note the following: Similarly as above, we can derive the lower bound

$$\left\| \mathcal{D}_k^\dagger w \right\|_{m^\dagger} \geq \left(1 - \frac{1}{2} C_{\mathcal{D}} \|e_k^\dagger\|_{m_k}^2\right) \|w\|_{m_k}$$

for $w \in \mathcal{T}_{m_k}\mathcal{M}$. If $m_k \in B_\varrho(m^\dagger)$ and ϱ are sufficiently small this guarantees invertibility and bounds the norm. The estimates for $\mathcal{D}_\dagger^k$ work analogously.

Recall that (5.9) is given by $q_k(L^*L) = (\alpha_k I + L^*L)^{-1}$ using the functional calculus. Then from (5.16) we obtain

$$\begin{aligned} v_k - e_k^\dagger = {} & \mathcal{W}_k q_k(L_k^* L_k) L_k^* \big(y^\delta - F(m^\dagger)\big) + \mathcal{W}_k q_k(L_k^* L_k) L_k^* \big(F(m^\dagger) - F(m_k) - T_k e_k^\dagger\big) \\ & - \alpha_k \mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k g_k - \alpha_k \mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k \mathcal{H}_k e_k^\dagger. \end{aligned}$$

Furthermore, recalling that $r_k(\lambda) = \alpha_k\, q_k(\lambda)$, we can write

$$\begin{aligned} \mathcal{W}_k r_k(L_k^* L_k) \mathcal{W}_k g_k = {} & \mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g + \mathcal{W}_k r_k(L_k^* L_k) \mathcal{W}_k \big(g_k - \mathcal{D}_\dagger^k g\big) \\ & + \mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_\dagger^k \mathcal{W}^{-1} (\alpha_k I + L^*L) \mathcal{W}^{-1} \big(\mathcal{D}_\dagger^k\big)^{-1} \\ & \qquad - \mathcal{W}_k^{-1} (\alpha_k I + L_k^* L_k) \mathcal{W}_k^{-1}\Big) \mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g \\ = {} & \mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g + \mathcal{W}_k r_k(L_k^* L_k) \mathcal{W}_k (g_k - \mathcal{D}_\dagger^k g) \\ & + \alpha_k \mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_\dagger^k \mathcal{H} \big(\mathcal{D}_\dagger^k\big)^{-1} - \mathcal{H}_k\Big) \mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g \\ & + \mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_\dagger^k T^* T \big(\mathcal{D}_\dagger^k\big)^{-1} - T_k^* T_k\Big) \mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g. \end{aligned}$$

From this decomposition we can distinguish general types of errors: approximation, noise, nonlinearity and Taylor remainder. Both of the latter two are split further into terms arising from the forward operator and the regularization functional. We denote them by

$$v_k - e_k^\dagger = \big(e_{k+1}^\dagger\big)^{\text{app}} + \big(e_{k+1}^\dagger\big)^{\text{noi}} + \big(e_{k+1}^\dagger\big)^{\text{nl-F}} + \big(e_{k+1}^\dagger\big)^{\text{tay-F}} + \big(e_{k+1}^\dagger\big)^{\text{nl-}\mathcal{E}} + \big(e_{k+1}^\dagger\big)^{\text{tay-}\mathcal{E}}, \tag{5.22a}$$

where

$$\big(e_{k+1}^\dagger\big)^{\text{app}} := -\mathcal{D}_\dagger^k \mathcal{W} r_k(L^*L) \mathcal{W} g, \tag{5.22b}$$

$$\big(e_{k+1}^\dagger\big)^{\text{noi}} := \mathcal{W}_k q_k(L_k^* L_k) L_k^* \big(y^\delta - F(m^\dagger)\big), \tag{5.22c}$$

$$\left(e_{k+1}^{\dagger}\right)^{\text{tay-F}} := \mathcal{W}_k q_k(L_k^* L_k) L_k^* \Big(F(m^{\dagger}) - F(m_k) - T_k e_k^{\dagger}\Big), \tag{5.22d}$$

$$\left(e_{k+1}^{\dagger}\right)^{\text{nl-F}} := -\mathcal{W}_k q_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_{\dagger}^{k} T^* T \big(\mathcal{D}_{\dagger}^{k}\big)^{-1} - T_k^* T_k\Big) \mathcal{D}_{\dagger}^{k} \mathcal{W} r_k(L^* L) \mathcal{W} g, \tag{5.22e}$$

$$\left(e_{k+1}^{\dagger}\right)^{\text{tay-}\mathcal{E}} := \mathcal{W}_k r_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_{\dagger}^{k} g - g_k - \mathcal{H}_k e_k^{\dagger}\Big), \tag{5.22f}$$

$$\left(e_{k+1}^{\dagger}\right)^{\text{nl-}\mathcal{E}} := -\mathcal{W}_k r_k(L_k^* L_k) \mathcal{W}_k \Big(\mathcal{D}_{\dagger}^{k} \mathcal{H} \big(\mathcal{D}_{\dagger}^{k}\big)^{-1} - \mathcal{H}_k\Big) \mathcal{D}_{\dagger}^{k} \mathcal{W} r_k(L^* L) \mathcal{W} g. \tag{5.22g}$$

Note that since $v_k - e_k^{\dagger} \in \mathcal{T}_{m_k}\mathcal{M}$ also the error pieces are elements of the tangent space $\mathcal{T}_{m_k}\mathcal{M}$. In general the terms $(e_{k+1}^{\dagger})^{\text{app}}$, $(e_{k+1}^{\dagger})^{\text{noi}}$, $(e_{k+1}^{\dagger})^{\text{tay-F}}$, $(e_{k+1}^{\dagger})^{\text{nl-F}}$, $(e_{k+1}^{\dagger})^{\text{tay-}\mathcal{E}}$ and $(e_{k+1}^{\dagger})^{\text{nl-}\mathcal{E}}$ do not form a decomposition of $e_{k+1}^{\dagger} \in \mathcal{T}_{m_{k+1}}\mathcal{M}$. Although this is true if $\mathcal{M}$ is flat as for example a Hilbert space.

Moreover, we denote

$$\left(e_{\dagger}^{k+1}\right)^{\text{app}} := \mathcal{W} r_k(L^* L) \mathcal{W} g.$$

In the following lemma we derive error estimates using this decomposition.

Lemma 5.12. *Suppose the Assumptions 5.2, 5.4, 5.5, 5.8, 5.9 with $\varphi \geqslant \sqrt{\mathrm{id}}$ hold true and assume (5.1) and $m_k \in B_{\varrho}(m^{\dagger})$. Then we obtain*

$$\left\|\left(e_{k+1}^{\dagger}\right)^{\text{app}}\right\|_{m_k} \leq \left(1 + \frac{1}{2} C_{\mathcal{D}} \left\|e_k^{\dagger}\right\|_m^2\right) \left\|\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{m^{\dagger}}, \tag{5.23a}$$

$$\left\|\left(e_{k+1}^{\dagger}\right)^{\text{noi}}\right\|_{m_k} \leq C_{\mathcal{H}} \frac{1}{2\sqrt{\alpha_k}} \delta, \tag{5.23b}$$

$$\left\|\left(e_{k+1}^{\dagger}\right)^{\text{tay-F}}\right\|_{m_k} \leq C_{\mathcal{H}} \frac{1}{2} \left(\frac{3}{2} C_Q \left\|e_{\dagger}^{k}\right\|_{m^{\dagger}} + 2 C_S\right) \frac{\left\|T e_{\dagger}^{k}\right\|_{\Upsilon}}{\sqrt{\alpha_k}}, \tag{5.23c}$$

$$\begin{aligned} \left\|\left(e_{k+1}^{\dagger}\right)^{\text{nl-F}}\right\|_{m_k} \leq{}& C_{\mathcal{H}} C_S \frac{\left\|T\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{\Upsilon}}{\sqrt{\alpha_k}} \\ &+ C_{\mathcal{H}}^2 C_Q \left(\frac{1}{2} \left\|\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{m^{\dagger}} + \frac{\left\|T\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{\Upsilon}}{\sqrt{\alpha_k}}\right) \frac{\left\|T e_{\dagger}^{k}\right\|_{\Upsilon}}{\sqrt{\alpha_k}}, \end{aligned} \tag{5.23d}$$

$$\left\|\left(e_{k+1}^{\dagger}\right)^{\text{tay-}\mathcal{E}}\right\|_{m_k} \leq \frac{1}{2} C_{\mathcal{H}}^2 \left(C_{\mathcal{E}} + C_{\mathcal{D}} C_{\mathcal{H}} \|v\|_{m^{\dagger}}\right) \left\|e_k^{\dagger}\right\|_{m_k}^2, \tag{5.23e}$$

$$\left\|\left(e_{k+1}^{\dagger}\right)^{\text{nl-}\mathcal{E}}\right\|_{m_k} \leq C_{\mathcal{H}}^2 \left(C_{\mathcal{E}} + C_{\mathcal{H}} C_{\mathcal{D}} \|e_k^{\dagger}\|_{m_k}\right) \left\|\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{m^{\dagger}} \left\|e_k^{\dagger}\right\|_{m_k} \tag{5.23f}$$

and

$$\left\|\left(e_{\dagger}^{k+1}\right)^{\text{app}}\right\|_{m^{\dagger}} \leq C_{\mathcal{H}}\, c_{\varphi} \|v\|_{m^{\dagger}} \varphi(\alpha_k). \tag{5.23g}$$

Proof. In the following we use the standard estimates as already used in Section 2.3

$$\left\|\left(\alpha_k I + L_k^* L_k\right)^{-1}\right\| \leq \frac{1}{\alpha_k} \quad \text{and} \quad \left\|\left(\alpha_k I + L_k^* L_k\right)^{-1} L_k^*\right\| \leq \frac{1}{2\sqrt{\alpha_k}}. \tag{5.24}$$

The estimate (5.23a) follows by (5.21) and (5.23g) from (5.5), (5.10a) and (5.15). The inequality (5.23b) can be derived from (5.1), (5.15) and (5.24).

Now let us study the errors arising from the nonlinearity of F. First the Taylor remainder:

$$
\begin{aligned}
&\left\|T_k e_k^\dagger + F(m_k) - F(m^\dagger)\right\|_{\mathrm{Y}} \\
&\le \int_0^1 \left\|T_k e_k^\dagger + \mathrm{D}F\big(exp_{m^\dagger}(te_\dagger^k)\big)\,\mathrm{D}exp_{m^\dagger}(te_\dagger^k)\,(e_\dagger^k)\right\|_{\mathrm{Y}}\,\mathrm{d}t \\
&= \int_0^1 \left\|\mathrm{D}F\big(exp_{m^\dagger}(te_\dagger^k)\big)\,\mathrm{D}exp_{m^\dagger}(te_\dagger^k)\,(e_\dagger^k) - T_k\,\mathrm{D}exp_{m^\dagger}(e_\dagger^k)\,(e_\dagger^k)\right\|_{\mathrm{Y}}\,\mathrm{d}t \\
&= \int_0^1 \Big\|\Big(S(exp_{m^\dagger}(te_\dagger^k), m^\dagger) - S(m_k, m^\dagger)\Big)\,Te_\dagger^k \\
&\qquad\qquad + Q(exp_{m^\dagger}(te_\dagger^k), m^\dagger)e_\dagger^k - Q(m_k, m^\dagger)e_\dagger^k\Big\|_{\mathrm{Y}}\,\mathrm{d}t \\
&\le 2C_S\left\|Te_\dagger^k\right\|_{\mathrm{Y}} + \int_0^1 \Big(C_Q\left\|T(te_\dagger^k)\right\|_{\mathrm{Y}}\left\|e_\dagger^k\right\|_{m^\dagger} + C_Q\left\|Te_\dagger^k\right\|_{\mathrm{Y}}\left\|e_\dagger^k\right\|_{m^\dagger}\Big)\,\mathrm{d}t \\
&\le \Big(\frac{3}{2}C_Q\left\|e_\dagger^k\right\|_{m^\dagger} + 2C_S\Big)\left\|Te_\dagger^k\right\|_{\mathrm{Y}}.
\end{aligned}
\tag{5.25}
$$

Inequality (5.25) provides the estimate for (5.23c). For the nonlinearity error we use the identity

$$
\begin{aligned}
\mathcal{D}_\dagger^k T^*T - T_k^*T_k\mathcal{D}_\dagger^k &= (T\mathcal{D}_k^\dagger)^*T - T_k^*T_k\mathcal{D}_\dagger^k \\
&= T_k^*(S^*(m^\dagger, m_k) - S(m_k, m^\dagger))T + Q^*(m^\dagger, m_k)T + T_k^*Q(m_k, m^\dagger).
\end{aligned}
$$

By straightforward computations using (5.10) and (5.14) we obtain the estimate (5.23d).

Next we consider the errors arising from the approximation of $\mathcal{E}$. Using (A.18) and (5.13) we get

$$
\|\mathrm{D}exp_{m_k}(e_k^\dagger)\,(w) - \mathcal{P}_{m_k}^{m^\dagger} w\|_{m^\dagger} \le \frac{1}{2}C_{\mathcal{D}}\|e_k^\dagger\|_{m_k}^2\|w\|_{m_k}. \tag{5.26}
$$

Recall that the Hessian is given as the covariant derivative of the gradient field (A.4) and that the covariant derivative can be written as a limit of difference quotients using parallel transport (A.6). Due to this, the fundamental theorem of calculus and (5.17) using the path $t \mapsto exp_{m_k}(te_k^\dagger)$ implies

$$
\begin{aligned}
&\|\mathcal{P}_{m^\dagger}^{m_k} g - g_k - \mathcal{H}_k e_k^\dagger\|_{m_k} \\
&= \left\|\int_0^1 \mathcal{P}_{exp_{m_k}(te_k^\dagger)}^{m_k}\,\mathcal{H}ess\,\mathcal{E}(exp_{m_k}(te_k^\dagger))\,\mathrm{D}exp_{m_k}(te_k^\dagger)\,(e_k^\dagger) - \mathcal{H}_k e_k^\dagger\,\mathrm{d}t\right\|_{m_k} \\
&\le \int_0^1 \left\|\Big(\mathcal{P}_{exp_{m_k}(te_k^\dagger)}^{m_k}\,\mathcal{H}ess\,\mathcal{E}(exp_{m_k}(te_k^\dagger))\,\mathcal{P}_{m_k}^{exp_{m_k}(te_k^\dagger)} - \mathcal{H}_k\Big)\,e_k^\dagger\right\|_{m_k}\,\mathrm{d}t
\end{aligned}
$$

$$\leq C_{\mathcal{E}} \left\|e_k^\dagger\right\|_{m_k} \int_0^1 dist\left(m_k, exp_{m_k}(te_k^\dagger)\right) \mathrm{d}t$$
$$= \frac{1}{2} C_{\mathcal{E}} \left\|e_k^\dagger\right\|_{m_k}^2.$$

Thus, together with (5.5), (5.8) and since $r_k(\lambda) \leq 1$ we derive the estimate (5.23e) for the Taylor remainder arising from the regularization functional. Only (5.23f) is left to show. By the decomposition

$$\left(\mathcal{D}_\dagger^k \mathcal{H} \left(\mathcal{D}_\dagger^k\right)^{-1} - \mathcal{H}_k\right)\mathcal{D}_\dagger^k = \mathcal{P}_{m^\dagger}^{m_k} \mathcal{H} - \mathcal{H}_k \mathcal{P}_{m^\dagger}^{m_k} + \left(\mathcal{D}_\dagger^k - \mathcal{P}_{m^\dagger}^{m_k}\right)\mathcal{H} - \mathcal{H}_k\left(\mathcal{D}_\dagger^k - \mathcal{P}_{m^\dagger}^{m_k}\right)$$

and (5.10a), (5.17) and (5.26) the inequality (5.23f) follows. □

The next lemma provides us error estimates in the observation space under $\mathrm{D}F(m_k)$.

Lemma 5.13. *Let the same assumptions as in Lemma 5.12 hold. Then*

$$\left\|T_k\left(e_{k+1}^\dagger\right)^{\mathrm{app}}\right\|_{\Upsilon} \leq (1 + C_S) \left\|T\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{\Upsilon} + C_Q \left\|\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{m^\dagger} \left\|Te_\dagger^k\right\|_{\Upsilon} \tag{5.27a}$$

$$\left\|T_k\left(e_{k+1}^\dagger\right)^{\mathrm{noi}}\right\|_{\Upsilon} \leq \delta \tag{5.27b}$$

$$\left\|T_k\left(e_{k+1}^\dagger\right)^{\text{tay-F}}\right\|_{\Upsilon} \leq \left(\frac{3}{2} C_Q \left\|e_\dagger^k\right\|_{m^\dagger} + 2C_S\right) \left\|Te_\dagger^k\right\|_{\Upsilon} \tag{5.27c}$$

$$\begin{aligned}\left\|T_k\left(e_{k+1}^\dagger\right)^{\text{nl-F}}\right\|_{\Upsilon} \leq{}& 2C_S \left\|T\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{\Upsilon} \\ &+ C_Q\left(\frac{1}{2} C_{\mathcal{H}} \frac{\left\|T\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{\Upsilon}}{\sqrt{\alpha_k}} + \left\|\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{m^\dagger}\right) \left\|Te_\dagger^k\right\|_{\Upsilon}\end{aligned} \tag{5.27d}$$

$$\left\|T_k\left(e_{k+1}^\dagger\right)^{\text{tay-}\mathcal{E}}\right\|_{\Upsilon} \leq \frac{1}{4} C_{\mathcal{H}} \left(C_{\mathcal{E}} + C_{\mathcal{H}} C_{\mathcal{D}} \|v\|_{m^\dagger}\right) \sqrt{\alpha_k} \|e_k^\dagger\|_{m_k}^2 \tag{5.27e}$$

$$\left\|T_k\left(e_{k+1}^\dagger\right)^{\text{nl-}\mathcal{E}}\right\|_{\Upsilon} \leq \frac{1}{2} C_{\mathcal{H}} \left(C_{\mathcal{E}} + C_{\mathcal{H}} C_{\mathcal{D}} \left\|e_k^\dagger\right\|_{m_k}\right) \sqrt{\alpha_k} \left\|\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{m^\dagger} \left\|e_k^\dagger\right\|_{m_k} \tag{5.27f}$$

and

$$\left\|T\left(e_\dagger^{k+1}\right)^{\mathrm{app}}\right\|_{\Upsilon} \leq c_\varphi \|v\|_{m^\dagger} \sqrt{\alpha_k} \varphi(\alpha_k). \tag{5.28}$$

Proof. (5.28) follows from (5.10b) and (5.5). To prove (5.27a), we apply (5.14). By

$$\left\|L_k \left(\alpha_k I + L_k^* L_k\right)^{-1} L_k^*\right\| \leq 1$$

we have (5.27b). Next we can combine again (5.24) with (5.25) to obtain (5.27c).

For the nonlinearity error with respect to F we apply the same decomposition as above and obtain the estimate (5.27d). The combination of the estimate for the Taylor remainder of the bending energy with (5.26) induces (5.27e). The nonlinearity error from the bending energy can be estimated by (5.10a), (5.17) and (5.26) to conclude (5.27f). □

We combine the Lemmas 5.12 and 5.13 in the next result to sort the upper bounds as preparation for the main theorem.

Lemma 5.14. *Let the assumptions of Lemma 5.12 hold true. Then using the stopping rule (5.12) the inequalities*

$$\left\|v_k - e_k^\dagger\right\|_{m_k} \leq \tilde{a}\,\varphi(\alpha_k) + 2\hat{a}\left\|e_k^\dagger\right\|_{m_k} + \tilde{c}\,\frac{\left\|Te_\dagger^k\right\|_{\mathrm{Y}}}{\sqrt{\alpha_k}}, \tag{5.29}$$

$$\left\|T_k\big(v_k - e_k^\dagger\big)\right\|_{\mathrm{Y}} \leq \overline{a}\,\sqrt{\alpha_k}\varphi(\alpha_k) + \hat{a}\sqrt{\alpha_k}\left\|e_k^\dagger\right\|_{m_k} + \tilde{b}\left\|Te_\dagger^k\right\|_{\mathrm{Y}} \tag{5.30}$$

follow, where

$$\begin{aligned}
\tilde{a} &:= C_{\mathcal{H}}\left(\left(\left(1 + \frac{1}{2}C_{\mathcal{D}}\varrho^2\right) + C_{\mathcal{H}}^2\left(C_{\mathcal{E}} + C_{\mathcal{D}}C_{\mathcal{H}}\varrho\right)\varrho\right) + C_S\right)c_\varphi\|v\|_{m^\dagger} + \frac{1}{2}C_{\mathcal{H}}\tau\\
\hat{a} &:= \frac{1}{4}C_{\mathcal{H}}^2\left(C_{\mathcal{E}} + C_{\mathcal{D}}C_{\mathcal{H}}\|v\|_{m^\dagger}\right)\varrho\\
\tilde{c} &:= \frac{1}{2}C_{\mathcal{H}}\left(\left(\frac{3}{2}C_Q\varrho + 2C_S\right) + C_{\mathcal{H}}C_Q\big(2 + C_{\mathcal{H}}\big)c_\varphi\|v\|_{m^\dagger}\right)\\
\overline{a} &:= \left(\left(1 + 3C_S\right) + \frac{1}{2}C_{\mathcal{H}}^2\big(C_{\mathcal{E}} + C_{\mathcal{H}}C_{\mathcal{D}}\varrho\big)\varrho\right)c_\varphi\|v\|_{m^\dagger} + \tau\\
\tilde{b} &:= 2C_S + \frac{3}{2}C_Q\varrho + \frac{5}{2}C_QC_{\mathcal{H}}c_\varphi\|v\|_{m^\dagger}.
\end{aligned}$$

Proof. We sum up the estimates in Lemma 5.12 and 5.13 and apply (5.10) and (5.12) to obtain the stated formulas. □

Now we are able to state the main result of this chapter and prove convergence rates for exact data.

Theorem 5.15. *Suppose the Assumptions 5.2, 5.4, 5.5, 5.8, 5.9 with concave $\varphi \succcurlyeq \sqrt{\mathrm{id}}$ hold true and assume (5.1). Assume C_S, C_Q, C_α, τ and $\|v\|_{m^\dagger}$ are sufficiently small and α_0 sufficiently large as specified in the proof. Then there exits $\varrho > 0$ such that the iterates m_k, $0 \leq k \leq \boldsymbol{k}$, defined by (5.2) are well defined for every $m_0 \in D(F)$ satisfying*

$$dist(m_0, m^\dagger) \leq \varrho\varphi(\alpha_0), \tag{5.31}$$

if the stopping index $\boldsymbol{k} = \boldsymbol{k}(\delta, y^\delta)$ is given by (5.12). Moreover, it holds that

$$dist(m_k, m^\dagger) \leq \varrho\varphi(\alpha_k), \qquad 0 \leq k \leq \boldsymbol{k}. \tag{5.32}$$

If $\delta = 0$ and $\boldsymbol{k} = \infty$, (5.32) holds true for all $k \in \mathbb{N}$.

Proof. The combination of the Lemmas 5.10 and 5.11 provides the estimate

$$\begin{aligned}\|Te_\dagger^{k+1}\|_\Upsilon &\le \frac{1+\frac{1}{2}C_Q\|v_k-e_k^\dagger\|_{m_k}}{1-C_S-\frac{1}{2}C_Q\|e_\dagger^{k+1}\|_{m^\dagger}}\Big((1+C_S)\|T_k(v_k-e_k^\dagger)\|_\Upsilon + C_Q\|v_k-e_k^\dagger\|_{m_k}\|Te_\dagger^k\|_\Upsilon\Big)\\ &\le \frac{\big(1+\frac{1}{2}C_Q\|v_k-e_k^\dagger\|_{m_k}\big)\Big((1+C_S)\|T_k(v_k-e_k^\dagger)\|_\Upsilon + C_Q\|v_k-e_k^\dagger\|_{m_k}\|Te_\dagger^k\|_\Upsilon\Big)}{1-C_S-\frac{1}{2}C_Q\left(\big(1+C_{\mathcal{D}}\,\|e_k^\dagger\|_{m_k}^2\big)\,\|v_k-e_k^\dagger\|_{m_k}+\frac{1}{3}\,C_{\mathcal{D}}\,\|v_k-e_k^\dagger\|_{m_k}^3\right)}\end{aligned}$$

if one can guarantee that

$$C_S+\frac{1}{2}C_Q\left(\big(1+C_{\mathcal{D}}\,\|e_k^\dagger\|_{m_k}^2\big)\,\|v_k-e_k^\dagger\|_{m_k}+\frac{1}{3}\,C_{\mathcal{D}}\,\|v_k-e_k^\dagger\|_{m_k}^3\right)<1. \tag{5.33}$$

In this case we denote

$$\xi_k := \frac{1+\frac{1}{2}C_Q\|v_k-e_k^\dagger\|_{m_k}}{1-C_S-\frac{1}{2}C_Q\left(\big(1+C_{\mathcal{D}}\,\|e_k^\dagger\|_{m_k}^2\big)\,\|v_k-e_k^\dagger\|_{m_k}+\frac{1}{3}\,C_{\mathcal{D}}\,\|v_k-e_k^\dagger\|_{m_k}^3\right)}$$

and by application of Lemma 5.14 we obtain

$$\begin{aligned}\|Te_\dagger^{k+1}\|_\Upsilon &\le \xi_k\left((1+C_S)\|T_k(v_k-e_k^\dagger)\|_\Upsilon + C_Q\|v_k-e_k^\dagger\|_{m_k}\|Te_\dagger^k\|_\Upsilon\right)\\ &\le \xi_k\,(1+C_S)\left(\bar{a}\,\sqrt{\alpha_k}\varphi(\alpha_k)+\hat{a}\sqrt{\alpha_k}\left\|e_k^\dagger\right\|_{m_k}+\tilde{b}\left\|Te_\dagger^k\right\|_\Upsilon\right)\\ &\quad + \xi_k\,C_Q\left(\tilde{a}\,\varphi(\alpha_k)+2\hat{a}\left\|e_k^\dagger\right\|_{m_k}+\tilde{c}\,\frac{\left\|Te_\dagger^k\right\|_\Upsilon}{\sqrt{\alpha_k}}\right)\|Te_\dagger^k\|_\Upsilon.\end{aligned} \tag{5.34}$$

Using the notation

$$\chi_k := \frac{\|Te_\dagger^k\|}{\Psi(\alpha_k)} \qquad \text{and} \qquad \Psi(\alpha) := \sqrt{\alpha}\varphi(\alpha)$$

we prove by induction that

$$\chi_k \le C_\chi \tag{5.35a}$$

$$dist(m_k, m^\dagger) \le \varrho\varphi(\alpha_k) \tag{5.35b}$$

for $0 \le k \le \boldsymbol{k}$ for some $C_\chi > 0$. We will show in Proposition 5.16 below that there exist positive constants C_χ, a, b, c_χ, C_b satisfying the system of equations

$$C_\chi = \max\left\{\chi_0, \frac{2a}{1-b+\sqrt{(1-b)^2-4ac}}\right\}, \tag{5.36a}$$

$$a = c_\chi(1+C_S)(\bar{a}+\hat{a}\varrho), \tag{5.36b}$$

$$b = c_\chi\big((1 + C_S)\tilde{b} + C_Q(\tilde{a} + 2\hat{a}\varrho)\big), \tag{5.36c}$$

$$c = c_\chi C_Q \tilde{c}, \tag{5.36d}$$

$$c_\chi = \sqrt{C_\alpha} C_\alpha \frac{1 + \frac{1}{2} C_Q C_b}{1 - C_S - \frac{1}{2} C_Q\big((1 + C_{\mathcal{D}}\varrho^2)C_b + \frac{1}{3} C_{\mathcal{D}} C_b^3\big)}, \tag{5.36e}$$

$$C_b = \tilde{a} + 2\hat{a}\varrho + \tilde{c} C_\chi \tag{5.36f}$$

with the constants from Lemma 5.14 as well as the inequalities

$$C_S + \frac{1}{2} C_Q \Big((1 + C_{\mathcal{D}}\varrho^2) C_b + \frac{1}{3} C_{\mathcal{D}} C_b^3 \Big) < 1, \tag{5.37a}$$

$$b + 2\sqrt{ac} < 1, \tag{5.37b}$$

$$\chi_0 \le \frac{1 - b + \sqrt{(1-b)^2 - 4ac}}{2c}, \tag{5.37c}$$

$$\Big(1 + C_{\mathcal{D}}\varrho^2 + \frac{1}{3} C_{\mathcal{D}}\, C_b^2\Big) \left(\tilde{a} + 2\hat{a}\varrho + \tilde{c} C_\chi\right) C_\alpha \le \varrho. \tag{5.37d}$$

By construction (5.35), for $k = 0$, follows from (5.31) and (5.36a). Assume now that (5.35) is true for some $k < \boldsymbol{k}$. Thus the assumptions of the Lemma 5.10 are satisfied and by the choice of C_b it holds that $\|v_k - e_k^\dagger\|_{m_k} \le C_b$. Therefore, (5.33) is satisfied by (5.37a) and $c_\chi > 0$.

From (5.35a) and (5.35b) we have $m_k \in B_\varrho(m^\dagger)$ and by (5.11) it holds that

$$\|e_k^\dagger\| \le \varrho\varphi(\alpha_k), \qquad \frac{\sqrt{\alpha_k}}{\sqrt{\alpha_{k+1}}} \le \sqrt{C_\alpha}, \qquad \frac{\varphi(\alpha_k)}{\varphi(\alpha_{k+1})} \le \frac{\varphi(C_\alpha \alpha_{k+1})}{\varphi(\alpha_{k+1})} \le C_\alpha$$

and therefore with (5.36e) it is

$$\frac{\xi_k}{\Psi(\alpha_{k+1})} \le \frac{c_\chi}{\Psi(\alpha_k)}$$

By combining these with (5.34) and (5.36) we get the recursive error inequality

$$\chi_{k+1} \le a + b\chi_k + c\chi_k^2.$$

Denote by t_1 and t_2 the solutions to $t = a + bt + ct^2$, i.e.

$$t_1 = \frac{2a}{1 - b + \sqrt{(1-b)^2 - 4ac}}, \qquad t_2 = \frac{1 - b + \sqrt{(1-b)^2 - 4ac}}{2c},$$

where the expression for t_1 follows by a binomial identity $(x - y)(x + y) = x^2 - y^2$. By (5.37b) the values t_1 and t_2 satisfy $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$. By construction of C_χ and (5.35a) either $0 \le \chi_k \le t_1$ or $t_1 < \chi_k \le \chi_0$. If $0 \le \chi_k \le t_1$ holds true, then $a, b, c \ge 0$

implies

$$\chi_{k+1} \leq a + b\chi_k + c\chi_k^2 \leq a + bt_1 + ct_1^2 = t_1.$$

In the second case (5.37c) and $a + (b-1)t + ct^2 \leq 0$, for $t_1 \leq t \leq t_2$ implies

$$\chi_{k+1} \leq a + b\chi_k + c\chi_k^2 \leq \chi_k \leq \chi_0.$$

This proves (5.35a) in the induction step. Now (5.19) and (5.29) imply

$$\begin{aligned}
\left\|e_{k+1}^\dagger\right\|_{m_{k+1}} &\leq \left(1 + C_{\mathcal{D}}\left\|e_k^\dagger\right\|_{m_k}^2 + \frac{1}{3}C_{\mathcal{D}}\left\|v_k - e_k^\dagger\right\|_{m_k}^2\right)\left\|v_k - e_k^\dagger\right\|_{m_k} \\
&\leq \left(1 + C_{\mathcal{D}}\varrho^2 + \frac{1}{3}C_{\mathcal{D}}\,C_b^2\right)\left(\tilde{a}\varphi(\alpha_k) + 2\hat{a}\left\|e_k^\dagger\right\|_{m_k} + \tilde{c}\varphi(\alpha_k)\frac{\left\|Te_\dagger^k\right\|_{\Upsilon}}{\sqrt{\alpha_k}\varphi(\alpha_k)}\right) \\
&\leq \left(1 + C_{\mathcal{D}}\varrho^2 + \frac{1}{3}C_{\mathcal{D}}\,C_b^2\right)\left(\tilde{a} + 2\hat{a}\varrho + \tilde{c}C_\chi\right)\frac{\varphi(\alpha_k)}{\varphi(\alpha_{k+1})}\varphi(\alpha_{k+1}) \\
&\leq \left(1 + C_{\mathcal{D}}\varrho^2 + \frac{1}{3}C_{\mathcal{D}}\,C_b^2\right)\left(\tilde{a} + 2\hat{a}\varrho + \tilde{c}C_\chi\right)C_\alpha\varphi(\alpha_{k+1}).
\end{aligned}$$

This together with the smallness condition (5.37d) show (5.35b) and in total (5.32). □

Proposition 5.16. *There are positive constants satisfying (5.36) and (5.37) if one assumes C_S, C_Q, $\|v\|_{m^\dagger}$ and ϱ to be small enough and α_0 to be big enough.*

Proof. We can replace (5.36a) by

$$C_\chi = \frac{2a}{1 - b + \sqrt{(1-b)^2 - 4ac}} \tag{5.38}$$

since χ_0 is smaller than the right-hand side of (5.38) if α_0 is sufficiently large. Therefore we start showing that there are constants solving (5.36) with (5.36a) replaced by (5.38). To prove the existence of such a constant one has to solve the equation with respect to C_χ and show that this solution is a positive real number. Define

$$\begin{aligned}
A &:= \sqrt{C_\alpha}C_\alpha(1 + C_S)(\bar{a} + \hat{a}\varrho), \\
B &:= \sqrt{C_\alpha}C_\alpha\left((1 + C_S)\tilde{b} + C_Q(\tilde{a} + 2\hat{a}\varrho)\right), \\
C &:= \sqrt{C_\alpha}C_\alpha C_Q\tilde{c}
\end{aligned}$$

and functions for the nominator and denominator of the fraction in (5.36e) with C_b replaced by (5.36f)

$$\begin{aligned}
N(C_\chi) &:= 1 + \frac{1}{2}C_Q(\tilde{a} + 2\hat{a}\varrho) + \frac{1}{2}C_Q\tilde{c}C_\chi \\
D(C_\chi) &:= 1 - C_S - \frac{1}{2}C_Q\left(\left(1 + C_{\mathcal{D}}\varrho^2\right)\left(\tilde{a} + 2\hat{a}\varrho + \tilde{c}C_\chi\right) + \frac{1}{3}C_{\mathcal{D}}\left(\tilde{a} + 2\hat{a}\varrho + \tilde{c}C_\chi\right)^3\right).
\end{aligned}$$

We may write D in the form

$$D(C_\chi) = 1 - C_S - \frac{1}{2} C_Q \left(d_0 + d_1 C_\chi + d_2 C_\chi^2 + d_3 C_\chi^3 \right)$$

with coefficients

$$\begin{aligned}
d_0 &:= \left(1 + C_{\mathcal{D}} \varrho^2\right) \left(\tilde{a} + 2\hat{a}\varrho\right) + \frac{1}{3} C_{\mathcal{D}} \left(\tilde{a} + 2\hat{a}\varrho\right)^3, \\
d_1 &:= \left(1 + C_{\mathcal{D}} \varrho^2\right) \tilde{c} + C_{\mathcal{D}} \left(\tilde{a} + 2\hat{a}\varrho\right)^2 \tilde{c}, \\
d_2 &:= C_{\mathcal{D}} \left(\tilde{a} + 2\hat{a}\varrho\right) \tilde{c}^2, \\
d_3 &:= \frac{1}{3} C_{\mathcal{D}} \, \tilde{c}^3.
\end{aligned}$$

Note that d_j, $j = 0, 1, 2, 3$, tend to 0 as C_S, C_Q, $\|v\|_{m^\dagger}$ and ϱ tend to 0. Using this notation, we choose

$$a = A \frac{N(C_\chi)}{D(C_\chi)}, \qquad b = B \frac{N(C_\chi)}{D(C_\chi)}, \qquad c = C \frac{N(C_\chi)}{D(C_\chi)}.$$

Now we can reformulate the equation (5.38) by multiplying with the denominator, substracting $C_\chi(1-b)$ and squaring the equation to obtain

$$C_\chi^2 \left((1-b)^2 - 4ac \right) = 4a^2 - 4a C_\chi (1-b) + C_\chi^2 (1-b)^2. \tag{5.39}$$

Note that by squaring we enlarge the solution set of the equation by allowing the right-hand side to become smaller than zero, i.e. $2a - C_\chi(1-b) < 0$. Plugging in the notation for a and b and multiplying with the denominator function we obtain

$$0 < C_\chi \left(D(C_\chi) - B \, N(C_\chi) \right) - 2A \, N(C_\chi)$$

For the right-hand side we get a polynomial expression such that

$$0 < q_0 + q_1 C_\chi + q_2 C_\chi^2 + q_3 C_\chi^3 + q_4 C_\chi^4$$

with coefficients

$$\begin{aligned}
q_0 &:= -2A\big(1 + \frac{1}{2} C_Q (\tilde{a} + 2\hat{a}\varrho)\big) \\
q_1 &:= 1 - C_S - \frac{1}{2} C_Q \, d_0 - B \left(1 + \frac{1}{2} C_Q (\tilde{a} + 2\hat{a}\varrho)\right) - 2A \frac{1}{2} C_Q \tilde{c} \\
q_2 &:= -\frac{1}{2} C_Q \, d_1 - \frac{1}{2} B \, C_Q \tilde{c} \\
q_3 &:= -\frac{1}{2} C_Q \, d_2
\end{aligned}$$

$$q_4 := -\frac{1}{2} C_Q \, d_3.$$

For C_S, C_Q, $\|v\|_{m^\dagger}$ and ϱ tending to 0 the coefficients q_0, q_2, q_3, q_4 are all negative and tend to zero and on the other hand q_1 is positive and tend to 1. Therefore the only real solutions of the inequality are in the negative axis for choosing the parameters small enough, which consequently cannot be a solution of the whole system.

Back to the equation (5.39), which is equivalent to

$$0 = 4(a^2 + acC_\chi^2) - 4aC_\chi(1 - b).$$

We divide this equation by a, since $a = 0$ if and only if $N(C_\chi) = 0$ and this is only the case for

$$C_\chi = \frac{-2(1 + \frac{1}{2} C_{\mathcal{D}}(\tilde{a} + 2\hat{a}_Q))}{C_{\mathcal{D}}\tilde{c}} < 0.$$

Using the representation of a, b and c by nominator and denominator functions, we multiply equation with $D(C_\chi)/4$ and obtain

$$0 = N(C_\chi)(A + CC_\chi^2) - C_\chi(D(C_\chi) - BN(C_\chi)).$$

As above using the expression for $N(C_\chi)$ and $D(C_\chi)$ we get the polynomial equation

$$0 = p_0 + p_1 C_\chi + p_2 C_\chi^2 + p_3 C_\chi^3 + p_4 C_\chi^4$$

with coefficients

$$\begin{aligned}
p_0 &:= \Big(1 + \frac{1}{2} C_Q \left(\tilde{a} + 2\hat{a}_Q\right)\Big) A \\
p_1 &:= \frac{1}{2} C_Q \tilde{c} A + \Big(1 + \frac{1}{2} C_Q \left(\tilde{a} + 2\hat{a}_Q\right)\Big) B + C_S + \frac{1}{2} C_Q d_0 - 1 \\
p_2 &:= \Big(1 + \frac{1}{2} C_Q \big(\tilde{a} + 2\hat{a}_Q\big)\Big) C + \frac{1}{2} C_Q d_1 + \frac{1}{2} C_Q \tilde{c} B \\
p_3 &:= \frac{1}{2} C_Q \tilde{c} C + \frac{1}{2} C_Q d_2 \\
p_4 &:= \frac{1}{2} C_Q d_3.
\end{aligned}$$

Now if the polynomial on the right-hand side has a root in the positive real numbers, we can choose this point as the constant C_χ.

Note that the coefficients p_0, p_2, p_3 and p_4 are positive can be chosen arbitrarily small by sufficiently decreasing $C_S, C_Q, \|v\|_{m^\dagger}$ and ϱ. In this case the last coefficient p_1 becomes negative and tends to -1. Hence, the polynomial on the right-hand side is monotonically decreasing in a neighborhood of 0 for positive real numbers so that it has a zero if C_S, C_Q, $\|v\|_{m^\dagger}$ and ϱ are chosen sufficiently small.

Now we consider the inequalities (5.37). By choosing C_S, C_Q, $\|v\|_{m^\dagger}$ and ϱ small enough the conditions (5.37a), (5.37b) and (5.37d) follow and selecting α_0 big enough one can satisfy (5.37c). □

In the case of noisy data for a-priori stopping rules: the algorithm has the following convergence rate.

Corollary 5.17. *Let the assumptions of Theorem 5.15 hold true. Then the convergence rate of the algorithm (5.2)*

$$\mathit{dist}(m_k, m^\dagger) < \varrho\, \varphi\left(\Psi^{-1}\left(\frac{\delta}{\tau}\right)\right) \tag{5.40}$$

in the noise level δ follows, where $\Psi(\lambda) = \sqrt{\lambda}\varphi(\lambda)$.

Proof. From (5.12) we obtain the estimate

$$\alpha_k < \Psi^{-1}\left(\frac{\delta}{\tau}\right).$$

From this, (5.32) and the monotonicity of φ the statement follows. □

Remark 5.18. Under the above assumptions with source conditions of the same form in our case (5.5) and in the Hilbert space case (2.27), our iteratively regularized algorithm has the same convergence rates as the iteratively regularized Gauss-Newton method in a Hilbert space setting, as discussed in Theorem 2.13.

5.5 Application to inverse obstacle scattering problems: the BERGN method

In Section 4.5 we proved that the forward operator F defined by the Problem 1.1 is continuous and compact. Furthermore, in [25, Thm. 1.9] it is shown that this operator is differentiable on general non-self-intersecting C^2 curves, which includes the class of shape manifolds $\mathcal{M}$ introduced in Chapter 3. Besides this the injectivity of $\mathrm{D}F(m)$ was proven in [25, Lem. 1.25].

Recall that by Lemma 4.5, the bending energy functional $\mathcal{E}_\mathrm{b}$ is twice continuously differentiable on $\mathcal{M}$. Hence, we can apply the algorithm (5.2) to our setting using $\mathcal{E}(\cdot) := \mathcal{E}_\mathrm{b}(\cdot, m_0)$ for a given initial curve $m_0 \in \mathcal{M}$. In this context we call the algorithm *iteratively bending energy regularized Gauss-Newton method* (BERGN method).

In this section we discuss the assumptions we made for the general framework to conclude that the convergence statements of the last section applies to the BERGN method. In fact, all assumptions arising by the generalization from Hilbert spaces to Riemannian manifolds are verified.

Concerning Assumption 5.2: As already mentioned in Section 2.3 it was proven in [23] that there always exists a source condition. It is known that inverse obstacle scattering problems are exponentially ill-posed and as pointed out in Section 2.5. One observes heuristically that for these problems a source condition of logarithmic type (5.7) is typically satisfied. The weight operator in (5.5) is less important in this context, since we assumed (see discussion below) that $\mathcal{W}$ is boundedly invertible.

The a-priori stopping rule (5.12) we used in the theory has a major drawback in applications: one needs to know the index function from (5.5). One can interpret this as the a-priori knowledge of the smoothness order of $m^\dagger$ (or more precisely of $\mathit{grad}\, \mathcal{E}_{\mathrm{b}}(m^\dagger, m_0)$). Unfortunately, in application the smoothness of the obstacle to be reconstructed is typically unknown. To prevent this disadvantage we may exchange the stopping rule (5.12) by the discrepancy principle

$$\tau\alpha_{\boldsymbol{k}} < \delta \le \tau\alpha_k, \qquad 0 \le k < \boldsymbol{k}. \tag{5.41}$$

Here, no prior knowledge of the index function is needed anymore, which improves the usability of this stopping rule. Assuming that a logarithmic source condition with index function φ is satisfied we obtain finiteness of

$$\gamma_\varphi := \sup_{0<\lambda\le\|L\|^2} \frac{\sqrt{\lambda}}{\varphi(\lambda)} < \infty \tag{5.42}$$

which implies an estimate $\delta \le \tau\gamma_\varphi\Psi(\alpha_k)$. Using this inequality, we can prove results similar to the ones in the previous section. In analogy to Corollary 5.17 we obtain the convergence rate

$$\mathit{dist}(m_{\boldsymbol{k}}, m^\dagger) < \varrho\, \varphi\left(\Psi^{-1}\left(\frac{\delta}{\tau}\right)\right).$$

One key ingredient in the general framework is Assumption 5.4, that one can bound the curvature tensor in a ball $B_\varrho(m^\dagger)$. From Theorem 3.7 we know that the second fundamental form $\mathbb{I}_m^{\mathcal{M}}$ can be bounded by $C_{\mathbb{I}}(\theta)$, which is by construction bounded in $B_\varrho(m^\dagger)$ in terms of ϱ and $\|m^\dagger\|_{H^1}$. Then by applying the Gauss-equation (A.24) one gets a local bound on the curvature tensor. Alternatively we can apply Theorem 3.17, which directly gives a bound of $\mathcal{R}_m$ in $B_\varrho(m^\dagger)$ in terms of ϱ and $\|m^\dagger\|_{H^1}$.

We could not verify Assumption 5.5 dealing with the nonlinearity of F by a decomposition of DF. Even in the case of star-shaped obstacles and F operating on Hilbert spaces (compare Section 2.5) the original nonlinearity assumption could not be verified yet. This is discussed in [24, 25] for inverse obstacle scattering problems. In the case of nonlinear ill-posed operators on Hilbert spaces there are many interesting forward mappings arising from applications for which there is no verification of such an assumption available. But there are also forward operator, which satisfies a nonlinearity condition. A list of examples can be found in [27, Ex. 2.8]. At least empirically one can observe the theoretical convergence rates under the nonlinearity assumption in simulations, see [24]. Even it could not yet be verified, this kind of assumption is widely used in the literature and accepted by the community.

The local bound on $\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_0)$ on $B_\varrho(m^\dagger)$ in Assumption 5.8 is satisfied by Lemma 4.7. Moreover, if one chooses ϱ small enough and $m_0 = (\theta_0, L_0, p_0)$ close enough to $m^\dagger$ such that

$$2\,\|\theta' - \theta_0'\|_{L^2}\, C_{\mathrm{I\!I}}(\theta) < 1 \tag{5.43}$$

for all $m = (\theta, L, p) \in B_\varrho(m^\dagger)$, then the operator $\mathcal{H}ess\,\mathcal{E}_\mathrm{b}(m, m_0)$ is positive definite on $B_\varrho(m^\dagger)$ by the second part of Lemma 4.7. If this holds, Assumption 5.8 is satisfied.

Assumption 5.9 to controlling the nonlinearity of the regularization term follows from Theorem 4.9. Here the function $C_{\mathcal{H}ess}(m)$ is bounded in $B_\varrho(m^\dagger)$ in terms of ϱ and $\|m^\dagger\|_{H^1}$ following from its definition in the theorem.

Summarizing this discussion, all assumptions are verified, which arise new in our framework on shape manifolds with bending-energy-based regularization in comparison to standard concepts in Hilbert spaces.

6 Implementation and numerical results

Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation, and eventually build a structure which has no relation to reality.

— Nikola Tesla

In the first part of this chapter we establish a discrete version of our setting introduced in Chapter 3 and explain how one can find minimizers to the Tikhonov functional proposed in Chapter 4. By small changes one obtains directly the discrete version of the BERGN method introduced in Chapter 5.

We demonstrate on simulations that the geometrical approach for solving inverse obstacle scattering problems using the bending energy functional works for several examples of shapes. The two major benefits of our concept in comparison with the star-shape approach (see Section 2.5), i.e. the independence of the parametrization and the extension to nonstar-shape obstacles, are emphasized by examples. Furthermore we highlight the convergence behavior of our algorithm with on the one hand exact data and on the other hand different noise levels.

6.1 Discrete setting

The main part of this section is published in the article [12].

In order to treat bending energy computationally, we represent closed curves by closed polygons. To this end, consider an arbitrary (but fixed) partition $(0 = \tau_0 < \tau_1 < \cdots < \tau_n = 1)$ of the unit interval and let the angle variable be given by a *piecewise constant* function represented by a vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$, i.e. $\theta(t) = \theta_j$ for $t \in (\theta_{j-1}, \theta_j]$. In perfect analogy to (3.2), we then define a polygon of length L by

$$\gamma(t) \coloneqq p + L \int_0^t \big(\cos(\theta(\tau)), \sin(\theta(\tau))\big)\, \mathrm{d}\tau. \tag{6.1}$$

Analogously to the smooth case, in order to fulfill the closing conditions (3.1), $\boldsymbol{\theta}$ needs to satisfy

$$\Phi(\boldsymbol{\theta}) = 0, \quad \text{where} \quad \Phi(\boldsymbol{\theta}) = \int_0^1 \big(\cos(\theta(t)), \sin(\theta(t))\big)\, \mathrm{d}t. \tag{6.2}$$

Define the *turning angles* by $[\boldsymbol{\theta}]_i \coloneqq (\theta_{i+1} - \theta_i)$, where indices are taken modulo n and $[\boldsymbol{\theta}]_i$ is shifted such that $[\boldsymbol{\theta}]_i \in (-\pi, \pi]$ for all i. The number $\left(\sum_i [\boldsymbol{\theta}]_i\right)/2\pi$ is known as the *discrete turning number* of γ.

Let $\Theta_n \coloneqq \{\boldsymbol{\theta} \in \mathbb{R}^n \mid \Phi(\boldsymbol{\theta}) = 0\}$ and define the space of discrete curves by

$$\mathcal{M}_n \coloneqq \Theta_n \times [L_1, L_2] \times B \quad \subset \quad \mathbb{X}_n \coloneqq \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^2,$$

for a compact, convex set of base points $B \subset \mathbb{R}^2$ and minimal and maximal acceptable curve lengths $0 < L_1 \leq L_2 < \infty$. On this space, the scale-invariant version of discrete bending energy for a curve $\boldsymbol{m} \in \mathcal{M}_n$ is readily defined as

$$\mathcal{E}_{\mathrm{b,n}}(\boldsymbol{m}) \coloneqq \sum_{i=1}^n \left(\frac{[\boldsymbol{\theta}]_i}{h_i}\right)^2 h_i = \sum_{i=1}^n \frac{([\boldsymbol{\theta}]_i)^2}{h_i}, \tag{6.3}$$

see, e.g. [21]. Here the *dual edge lengths* are given by $h_i \coloneqq (\tau_{i+1} - \tau_{i-1})/2$ for $i \in \{1, \ldots, n\}$, where we set $\tau_{n+1} = 1 + \tau_1$. This expression provides the natural analogue[1] of the smooth version (4.1). It goes back to the work of Hencky in his 1921 PhD thesis [21] and is in the spirit of discontinuous Galerkin (DG) methods (see [2]). A completely analogous discrete version of this energy can be defined for *open* polygons. In this case, for clamped boundary conditions and under the constraint of fixed total curve length, the set of minimizers of this discrete energy converges in Hausdorff distance to the corresponding set of smooth minimizers under mesh refinement, see [55]. More specifically, the angle variables converge in L^∞ and in $W^{1,p}$ for $p \in [2, \infty)$ under a suitable smoothing operator for the angle variables. Finally, a

[1]Notice that discrete bending energy corresponds to its smooth counterpart in the sense that turning angles at vertices correspond to curvatures *integrated* over dual edges, i.e., $[\boldsymbol{\theta}]_i \cong \int_{(\tau_i+\tau_{i-1})/2}^{(\tau_{i+1}+\tau_i)/2} \kappa(s)\, \mathrm{d}s$. This perspective naturally leads to formulation (6.3).

discrete analogue $\mathcal{E}_{\mathrm{b,n}}(\boldsymbol{m}, \boldsymbol{m}_*)$ of the smooth pre-curved energy $\mathcal{E}_{\mathrm{b}}(m, m_*)$ is readily obtained by replacing $[\boldsymbol{\theta}]$ by $([\boldsymbol{\theta}] - [\boldsymbol{\theta}]_*)$ in (6.3).

For convenience, we briefly sketch here the implementation of an algorithm arising from Tikhonov regularization (see Section 4.5), which will turn out to be applicable to the BERGN method (see Section 5.5). The regularized functional that we seek to minimize on the space $\mathcal{M}_n \subset \mathbb{X}_n$ is of the form

$$\mathcal{J}^{\alpha} \colon \mathcal{M}_n \to \mathbb{R}, \quad \boldsymbol{m} \mapsto \tfrac{1}{2}\left\|F_n(\boldsymbol{m}) - \boldsymbol{y}^{\delta}\right\|_{\mathbb{Y}_n}^2 + \alpha\, \mathcal{E}_n(\boldsymbol{m}). \tag{6.4}$$

Here, $F_n : \mathcal{M}_n \to \mathbb{Y}_n$ is some discretization for polygonal closed curves of the forward operator F, the term $\boldsymbol{y}^{\delta} \in \mathbb{Y}_n$ represents the measured data in some finite dimensional Euclidean space $\mathbb{Y}_n$, the scalar $\alpha \geq 0$ is the regularization parameter, and $\mathcal{E}_n = \mathcal{E}_{\mathrm{b,n}}$ or $\mathcal{E}_n = \mathcal{E}_{\mathrm{b,n}} + \mathcal{E}_{\mathrm{M,n}}$ with a discrete approximation $\mathcal{E}_{\mathrm{M,n}}$ of the Möbius energy $\mathcal{E}_{\mathrm{M}}$.

Remark 6.1. We skip the requisite details on the definition of $\mathcal{E}_{\mathrm{M,n}}$ since our numerical experiments show that in practice the tracking term $\frac{1}{2}\left\|F_n(\boldsymbol{m}) - \boldsymbol{y}^{\delta}\right\|_{\mathbb{Y}_n}^2$ (see (6.4) below) is sufficient to prevent iterates from developing self-intersections. Notwithstanding, for details on discrete Möbius energy, see [34, 35], and for Γ-convergence to the smooth case see [54].

The discrete nonlinear Tikhonov regularization on $\mathcal{M}_n$ may then be written as the following constrained minimization problem:

$$\text{Minimize } \ \mathcal{J}^{\alpha}(\boldsymbol{m}) \ \text{ subject to } \ \Phi(\boldsymbol{m}) = 0 \ \text{ and } \ (L, p) \in [L_1, L_2] \times B. \tag{6.5}$$

We will ignore the inequality constraints $(L, p) \in [L_1, L_2] \times B$ for simplicity, although it would not be difficult to include them. In particular, these constraints never became active in our numerical experiments. We only require these constraints for the theoretical analysis in Section 4.4.

Since F_n does not have a natural extension outside the discrete shape space $\mathcal{M}_n = \{\boldsymbol{m} \mid \Phi(\boldsymbol{m}) = 0\}$, standard methods of constrained nonlinear programming are not applicable. When using iterative methods for minimizing $\mathcal{J}^{\alpha}$, we require an *intrinsic* stepping method on the constraint manifold $\mathcal{M}_n$ in order to supply the forward operator F_n with meaningful input. Prominent examples of such methods are intrinsic Newton-type algorithms on Riemannian manifolds, see, e.g. [51]. In such methods, one determines the update direction $\boldsymbol{u} \in \mathbb{X}_n$ by solving a saddle point system of the form

$$\begin{pmatrix} H(\boldsymbol{m}) & D\Phi^{\top}(\boldsymbol{m}) \\ D\Phi(\boldsymbol{m}) & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \mu \end{pmatrix} = \begin{pmatrix} -D\mathcal{J}^{\alpha}(\boldsymbol{m}) \\ 0 \end{pmatrix}, \tag{6.6}$$

where H is (a surrogate for) the Hessian of the objective functional, the manifold $\mathcal{M}_n$ is given by the constraint equations (6.2), which we encode by a function $\Phi : \mathbb{X}_n \to \mathbb{R}^2$, and $\mu \in \mathbb{R}^2$ denotes a Lagrange multiplier. The resulting linear systems have roughly the size $n \times n$ and can be solved using a direct solver. In our implementation, we usually use $n = 150$.

A first example is the full intrinsic Hessian, which can be obtained from the Lagrange function $\mathcal{L}(\boldsymbol{m}, \lambda) := \mathcal{J}^\alpha(\boldsymbol{m}) + \lambda^\top \Phi(\boldsymbol{m})$ of (6.5) as

$$H(\boldsymbol{m}) = D^2_{\boldsymbol{m}} \mathcal{L}(\boldsymbol{m}, \lambda_{\boldsymbol{m}}) = D^2 \mathcal{J}^\alpha(\boldsymbol{m}) + \lambda_{\boldsymbol{m}}^\top D^2 \Phi(\boldsymbol{m}). \tag{6.7}$$

The requisite Lagrange multiplier $\lambda_{\boldsymbol{m}}^\top \in \mathbb{R}^2$ is obtained by multiplying the equation $D_{\boldsymbol{m}} \mathcal{L}(\boldsymbol{m}, \lambda) = 0$ by $D\Phi^\dagger(\boldsymbol{m})$ from the right, i.e.,

$$\lambda_{\boldsymbol{m}}^\top = -D\mathcal{J}^\alpha(\boldsymbol{m})\, D\Phi^\dagger(\boldsymbol{m}).$$

Here $D\Phi^\dagger(\boldsymbol{m})$ denotes the Moore-Penrose inverse with respect to a finite difference approximation of the H^1-inner product.

Notice that assembling the system with the full intrinsic Hessian contains a contribution of the form $\langle F_n(\boldsymbol{m}) - \boldsymbol{y}^\delta, D^2 F_n(\boldsymbol{m})(\cdot,\cdot)\rangle_{\mathbb{Y}_n}$, which is dense and costly to compute. We therefore use a Gauß-Newton inspired surrogate, which is given in bilinear form as[2]

$$H(\boldsymbol{m}) = \langle DF_n(\boldsymbol{m})\,\cdot\,, DF_n(\boldsymbol{m})\,\cdot\,\rangle_{\mathbb{Y}_n} + \alpha \mathcal{H}ess\, \mathcal{E}_n(\boldsymbol{m}), \tag{6.8}$$

where we identify matrices with bilinear forms and where the *intrinsic* energy Hessian has the form

$$\mathcal{H}ess\, \mathcal{E}_n(\boldsymbol{m}) = D^2 \mathcal{E}_n(\boldsymbol{m}) - D\mathcal{E}_n(\boldsymbol{m})\, D\Phi^\dagger(\boldsymbol{m}) D^2 \Phi(\boldsymbol{m}). \tag{6.9}$$

Notice that the second term on the right-hand side of this equation arises from the second term on the right-hand side of (6.7). In the language of differential geometry (see Section A.5), the term $D\Phi^\dagger(\boldsymbol{m})\, D^2\Phi(\boldsymbol{m})$ encodes the *second fundamental form* of the discrete constraint manifold. In the continuous case we computed this explicitly and proved local bounds, see Section 3.2. The quantities on the right-hand side of (6.9) are easy to assemble for $\mathcal{E}_n = \mathcal{E}_{\mathrm{b,n}}$ due to the quadratic nature of $\mathcal{E}_{\mathrm{b,n}}$.

Another attractive alternative is to use

$$H(\boldsymbol{m}) = \langle DF_n(\boldsymbol{m})\,\cdot\,, DF_n(\boldsymbol{m})\,\cdot\,\rangle_{\mathbb{Y}_n} + \alpha\, \langle \cdot, \cdot \rangle_{\mathbb{X}}.$$

This way, $H(\boldsymbol{m})$ is always positive definite on the null space of $D\Phi(\boldsymbol{m})$ and the saddle-point matrix from (6.6) is guaranteed to be continuously invertible. Thus, in this case, the method boils down to a gradient descent in the manifold $\mathcal{M}_n$ with respect to the

[2]Notice that in this formulation we have also dropped the additional term of the form $\langle F_n(\boldsymbol{m}) - \boldsymbol{y}^\delta, DF_n(\boldsymbol{m})\, D\Phi^\dagger(\boldsymbol{m})\, D^2\Phi(\boldsymbol{m})\rangle_{\mathbb{Y}_n}$ since it does not lead to improved convergence rates.

Riemannian metric induced by H.

Once an update direction $\boldsymbol{u}$ has been computed in the above fashion, the next iterate is found by first setting $\boldsymbol{x}_0 = \boldsymbol{m} + t\,\boldsymbol{u}$ for some small $t > 0$. Restoring feasibility (i.e., ensuring that the next iterate resides on the constraint manifold) is then achieved by iterating

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - D\Phi^{\dagger}(\boldsymbol{x}_k)\,\Phi(\boldsymbol{x}_k), \tag{6.10}$$

until $\Phi(\boldsymbol{x}_k)$ is sufficiently small.[3] The step size t can be determined by a standard backtracking line search, while the matrix-vector product $\tilde{\boldsymbol{u}} = D\Phi^{\dagger}(\boldsymbol{x})\,\tilde{\boldsymbol{v}}$ is computed by solving the saddle point problem

$$\begin{pmatrix} G_{\mathbb{X}_n} & D\Phi^{\top}(\boldsymbol{x}) \\ D\Phi(\boldsymbol{x}) & 0 \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{u}} \\ \tilde{\mu} \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{\boldsymbol{v}} \end{pmatrix}.$$

Here $G_{\mathbb{X}_n}$ is the Gram matrix of the discrete H^1-inner product on $\mathbb{X}_n$, the upper left $n \times n$ block of which is a finite-difference Laplacian. Analogously, $D\mathcal{E}_{\mathrm{b,n}}(\boldsymbol{m})\,D\Phi^{\dagger}(\boldsymbol{m}) = (D\Phi^{\dagger}(\boldsymbol{m}))^{\top} D\mathcal{E}_{\mathrm{b,n}}(\boldsymbol{m})$ can be computed this way by utilizing the dual saddle point system. Finally, one updates $\boldsymbol{m}$ to the last iterate $\boldsymbol{x}_k$.

In this discrete setting the BERGN algorithm (5.2) can be computed analogously as above. The update direction is computed using (6.8) with $\alpha = \alpha_k$ in each step k. The evaluation of the Riemannian exponential map (5.2c) is simulated by the projection on the constraint manifold, which can be computed as above by (6.10).

6.2 Numerical simulations

In this section we demonstrate the benefits of our geometrical approach in numerical experiments for inverse obstacle scattering problems introduced in Section 4.5 and 5.5. The forward scattering problems were solved by a boundary integral equation method using a Nyström method with n points as described in [9, Sec. 3.6]. To this end we interpolated the polygonal curve approximations described in Section 6.1 trigonometrically. Both the evaluation of discrete forward operator F_n and the evaluation of its Jacobian DF_n as described e.g. in [25] require $O(n^3)$ flops.

We always use 8 equidistant incident plane waves and $n = 150$ points for the reconstructed curves; the far field pattern is measured at 16 equidistant measurement directions. The wavelength is chosen of the same order of magnitude as the diameter

[3] Notice that the Newton-type method (6.10) for underdetermined systems would correspond to a nearest point projection if the constraint were linear.

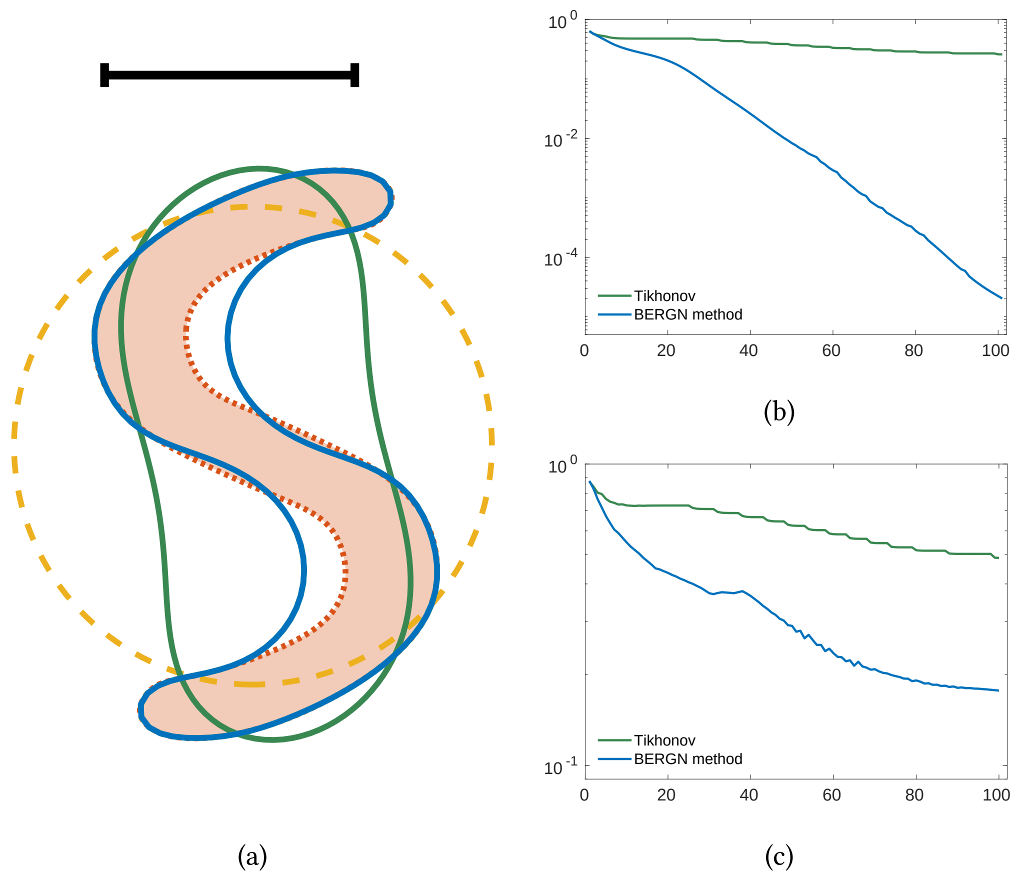


Figure 6.1: Reconstruction of a smooth nonstar-shaped domain (a) by Tikhonov regularization (dark green line) and the BERGN method (blue line) with exact data after 100 iterations. The light green plot (nearly identical to the blue one) indicates the reconstruction of Tikhonov regularization after 100 times decreasing α in 537 iterations. We use 8 equidistant incident waves, where the half of a wavelength is illustrated by the black scale bar. Red dotted lines indicate the exact solution, blue and green solid lines the reconstructions and yellow dashed lines initial guesses. In Panel (b) the error $\|F_n(\boldsymbol{m}_k) - \boldsymbol{y}^\dagger\|_{\mathbb{Y}_n}$ in the observation space (y-axis) is illustrated in the iterations k (x-axis). In Panel (c) the error in the solution space: on the y-axis the Hausdorff error $dist_{\mathrm{H}}(\boldsymbol{m}_k, \boldsymbol{m}^\dagger)$ over the iterations k (x-axis).

of the obstacle highlighted in every plot by a black scale bar, which highlights always the half of the wavelength.

For the following experiments we use simulated data, which is produced by evaluation of the forward operator described above. To make our numerical examples reliable we use a different discretization for the integral equation and therefore prevent an inverse crime.

The computation with Tikhonov regularization we do as follows. We first minimize the Tikhonov functional for a large α by an intrinsic Gauss-Newton-type method as described in Section 6.1 with update direction $\boldsymbol{u}$ defined by (6.6), (6.8), (6.9). The Gauss-Newton iteration is stopped when $\|\boldsymbol{u}\|$ or the norm of the gradient of the

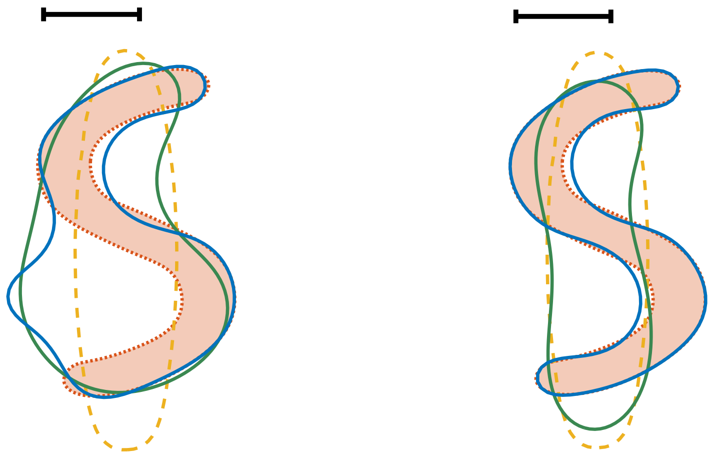


Figure 6.2: Reconstruction of a smooth nonstar-shaped domain by Tikhonov regularization (a) and the BERGN method (b) with 1% Gaussian white noise. We use 8 equidistant incident waves, where the half of a wavelength is illustrated by the black scale bar. Red dotted lines indicate the exact solution and yellow dashed lines initial guesses. The green line indicates in (a) the local minimum of $\mathcal{J}^{\alpha_0}(\boldsymbol{m})$ and the blue line the minimum of $\mathcal{J}^{\alpha_{10}}(\boldsymbol{m})$. Note that independent of how long we continue this iteration the lower concave part of the obstacle is not reconstructible. In (b) analogously the green solid line illustrates $\boldsymbol{m}_1$ and the blue line the reconstruction $\boldsymbol{m}_{10}$.

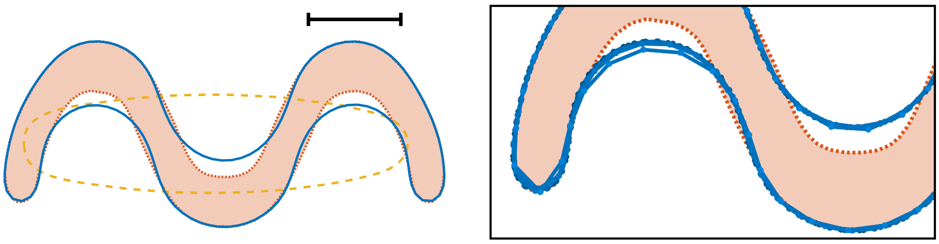
Tikhonov functional $\|D\mathcal{J}^{\alpha}(\boldsymbol{m})\|$ is smaller than 10^{-5}. Then we decrease α by a factor of $2/3$ and minimize the Tikhonov functional for this smaller α using the previous minimizer as an initial guess.

Moreover, the BERGN method is computed as described in the last section by using $\alpha_k = \alpha_0 \cdot (2/3)^k$ (if not stated otherwise). In Figure 6.1 we demonstrate the behavior of the algorithm with exact data. We can achieve a reasonable good reconstruction of a nonstar-shape domain by both algorithms, but the BERGN method is much faster than iteratively minimizing the Tikhonov functional. Note that (b) illustrates the good reconstruction in the observation space and from (c) one sees that the error in the solution space decreases over the iterations. Here the distance between two discrete curves is measured by the Hausdorff distance of point clouds.

In the rest of our examples we add independent, identically distributed, centered Gaussian random variables to the simulated far field data at each sampling point such that the relative noise level in the l^2-norm is 1%. Only in Figure 6.7 we use the higher noise levels 10% and 20%.

For the Tikhonov regularization the regularization parameter α was determined by the discrepancy principle. More precisely we iterate as described above as long as the condition $\|F_n(\boldsymbol{m}_\alpha) - \boldsymbol{y}^\delta\| \geq \tau\delta$ was satisfied. In most of the experiments τ is chosen to be 1.1.

In Section 4.5 it is already mentioned that purely minimizing the Tikhonov functional has the drawback that one can get stuck in a local minima. This is



(a) (b)

Figure 6.3: Reconstruction of a smooth non-star-shaped domain by BERGN method with 1% Gaussian white noise. Red dotted lines indicate the exact solution, blue solid lines the reconstructions and yellow dashed lines initial guesses. The half of the used wavelength is illustrated by the black scale bar. Panel (b) shows a magnification of reconstructions for different numbers of points ($n = 60$, 80, 100, 150, and 200) illustrating the asymptotic independence of the results on the choice of n.

illustrated in Figure 6.2. Using the wavelength and initial guess in Panel (a) by finding the minimum of $\mathcal{J}^\alpha$ for a fixed α one cannot compute the lower concave part. On the other hand in Panel (b) of the same figure the reconstruction by the BERGN method under the same parameters works out.

As a last part of the comparison between Tikhonov regularization and the BERGN method we point out the large number of iterations one needs for iteratively minimizing the Tikhonov functional compared to BERGN. In Table 6.1 the number of iterations needed by these algorithms is shown stopped by the discrepancy principle above. The BERGN method is clearly faster than the other one where the quality of the results are of the same order.

In Figure 6.3 we show a reconstruction of another non-star-shaped domain. Moreover, we demonstrate in (b) that the reconstructions are almost independent of the choice of the number n of points on the curves as long as n is large enough. Also the number of Gauß-Newton steps and the regularization parameter α determined by the discrepancy principle do not depend on n. Note that concave parts of the boundary where multiple reflections occur in a geometrical optics approximation are more difficult to reconstruct than convex parts.

In view of the fact that we use only one wave length which is almost of the size of the obstacle and a noise level of 1%, these reconstructions for this exponentially ill-posed problem are already remarkably good. The reconstructions could be further improved by using shorter wave lengths as illustrated in Figure 6.9 (b).

Figure 2.2 already illustrated the obvious limitation of the commonly used radial function parameterizations to star-shaped domains. In Figure 2.1 we demonstrated a further disadvantage of such parameterizations, which is the dependence on the choice of the center point. In comparison to this the BERGN method is independent of the parameterization of the curves. Indeed this is illustrated in the Figures 6.4 and 6.5. In both figures we computed for two different obstacles the reconstruction from respectively the same initial curves, but we used different parameterizations. Through all examples we gain the same quality of reconstructions.

In the theoretical investigation of the algorithm in Chapter 5 all statements are only locally and the assumption of starting close enough to the exact shape is elementary for the convergence analysis. Nevertheless, in Figure 6.6 it is shown that to some extent one can get the same quality of reconstruction for different initial guesses. Of course this will not hold true for arbitrary initial guesses.

Up to now all computations are done either without noise or with a relative noise level of 1%. It is reasonable to show the abilities of the BERGN method under idealized conditions. To be slightly more realistic in Figure 6.7 we demonstrate our algorithm performing with 10% respectively 20% relative noise. For illustration of the effect of the noise to the reconstruction we computed for each noise level and different domain the algorithm ten times and plot in all cases the best and the worst reconstruction.

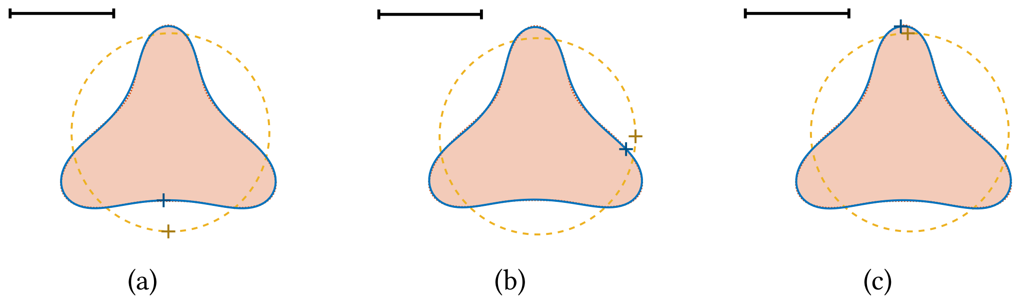


Figure 6.4: Comparison of reconstructions of a star-shaped domain by BERGN method with 1% Gaussian white noise using the same initial curve, but with different parametrizations. Parameters, line styles and colors are chosen as in Figure 6.3. The colored cross indicates the basepoint of the corresponding parametrization in the sense of (3.2).

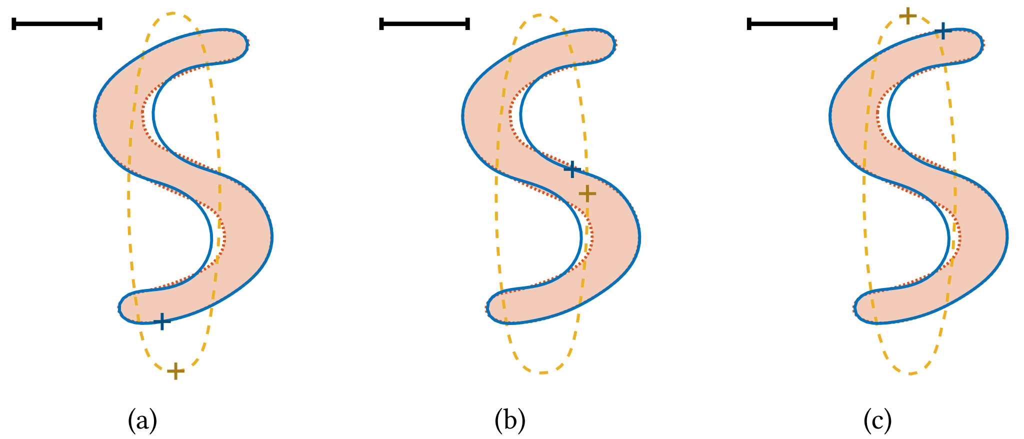


Figure 6.5: Comparison of reconstructions of a nonstar-shaped domain by BERGN method with 1% Gaussian white noise using the same initial curve, but with different parametrizations. Parameters, line styles and colors are chosen as in Figure 6.4.

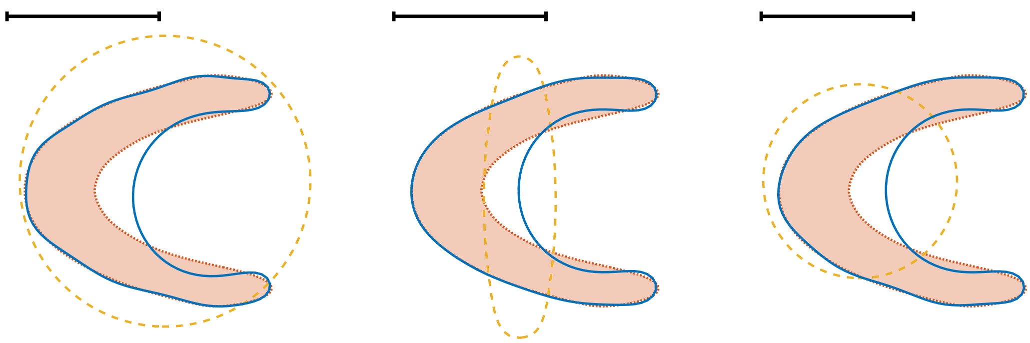
	Tikhonov regularization		BERGN method
	iteration	α_k decreased	iteration
letter C	479	22	28
letter S	271	25	25
letter M	324	18	20

Table 6.1: Comparison of number of iterations needed for reconstructing by iteratively minimizing the Tikhonov functional and BERGN method with 1% Gaussian white noise for different nonstar-shape obstacles. The reconstructed obstacles are plotted in the Figures 6.1, 6.3 and 6.6. In the first and the last column the number of iterations steps needed to reach the stopping criteria is written down. For Tikhonov regularization every step is counted independent if ones decreases α or not. The second row shows how often α is decreased while reconstructing.

One can see a significant difference to the quality of the reconstructions from less perturbed data, but by using such high noise levels the computations are exceptionally good.

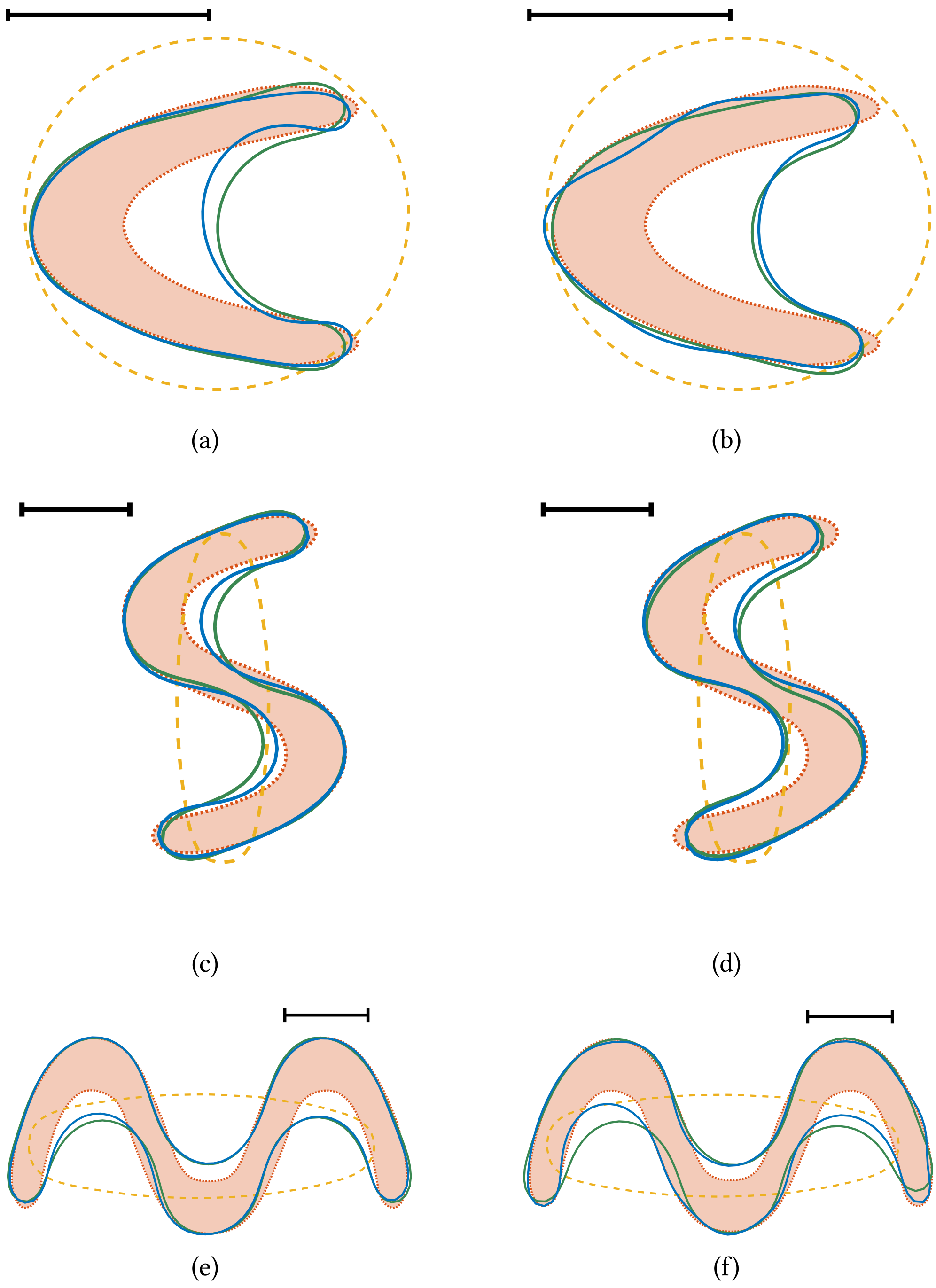
In the Figures 6.8 and 6.9 further reconstructions of more complicated domains are shown. Even for these example the BERGN method is able to calculate very precise and detailed reconstructions.

We summarize that the proposed approach for solving inverse obstacle problems on a shape manifold with bending energy penalization provides considerably better reconstructions than radial function parameterizations even for star-shaped obstacles and allows the reconstruction of more complicated curves.



(a) (b) (c)

Figure 6.6: Comparison of reconstructions of a nonstar-shaped domain by BERGN method with 1% Gaussian white noise using different initial guesses. Parameters, line styles and colors are chosen as in Figure 6.3.



(a) (b) (c) (d) (e) (f)

Figure 6.7: Comparison of reconstructions of different nonstar-shaped domains by BERGN method with higher noise levels. Parameters, line styles and colors are chosen as in Figure 6.3. In the Panels (a), (c) and (e) we used 10% and in (b), (d) and (f) 20% Gaussian white noise. Each reconstruction is simulated ten times and the blue line indicates the best reconstruction of them and the green the worst one out of the ten times.

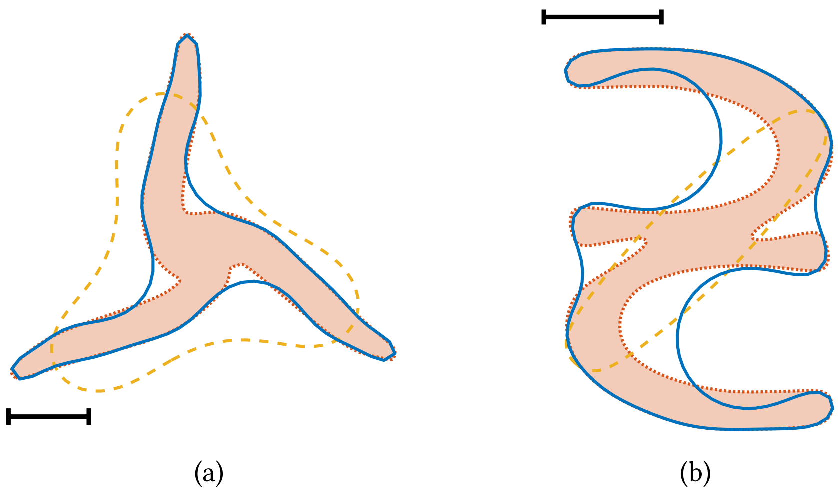


Figure 6.8: Reconstructions of nonstar-shaped domains by BERGN method with 1% Gaussian white noise. Parameters, line styles and colors are chosen as in Figure 6.3. For the reconstruction in Panel (b) we used $\alpha_k = 0.01 \cdot (5/6)^k$.

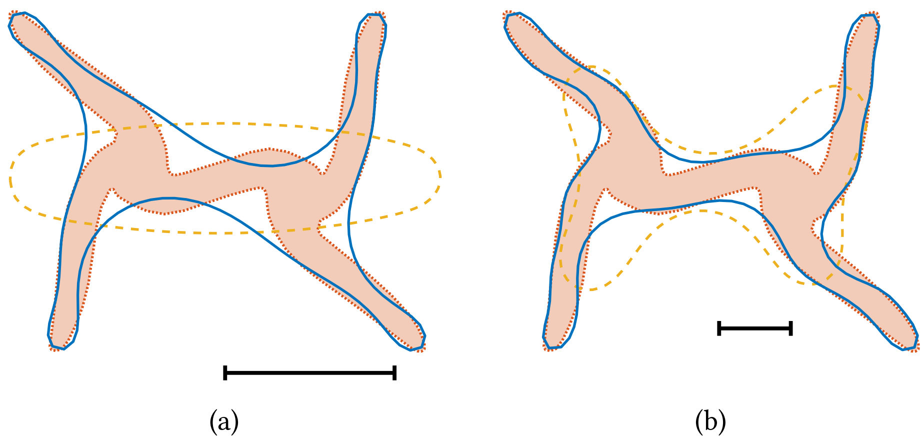


Figure 6.9: Reconstructions of a nonstar-shaped domain by BERGN method with 1% Gaussian white noise but different initial guesses and wavelengths. Parameters, line styles and colors are chosen as in Figure 6.3. The reconstruction in Panel (b) used $\alpha_k = 0.0001 \cdot (.97)^k$ and stopped with $\tau = 9$.

A Tools from infinite dimensional Riemannian geometry

There is no royal road to geometry.

— Euclid of Alexandria

In the following appendix we give a brief introduction to the essential basics and tools from Riemannian geometry used in this thesis. For more details we refer to the monographs [10, 36].

A.1 Riemannian manifolds

Let $(\{U_i\}_{i\in\mathcal{I}}, \{\varphi_i\}_{i\in\mathcal{I}})$ be a C^p*-atlas* on $\mathcal{M}$ over some index set $\mathcal{I}$, where $\{U_i\}_{i\in\mathcal{I}}$ forms a cover of $\mathcal{M}$ and φ_i maps U_i bijectively onto an open subset $\varphi_i(U_i)$ of some Banach space $\mathbb{X}$ for all $i \in \mathcal{I}$. Further $\varphi_i(U_i \cap U_j)$ is open in $\mathbb{X}$ and $\varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is a C^p-isomorphism for any $i, j \in \mathcal{I}$. Each pair (U_i, φ_i) is called a *chart*.

A set $\mathcal{M}$ with a C^p-atlas $(\{U_i\}_{i\in\mathcal{I}}, \{\varphi_i\}_{i\in\mathcal{I}})$ is called a C^p*-manifold*. Moreover the manifold $\mathcal{M}$ is said to be modeled on $\mathbb{X}$. This structure describes the set $\mathcal{M}$ locally since a chart U_i inherits the properties of the model space $\mathbb{X}$.

The notion of differentiability in $\mathcal{M}$ carries over from finite dimensional geometry. A path $\gamma\colon (-\epsilon, \epsilon) \to \mathcal{M}$ with $\gamma(0) = x$ is called differentiable if it is differentiable in the local structure, i.e. if $\varphi \circ \gamma\colon (-\epsilon, \epsilon) \to \varphi(U)$ is differentiable where (U, φ) is a chart with $x \in U$. The element $(\varphi \circ \gamma)'(0)$ in $\mathbb{X}$ is called a tangent vector and the space of all such tangent vectors at x forms the tangent space $\mathcal{T}_x\mathcal{M}$. The tangent bundle $\mathcal{T}\mathcal{M}$ is the set of pairs (x, v) with $x \in \mathcal{M}$ and $v \in \mathcal{T}_x\mathcal{M}$.

A function $f\colon \mathcal{M} \to \mathcal{N}$ between manifolds is called differentiable if the composition with the corresponding charts is differentiable. Therefore one can introduce the derivative of f by the linear map $\mathrm{D}f(x)\colon \mathcal{T}_x\mathcal{M} \to \mathcal{T}_{f(x)}\mathcal{N}$ between the corresponding tangent spaces.

One standard way for proving that a set has a manifold structure is characterizing it by a submersion. A function $f\colon \mathcal{M} \to \mathcal{N}$ is called a *submersion* at $x \in \mathcal{M}$ if the derivative $\mathrm{D}f(x)$ is surjective and its kernel splits, which means that one can decompose $\mathcal{T}_x\mathcal{M}$ into the splitting $\ker(\mathrm{D}f(x)) \times W$ for some subspace W or in other words $\mathrm{D}f(x)$ restricted to W is an isomorphism. In this case for all $y \in f(\mathcal{M})$ the set $\mathcal{U} := f^{-1}(y)$ is a submanifold in $\mathcal{M}$ and the tangent space equals the kernel of $\mathrm{D}f(x)$, i.e. $\mathcal{T}_x\mathcal{U} = \ker(\mathrm{D}f(x))$ (see [36, II.2] and [59, Thm. 73.C]).

In the infinite dimensional setting a manifold is called *Riemannian* if it is modeled over a Hilbert space $\mathbb{X}$ and the topology on the tangent space $\mathcal{T}_x\mathcal{M}$ is induced by a nonsingular symmetric bilinear positive definite inner product $g_x(\cdot,\cdot)$ or $\langle\cdot,\cdot\rangle_x$ for all $x \in \mathcal{M}$. Such a family of inner products is called the *Riemannian metric*. The corresponding norms are denoted by $\|\cdot\|_x$.

The major advantage of having an inner product is the ability of measure distances and angles. The length of a curve $\gamma\colon [0,1] \to \mathcal{M}$ on a Riemannian manifold is given by

$$L(\gamma) := \int_0^1 \sqrt{\langle\gamma'(t), \gamma'(t)\rangle_{\gamma(t)}}\,\mathrm{d}t = \int_0^1 \|\gamma'(t)\|_{\gamma(t)}\,\mathrm{d}t. \tag{A.1}$$

This endows $\mathcal{M}$ with a distance function $dist^{\mathcal{M}}\colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ via

$$dist^{\mathcal{M}}(x,y) = \inf\left\{ L(\gamma) \;\middle|\; \gamma \text{ is a path with } \gamma(0) = x \text{ and } \gamma(1) = y \right\}. \tag{A.2}$$

It is a well-known fact that $\mathcal{M}$ with this distance function becomes a metric space.

For a mapping f as above if $\mathcal{N} = \mathbb{Y}$ is itself a Hilbert space, one can give a definition of Fréchet differentiability using curves in $\mathcal{M}$ as follows. A function $f\colon \mathcal{M} \to \mathbb{Y}$ is called Fréchet differentiable at x if there exists a linear operator $\mathrm{D}f(x)\colon \mathcal{T}_x\mathcal{M} \to \mathbb{Y}$ such that

$$\lim_{t\to 0} \sup_{\|v\|_x = 1} \frac{\|f(\gamma_v(t)) - f(\gamma_v(0)) - \mathrm{D}f(x)\,(\gamma_v'(0))\|_{\mathbb{Y}}}{t} = 0 \tag{A.3}$$

where the paths $\gamma_v\colon (-\epsilon, \epsilon) \to \mathcal{M}$ satisfy $\gamma_v(0) = x$ and $\gamma_v'(0) = v$.

A.2 Vector fields and covariant derivatives

One key benefit of Riemannian geometry is that one can define a gradient. For a real valued differentiable function f on $\mathcal{M}$ the derivative $\mathrm{D}f(x)\colon \mathcal{T}_x\mathcal{M} \to \mathbb{R}$ is a bounded linear map on a Hilbert space. Consequently by the Riesz theorem there is an element $\mathit{grad}\, f(x) \in \mathcal{T}_x\mathcal{M}$ such that $\mathrm{D}f(x)\,v = \langle \mathit{grad}\, f(x), v\rangle_x$ for all $v \in \mathcal{T}_x\mathcal{M}$. The *gradient field* $\mathit{grad}\, f\colon \mathcal{M} \to \mathcal{T}\mathcal{M}$, $x \mapsto \mathit{grad}\, f(x)$ is one example for a vector field on $\mathcal{M}$. In general a vector field is a map of the form $V\colon \mathcal{M} \to \mathcal{T}\mathcal{M}$ and for convenience write $V_x \in \mathcal{T}_x\mathcal{M}$ for $V(x)$. Let $\Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$ denote the vector space of all smooth vector fields on $\mathcal{M}$ into the tangent bundle $\mathcal{T}\mathcal{M}$.

Furthermore, one can define a derivative of a vector field into a direction of a vector field. A mapping $\nabla\colon \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})\times\Gamma(\mathcal{M};\mathcal{T}\mathcal{M}) \to \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$, $(W,V) \mapsto \nabla_W V$ is called an *affine connection* if

1. $\nabla_{f\,W+g\,U}V = f\,\nabla_W V + g\,\nabla_U V$
2. $\nabla_W(a\,V + b\,U) = a\,\nabla_W V + b\,\nabla_W U$
3. (Leibniz' law) $\nabla_W(f\,V) = (W\,f)\,V + f\,\nabla_W V$,

where $V, W, U \in \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$, $f, g\colon \mathcal{M} \to \mathbb{R}$ and $a, b \in \mathbb{R}$. The expression Wf is a short notation for the real valued function $(Wf)(x) := W_x(f) = \langle \mathit{grad}\, f(x), W_x\rangle$ on $\mathcal{M}$. The vector field $\nabla_W V$ is called the *covariant derivative* of V with respect to W for the affine connection ∇. For a real valued function f and vector fields W, V define the vector field $[W,V]$ by evaluation $[W,V]\,f := W\,(V\,f) - V\,(W\,f)$. It can be shown that this defines a derivation of real valued functions, i.e. it satisfies the Leibniz' Law $[W,V]\,(f\,g) = f\,([W,V]\,g) + ([W,V]\,f)\,g$ for f and g real valued functions on $\mathcal{M}$ (see [36, Prop. V.1.3]).

Theorem A.1 ([36, Thm. VIII.4.1])**.** *On a Riemannian manifold $\mathcal{M}$ there exists a unique affine connection ∇ that satisfies*

1. $\nabla_W V - \nabla_V W = [W,V]$ *(symmetry)*
2. $U\langle W,V\rangle = \langle \nabla_U W, V\rangle + \langle W, \nabla_U V\rangle$ *(compatibility with the Riemannian metric),*

where $V, W, U \in \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$. This affine connection, called Levi-Civita connection *or* Riemannian connection*, is characterized by the Koszul formula*

$$
\begin{aligned}
2\langle \nabla_U W, V\rangle = U\langle W,V\rangle + W\langle V,U\rangle - V\langle U,W\rangle \\
- \langle U,[W,V]\rangle + \langle W,[V,U]\rangle + \langle V,[U,W]\rangle.
\end{aligned}
$$

For a real valued smooth function f on $\mathcal{M}$ the *Riemannian Hessian* of f at x in $\mathcal{M}$ as the linear mapping $\mathcal{H}ess\, f(x)\colon \mathcal{T}_x\mathcal{M} \to \mathcal{T}_x\mathcal{M}$ by

$$\mathcal{H}ess\, f(x)\,(v) := \nabla_v grad\, f(x) \tag{A.4}$$

for all $v \in \mathcal{T}_x\mathcal{M}$, where ∇ is the Riemannian connection on $\mathcal{M}$. For the sake of simplicity of notation in this thesis this operator is identified with the corresponding bilinear form $\mathcal{H}ess\, f(x)\colon \mathcal{T}_x\mathcal{M} \times \mathcal{T}_x\mathcal{M} \to \mathbb{R}$ given by

$$\mathcal{H}ess\, f(x)\,(v, w) = \langle \nabla_v grad\, f(x), w\rangle_x$$

for all $x \in \mathcal{M}$ and $v, w \in \mathcal{T}_x\mathcal{M}$.

A.3 Parallelism and the exponential map

Let $\gamma\colon [0,1] \to \mathcal{M}$ be a C^2 curve. We call a map $\beta\colon [0,1] \to \mathcal{T}\mathcal{M}$ a vector field along γ if $\beta(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}$ for all t and $V \in \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$ can be restricted to a vector field along γ by $\beta(t) := V(\gamma(t))$. A vector field V along γ is called γ-parallel if $\nabla_{\gamma'(t)}V = 0$ for all t, where $\gamma'(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}$. The definition of parallel transport along a path γ is given as follows. Let $v \in \mathcal{T}_{\gamma(0)}\mathcal{M}$ be given. There is a unique γ-parallel vector field $t \mapsto \zeta(t, v)$ along γ with $\zeta(0, v) = v$. The *parallel transport map* is defined by

$$\mathcal{P}_{\gamma(0)}^{\gamma(t)}\colon \mathcal{T}_{\gamma(0)}\mathcal{M} \to \mathcal{T}_{\gamma(t)}\mathcal{M}, \qquad \mathcal{P}_{\gamma(0)}^{\gamma(t)}(v) = \zeta(t, v). \tag{A.5}$$

The map $v \mapsto \mathcal{P}_{\gamma(0)}^{\gamma(t)}(v)$ is a linear isometric isomorphism between the tangent spaces (see [36, Thm. VIII.3.4]).

This vector transport is even connected to the covariant derivative. For $U, V \in \Gamma(\mathcal{M}, \mathcal{T}\mathcal{M})$ and a path γ with $\gamma(0) = x$ and $\gamma'(0) = U_x$ and $\mathcal{P}$ maps along γ then

$$\nabla_U V(x) = \lim_{t\to 0} \frac{\mathcal{P}_{\gamma(t)}^{x} V_{\gamma(t)} - V_x}{t} = \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{P}_{\gamma(t)}^{x} V_{\gamma(t)}\Big|_{t=0} \tag{A.6}$$

(see [10, Ch. 2, Ex.2]). In this manner, given the parallel transport of a manifold, the covariant derivative is reobtainable from it.

A curve γ on $\mathcal{M}$ is called a *geodesic* if it solves the differential equation

$$\nabla_{\gamma'}\gamma' = 0. \tag{A.7}$$

Therefore these objects have zero acceleration and in this sense they are the natural generalization of straight lines to nonlinear spaces.

Recall the construction of the exponential map which turns out to be one of the key tools in defining and analyzing algorithms on manifolds. For every $v \in \mathcal{T}_x\mathcal{M}$ there exists a unique geodesic γ, such that $\gamma(0) = x$ and $\gamma'(0) = v$. The *Riemannian exponential map* is defined by

$$exp_x \colon \mathcal{T}_x\mathcal{M} \to \mathcal{M}, \qquad v \mapsto exp_x(v) := \gamma(1) \tag{A.8}$$

(see [36, Cor. VIII.5.2]). It is a well-known fact that locally the unique geodesic given in the construction of the exponential map is also the geodesic which minimizes the Riemannian distance. That means if for $x, y \in \mathcal{M}$ there is exactly one $v \in \mathcal{T}_x\mathcal{M}$ such that $exp_x(v) = y$, then the distance between x and y is given by the norm of v:

$$dist(x, y) = dist(x, exp_x(v)) = L(\gamma) = \int_0^1 \|\gamma'(t)\|_{\gamma(t)} \,\mathrm{d}t = \|v\|_x, \tag{A.9}$$

where the last identity holds since geodesics have a vanishing second derivative and therefore a constant first derivative, which is here given by the initial direction v.

A.4 Curvature

The notion of curvature arises from investigating how the covariant derivative is influenced by the geometry of the manifold. Let V, W, $Z \in \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$ be vector fields. The operator

$$\mathcal{R}(V, W) \colon \Gamma(\mathcal{M};\mathcal{T}\mathcal{M}) \to \Gamma(\mathcal{M};\mathcal{T}\mathcal{M}), \quad \mathcal{R}(V, W)\,Z = \nabla_V\nabla_W Z - \nabla_W\nabla_V Z - \nabla_{[V,W]}Z \tag{A.10}$$

is called the *curvature tensor* and measures how curved the manifold is. If v, w, $z \in \mathcal{T}_x\mathcal{M}$ then this simplifies to

$$\mathcal{R}_x(v, w)\,z = \mathcal{R}(V(x), W(x))\,Z(x) = (\mathcal{R}(V, W)\,Z)\,(x)$$

for any vector fields V, W, Z which satisfy $V(x) = v$, $W(x) = w$ and $Z(x) = z$. The functional

$$\mathcal{R}_x(v, w, z, u) = \langle \mathcal{R}_x(v, w)\,z, u\rangle_x, \qquad v, w, z, u \in \mathcal{T}_x\mathcal{M} \tag{A.11}$$

is called the *Riemannian* 4*-tensor*. This tensor satisfies the four identities

1. $\mathcal{R}_x(v, w, z, u) = -\mathcal{R}_x(w, v, z, u)$
2. $\mathcal{R}_x(v, w, z, u) = -\mathcal{R}_x(v, w, u, z)$

3. $\mathcal{R}_x(v, w, z, u) + \mathcal{R}_x(w, z, v, u) + \mathcal{R}_x(z, v, w, u) = 0$ (Bianchi identity)
4. $\mathcal{R}_x(v, w, z, u) = \mathcal{R}_x(z, u, v, w)$.

Let $\gamma\colon [0,1] \to \mathcal{M}$ be a curve and W be a vector field along γ, i.e. $W\colon [0,1] \to \mathcal{TM}$, $W(t) = W(\gamma(t))$. Then W is called a *Jacobi field* if it satisfies the Jacobi differential equation

$$\nabla^2_{\gamma'} W = \mathcal{R}(\gamma', W)\,\gamma'. \tag{A.12}$$

If γ is a geodesic, then for any v, $w \in \mathcal{T}_{\gamma(0)}\mathcal{M}$ there exists a unique Jacobi field $W = W_{v,w}$ such that $W(0) = v$ and $\nabla_{\gamma'} W(0) = w$ (see [36, Thm. IX.2.1]).

Let $v \in \mathcal{T}_x\mathcal{M}$ be given and consider the construction above using the geodesic $\gamma(t) = \mathit{exp}_x(tv)$. Then $\gamma(0) = x$ and $\gamma'(0) = v$ by definition. If $w \in \mathcal{T}_x\mathcal{M}$ is an additional tangent, denote W_w as the Jacobi field with $W_w(0) = 0$ and $\nabla_{\gamma'} W_w(0) = w$. In the special case $w = v$ it follows immediately that

$$W_v(t) = t\gamma'(t) = t\,\mathcal{P}_{\gamma(0)}^{\gamma(t)} v, \tag{A.13}$$

since $\gamma'(t)$ transports v parallel along γ. Therefore the Jacobi field corresponds to the parallel transport in at least one direction.

For a given $v \in \mathcal{T}_x\mathcal{M}$ every element $w \in \mathcal{T}_x\mathcal{M}$ can be decomposed orthogonally by $w = cv + \tilde{w}$ with $\langle \tilde{w}, v\rangle_x = 0$. From this one obtains the decomposition

$$W_w(t) = ct\gamma'(t) + W_{\tilde{w}}(t), \tag{A.14}$$

where $\langle W_{\tilde{w}}(t), \gamma'(t)\rangle_{\gamma(t)} = 0$ for all t.

By keeping the notation above the Jacobi field is equivalent to the derivative of the exponential map

$$\mathrm{D}\mathit{exp}_x(tv)\,(w) = \frac{1}{t}\,W_w(t) \tag{A.15}$$

for $t > 0$ (see [36, Thm. IX.3.1]). For $y = \mathit{exp}_x(v)$ the operator $\mathrm{D}\mathit{exp}_x(v)$ maps from $\mathcal{T}_x\mathcal{M}$ to $\mathcal{T}_y\mathcal{M}$. Thus a notion of an adjoint operator is applicable and its characterization is as follows. Let $z \in \mathcal{T}_y\mathcal{M}$, $w \in \mathcal{T}_x\mathcal{M}$ and $v^* = -\gamma'(1) = -\mathcal{P}_x^y v$, where $\mathcal{P}_x^y$ maps parallel along γ. Then the equation

$$\langle \mathrm{D}\mathit{exp}_x(v)\,(w), z\rangle_y = \langle w, \mathrm{D}\mathit{exp}_y(v^*)\,(z)\rangle_x \tag{A.16}$$

holds true (see [36, Lem. IX.3.5]).

A Taylor expansion of the Jacobi field can be given as follows. The proof can be found in [36, Prop. IX.5.3]. For $\gamma(t) = \mathit{exp}_x(tv)$ and $w \in \mathcal{T}_x\mathcal{M}$ denote again W_w as the corresponding Jacobi lift. Then

$$W_w(t) = \mathcal{P}_x^{\gamma(t)}\Big(wt + \mathcal{R}_x(v, w)v\frac{t^3}{3!}\Big) + O(t^4) \qquad t \to 0. \tag{A.17}$$

The Jacobi field and the parallel transport can map between tangent spaces, but they do it in different ways. The difference between these two different vector transport mappings can be controlled by the curvature tensor.

Lemma A.2. *For all $v, w \in \mathcal{T}_x\mathcal{M}$ and $\gamma(t) = exp_x(tv)$ the inequality*

$$\left\| \mathrm{D}exp_x(v)\,(w) - \mathcal{P}_x^{exp_x(v)} w \right\|_{exp_x(v)} \leq \frac{1}{2}\, \|\mathcal{R}_x(v, w)\, v\|_x \tag{A.18}$$

holds true.

Proof. By using $\mathrm{D}exp_x(v)\,(w)$ is equal to the Jacobi field W_w on the curve γ and applying the defining differential equation (A.12) it is

$$\begin{aligned}
\left\|W_w(1) - W_w(0) - \mathcal{P}_{\gamma(0)}^{\gamma(1)} w\right\|_{\gamma(1)} &\leq \int_0^1 \left\|\mathcal{P}_{\gamma(t)}^{\gamma(1)} \nabla_{\gamma'} W_w(t) - \mathcal{P}_{\gamma(0)}^{\gamma(1)} \nabla_{\gamma'} W_w(0)\right\|_{\gamma(1)} \mathrm{d}t \\
&\leq \int_0^1 \int_0^t \left\|\mathcal{P}_{\gamma(s)}^{\gamma(1)} \nabla_{\gamma'}^2 W_w(s)\right\|_{\gamma(1)} \mathrm{d}s\, \mathrm{d}t \\
&\leq \frac{1}{2} \|\mathcal{R}_x(v, w) v\|_x
\end{aligned}$$

which proves the statement. □

A.5 Covariant derivatives on submanifolds

Let $\mathbb{X}$ be Hilbert space and $\iota \colon \mathcal{M} \to \mathbb{X}$ be a submanifold. For $x \in \mathcal{M}$ let $P_x \colon \mathbb{X} \to \mathcal{T}_x\mathcal{M}$ be the orthogonal projection. Then the Levi-Civita connection ∇ on $\mathcal{M}$ is given by

$$\nabla_W V(x) = P_x\big(\mathrm{D}V(x)\,(W(x))\big) \tag{A.19}$$

for $V, W \in \Gamma(\mathbb{X}; \mathcal{T}\mathbb{X})$, where $\mathcal{T}\mathbb{X} = \mathbb{X} \times \mathbb{X}$. Denote the normal bundle $\mathcal{N}\mathcal{M}$ of $\mathcal{M}$ component wise via $\mathcal{N}_x\mathcal{M} = \mathcal{T}_x\mathcal{M}^\perp$.

The *second fundamental form* $\mathbb{I}$ is a symmetric bilinear map $\mathbb{I} \colon \mathcal{T}\mathcal{M} \times \mathcal{T}\mathcal{M} \to \mathcal{N}\mathcal{M}$ given by

$$\mathbb{I}(V, W)\,(x) = \mathbb{I}(V(x), W(x)) = (I - P_x)\,\big(\mathrm{D}V(x)\,(W(x))\big). \tag{A.20}$$

For $v, w \in \mathcal{T}_x\mathcal{M}$ such that $V(x) = x$ and $W(x)$ one can write the fundamental form tensorial by

$$\mathbb{I}_x(v, w) = \mathbb{I}(V, W)\,(x).$$

The *Weingarten equation* is the combination of (A.19) and (A.20) given by

$$\nabla_W V(x) = \mathrm{D}V(x)\,(W(x)) + \mathbb{I}(V, W)\,(x). \tag{A.21}$$

Closely related to this is the so-called *Weingarten map* $\mathcal{S}$, which is a bilinear map $\mathcal{S}\colon \mathcal{N}\mathcal{M}\times\mathcal{T}\mathcal{M} \to \mathcal{T}\mathcal{M}$ with

$$\mathcal{S}_N(V)\,(x) = \mathcal{S}_{N(x)}(V(x)) = P_x\big(\mathrm{D}N(x)\,(V(x))\big) \tag{A.22}$$

with $N \in \Gamma(\mathcal{M};\mathcal{N}\mathcal{M})$ and $V \in \Gamma(\mathcal{M};\mathcal{T}\mathcal{M})$ or respectively for vectors $n \in \mathcal{N}_x\mathcal{M}$ and $v \in \mathcal{T}_x\mathcal{M}$ such that $N(x) = n$ and $V(x) = x$ one writes it tensorial by $\mathcal{S}_n(v) = \mathcal{S}_N(V)(x)$ (see [36, Thm. XIV.1.1]). These two maps are related through the equation

$$\langle \mathcal{S}_n(v), w\rangle_{\mathbb{X}} = \langle n, \mathrm{I\!I}_x(v, w)\rangle_{\mathbb{X}} \tag{A.23}$$

for all $n \in \mathcal{N}_x\mathcal{M}$ and $v, w \in \mathcal{T}_x\mathcal{M}$.

The operator $\mathrm{I\!I}$ measures the curvature of the manifold from an outside or *extrinsic* point of view and there is a well-known connection to the interior or *intrinsic* perspective: The so-called *Gauss-equation* connects $\mathrm{I\!I}$ with the curvature tensor $\mathcal{R}$ by

$$\langle \mathcal{R}_x(v, w)\,z, u\rangle_x = \langle \mathrm{I\!I}_x(v, u), \mathrm{I\!I}_x(w, z)\rangle_x - \langle \mathrm{I\!I}_x(v, z), \mathrm{I\!I}_x(w, u)\rangle_x \tag{A.24}$$

for all $u, v, w, z \in \mathcal{T}_x\mathcal{M}$ (see [36, Thm. XIV.5.1]).

Let $f\colon \mathbb{X} \to \mathbb{Y}$ be a smooth function with values in another Hilbert space $\mathbb{Y}$. Then the Hessian of f at $x \in \mathcal{M}$ is given by

$$\mathcal{H}ess\, f(x)\,(u, v) = \mathrm{D}^2 f(x)\,(u, v) + \mathrm{D}f(x)\;\mathrm{I\!I}_x(u, v) \tag{A.25}$$

for all $u, v \in \mathcal{T}_x\mathcal{M}$ (see [36, Prop. XIV.2.1]).

In the following lemma the second fundamental form is characterized in the case that $\mathcal{M}$ is given by a submersion.

Lemma A.3. *For an open neighborhood $U \subset \mathbb{X}$ of $\mathcal{M}$ and a submersion $\Phi\colon U \to \mathbb{Y}$ such that $\mathcal{M} = \Phi^{-1}(0)$. Then*

$$\mathrm{I\!I}_x(u, v) = -\mathrm{D}\Phi(x)^\dagger\, \mathrm{D}^2\Phi(x)\,(u, v) \tag{A.26}$$

holds true for all $u, v \in \mathcal{T}_x\mathcal{M}$. Here $\mathrm{D}\Phi(x)^\dagger$ denotes the Moore-Penrose pseudoinverse.

Proof. By using (A.25) with Φ one gets

$$\mathcal{H}ess\,\Phi(x)\,(u, v) = \mathrm{D}^2\Phi(x)\,(u, v) + \mathrm{D}\Phi(x)\;\mathrm{I\!I}_x(u, v).$$

Since $M = \Phi^{-1}(0)$ the map Φ restricted to $\mathcal{M}$ is constant and therefore $\mathcal{H}ess\,\Phi(x) = 0$ for $x \in \mathcal{M}$ and obtain the equation

$$\mathrm{D}\Phi(x)\;\mathrm{I\!I}_x(u, v) = -\mathrm{D}^2\Phi(x)\,(u, v).$$

Recall that for a submersion the tangent space is given by the kernel of $\mathrm{D}\Phi$, i.e. $\mathcal{T}_x\mathcal{M} = \ker(\mathrm{D}\Phi(x))$. From $\mathbb{I}_x(u,v) \in \mathcal{N}_x\mathcal{M} = \ker(\mathrm{D}\Phi(x))^\perp$ and the fact that the Moore-Penrose inverse is an isomorphism from $\ker(\mathrm{D}\Phi(x))^\perp$ onto $\mathbb{Y}$ one gets the statement. □

Under the assumptions of Lemma A.3 for a smooth real valued function F on M one obtains the expression

$$\mathcal{H}ess\, f(x)\,(u,v) = \mathrm{D}^2 f(x)\,(u,v) - \mathrm{D}f(x)\,\mathrm{D}\Phi(x)^\dagger\,\mathrm{D}^2\Phi(x)\,(u,v) \tag{A.27}$$

for the intrinsic Hessian of f at $x \in \mathcal{M}$ in terms of extrinsic operations in the Hilbert space $\mathbb{X}$.

A.6 Optimization on manifolds

Let $\mathcal{M}$ be a Riemannian manifold and f a smooth positive real valued function. The problem of finding a minimum of

$$\min_{x\in\mathcal{M}} f(x) \tag{A.28}$$

looks similar to known smooth optimization problems in Hilbert spaces and even the principle ideas for solving it carry over to curved spaces. Assume for a moment that $\mathcal{M} = \mathbb{X}$ is itself a Hilbert space and recall that the most common algorithms can be written as follows: Starting at some point x_0 in each iteration one chooses an update direction $v_k \in \mathbb{X}$ and computes the next iterate by $x_{k+1} = x_k + t_k\, v_k$ with an appropriate step size $t_k > 0$.

Back to the case of $\mathcal{M}$ a general Riemannian manifold. The concept above can be formulated also in the language of differential geometry. Let $x_0 \in \mathcal{M}$ be given. In each iteration $k = 1, 2, 3, \dots$ one chooses an update direction

$$v_k \in \mathcal{T}_{x_k}\mathcal{M} \tag{A.29}$$

and compute the next iterate via

$$x_{k+1} = exp_{x_k}(t_k\, v_k). \tag{A.30}$$

Of course there are a lot of possibilities how to choose the update direction in a suitable way depending on the problem. One typical example would be to take it as the solution of a Newton equation

$$\mathcal{H}ess\, f(x_k)\,(v,v) = -\langle grad\, f(x_k), v\rangle_{x_k} \tag{A.31}$$

with $v \in \mathcal{T}_{x_k}\mathcal{M}$. There is a large amount of literature dealing with such kind of optimization algorithms. For example the monograph [1] gives an extensive overview on different types of algorithms on finite dimensional manifolds, including proofs of convergence and possible modifications. In the paper [56] some Newton-type algorithms are also presented on infinite dimensional Riemannian manifolds applied to shape optimization.

One important new feature in the theoretical convergence investigations of algorithms on Riemannian manifolds is the incorporation of the curvature. Especially on infinite dimensional manifolds most of the interesting examples do not have a uniform bound on the curvature. In the paper [51] convergence of two Newton-type algorithms on infinite dimensional Riemannian manifolds is shown under local bounds on the intrinsic curvature.

This motivates the investigation in the second fundamental form in Chapter 3 such that we can obtain local bounds on the extrinsic curvature which are used in the convergence analysis in Chapter 5.

Bibliography

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Walter de Gruyter GmbH, Jan 2008.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02.

[3] B. Audoly, N. Clauvelin, P.-T. Brun, M. Bergou, E. Grinspun, and M. Wardetzky. A discrete geometric approach for simulating the dynamics of thin viscous threads. *Journal of Computational Physics*, 253:18–49, 2013.

[4] A. B. Bakushinskii. The problem of the convergence of the iteratively regularized gauss-newton method. *Comput. Math. Math. Phys.*, 32(9):1353–1359, Sept. 1992.

[5] M. Bergou, M. Wardetzky, S. Robinson, B. Audoly, and E. Grinspun. Discrete Elastic Rods. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, 27(3):63:1–63:12, 2008.

[6] S. Blatt. Boundedness and regularizing effects of O'Hara's knot energies. *J. Knot Theory Ramifications*, 21(1):1250010, 9, 2012.

[7] S. Blatt, P. Reiter, and A. Schikorra. Harmonic analysis meets critical knots. Critical points of the Möbius energy are smooth. *Trans. Amer. Math. Soc.*, 368(9):6391–6438, 2016.

[8] F. Cakoni and D. Colton. *Qualitative methods in inverse scattering theory.* Interaction of Mechanics and Mathematics. Springer-Verlag, Berlin, 2006.

[9] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer New York, 2013.

[10] M. do Carmo. *Riemannian Geometry.* Mathematics (Boston, Mass.). Birkhäuser, 1992.

[11] M. P. do Carmo. *Differential geometry of curves and surfaces.* Prentice Hall, 1976.

[12] J. Eckhardt, R. Hiptmair, T. Hohage, H. Schumacher, and M. Wardetzky. Elastic energy regularization for inverse obstacle scattering problems. *Inverse Problems*, 35(10):104009, Sept. 2019.

[13] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications.* Kluwer Academic Publishers Group, Dordrecht, 1996.

[14] L. Euler. Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti. In *Opera Omnia*, volume 24 of *series 1.* 1744.

[15] J. Flemming. *Generalized Tikhonov regularization and modern convergence rate theory in Banach spaces.* Shaker Verlag, Aachen, 2012.

[16] M. H. Freedman, Z.-X. He, and Z. Wang. Möbius energy of knots and unknots. *Ann. of Math. (2)*, 139(1):1–50, 1994.

[17] W. B. Gordon. The Riemannian structure of certain function space manifolds. *J. Differential Geom.*, 4(4):499–508, 1970.

[18] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Problems*, 26:115014 (16pp), 2010.

[19] J. Hadamard. Lectures on cauchy's problem in linear partial differential equations. 1923.

[20] Z.-X. He. The Euler-Lagrange equation and heat flow for the Möbius energy. *Comm. Pure Appl. Math.*, 53(4):399–431, 2000.

[21] H. Hencky. *Über die angenäherte Lösung von Stabilitätsproblemen im Raum mittels der elastischen Gelenkkette.* PhD thesis, Engelmann, 1921.

[22] F. Hettlich and W. Rundell. Iterative methods for the reconstruction of an inverse potential problem. *Inverse Problems*, 12(3):251–266, jun 1996.

[23] B. Hofmann, P. Mathé, and H. von Weizsäcker. Regularization in Hilbert space under unbounded operators and general source conditions. *Inverse Problems*, 25(11):115013, 15, 2009.

[24] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss - Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13(5):1279–1299, Oct. 1997.

[25] T. Hohage. *Iterative method in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems.* PhD thesis, Universität Linz, 1999.

[26] T. Hohage. Regularization of exponentially ill-posed problems. *Numerical Functional Analysis and Optimization*, 21(3-4):439–464, Jan 2000.

[27] T. Hohage and P. Miller. Optimal convergence rates for sparsity promoting wavelet-regularization in besov spaces. *Inverse Problems*, 35(6):065005, May 2019.

[28] T. Hohage and F. Weidling. Verification of a variational source condition for acoustic inverse medium scattering problems. *Inverse Problems*, 31(7):075006, 14, 2015.

[29] T. Hohage and F. Werner. Iteratively regularized Newton-type methods for general data misfit functionals and applications to poisson data. *Numerische Mathematik*, 123(4):745–779, Oct. 2012.

[30] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems.* Walter de Gruyter, Jan. 2008.

[31] J. B. Keller. Inverse problems. *The American Mathematical Monthly*, 83(2):107, Feb. 1976.

[32] A. Kirsch and N. Grinberg. *The factorization method for inverse problems*, volume 36 of *Oxford Lecture Series in Mathematics and its Applications.* Oxford University Press, Oxford, 2008.

[33] R. Kress. *Linear Integral Equations.* Applied Mathematical Sciences. Springer, 3ed. edition, 2014.

[34] R. B. Kusner and J. M. Sullivan. Möbius energies for knots and links, surfaces and submanifolds. In *Geometric topology (Athens, GA, 1993)*, volume 2 of *AMS/IP Stud. Adv. Math.*, pages 570–604. Amer. Math. Soc., Providence, RI, 1997.

[35] R. B. Kusner and J. M. Sullivan. Möbius-invariant knot energies. In *Ideal knots*, volume 19 of *Series on Knots and Everything*, pages 315–352. World Sci. Publ., River Edge, NJ, 1998.

[36] S. Lang. *Fundamentals of Differential Geometry.* Springer New York, 1999.

[37] S. Langer and T. Hohage. Convergence analysis of an inexact iteratively regularized Gauss-Newton method under general source conditions. *J. Inverse Ill-Posed Probl.*, 15(3):311–327, 2007.

[38] P. Mathé and B. Hofmann. How general are general source conditions? *Inverse Problems*, 24(1):015009, Jan. 2008.

[39] P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable hilbert scales. *Inverse Problems*, 19(3):789–803, May 2003.

[40] W. C. McLean. *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, 2000.

[41] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*, 8(1):1–48, 2006.

[42] D. S. Mitrinović. *Analytic Inequalities*. Springer Berlin Heidelberg, 1970.

[43] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems*. Springer, New York, Berlin, Heidelberg, 1984.

[44] J. R. Munkres. *Elementary differential topology*, volume 1961 of *Lectures given at Massachusetts Institute of Technology, Fall*. Princeton University Press, Princeton, N.J., 1966.

[45] M. Z. Nashed. *Generalized Inverses and Applications*. Academic Press, 1976.

[46] A. Neubauer. On converse and saturation results for Tikhonov regularization of linear ill-posed problems. *SIAM journal on numerical analysis*, 34(2):517–527, 1997.

[47] J. O'Hara. Energy of a knot. *Topology*, 30(2):241–247, 1991.

[48] R. Potthast. *Point sources and multipoles in inverse scattering theory*, volume 427 of *Chapman & Hall/CRC Research Notes in Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2001.

[49] M. Reed and B. Simon. *Methods of modern mathematical physics, Volume 1: Functional analysis*. Academic Press, 1980.

[50] W. Ring. Identification of a core from boundary data. *SIAM Journal on Applied Mathematics*, 55, 05 2000.

[51] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2):596–627, Jan 2012.

[52] M. Rumpf and B. Wirth. Variational time discretization of geodesic calculus. *IMA J. Numer. Anal.*, 35(3):1011–1046, 2015.

[53] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences.* Springer, New York, 2009.

[54] S. Scholtes. Discrete Möbius energy. *J. Knot Theory Ramifications*, 23(9):1450045–1–1450045–16, 2014.

[55] S. Scholtes, H. Schumacher, and M. Wardetzky. Variational Convergence of Discrete Elasticae. *arXiv e-prints*, 2019. arXiv:1901.02228.

[56] V. H. Schulz. A riemannian view on shape optimization. *Foundations of Computational Mathematics*, 14(3):483–501, Jun 2014.

[57] A. N. Tikhonov. On the regularization of ill-posed problems. *Dokl. Akad. Nauk SSSR*, 153:49–52, 1963.

[58] A. N. Tikhonov. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk SSSR*, 151:501–504, 1963.

[59] E. Zeidler. *Nonlinear Functional Analysis and its Applications IV: Applications to Mathematical Physics.* Springer New York, 1988.